An Evaluation of Factors Affecting Drug Quality: Evidence from the Antimalarial Market in Uganda

Esther Atukunda and Anne Fitzpatrick

WORKING PAPER 2015-03

DEPARTMENT OF ECONOMICS

UNIVERSITY OF MASSACHUSETTS BOSTON

An Evaluation of Factors Affecting Drug Quality: Evidence from the Antimalarial Market in Uganda

Atukunda, Esther Fitzpatrick, Anne*

October 28, 2015

Abstract

The quality of healthcare, and specifically medicines, is reportedly low in developing countries. We purchase and test 879 antimalarial drugs from 459 outlets in 44 randomly selected parishes (131 villages) in Uganda to estimate the average rate of drug quality. We focus on artemether-lumefantrine (AL), the first-line treatment for malaria in Uganda. Purchased drugs are tested for quality using a handheld spectrometer. Our methodology allows us to differentiate between counterfeit and substandard drugs; counterfeit drugs are different than a high quality drug of the same brand, while substandard drugs are different and also likely medically ineffective. Data are then linked to surveys of drug vendors at the same outlets to test hypotheses of how low quality drugs arrive at market. In contrast to previous literature, we find that AL is widely available and drug quality is relatively high in the study area. While 17% of samples are counterfeit, only 3.4% of purchased drugs are substandard. We subsequently establish three new empirical facts regarding low-quality drugs. First, substandard drugs are typically dilutions of high quality doses, rather than dosages of all ineffective tablets. Dilution increases noise and makes it more difficult for customers to recognize when they have been sold a substandard dosage. Second, we show that counterfeit dugs are priced slightly lower, but substandard drugs are priced the same as high quality drugs. These results are consistent with consumer deception as opposed to a low willingness to pay for quality. Third, a small percentage of vendors are complicit in selling deceptively ineffective medicines. However, identifying which vendors and outlets sell low-quality medicines is difficult.

JEL: D8, I15, L15

*Corresponding author: Department of Economics, University of Massachusetts- Boston, 100 Morrissey Blvd., Boston, MA, 02125. Email: Anne.Fitzpatrick@umb.edu; Atukunda: Department of Pharmacology, Mbarara University of Science and Technology, P.O. Box 1410, Mbarara University Rd, Mbarara Uganda. The online appendix is available here: www.sites.google.com/site/fitzpatrickanne. We are grateful for financial support from the following sources: University of Michigan (UM) Department of Economics MITRE grant, National Science Foundation Dissertation Improvement Grant (#1260911), the UM Department of Afro-American and African Studies; the UM African Studies Center; the Rackham Graduate School International Research and Graduate Research Awards; the Center for Public Policy in Diverse Societies Research Grant; Center for International Business Education Research Grant; the Center for the Education of Women Research Grant. This paper does not necessarily reflect the views of the NSF or other donors. We thank Edgar Agaba, Theopista Fokukora, and the entire research field staff for essential support in implementing this project. We thank Bob Brush and Thermo-Fisher Scientific for assistance with drug testing. This study received ethical approval from the UM-IRB (HUM00067957 and HUM00091197), the MUST-IRB (No. 18/09-12), and the UNCST (IS-93).

1 Introduction

According to a recent meta-analysis, between 12 to 50 percent of anti-malarial drugs sold in the private sector in sub-Saharan Africa are of "low quality": a catch-all term ranging from counterfeit, yet effective generic medicines to substandard or falsified dosages [Nayyar et al., 2012]. Bate et al. [2011] show that low drug quality is a global issue; in data from 17 middle and low-income countries, capturing 8 different types of medicines classified by the WHO as essential medicines, 15 percent are of low-quality. Low-quality drugs may harm individuals by delaying effective treatment or wasting money. They are also a public health concern, as they contribute to drug-resistant diseases and render [Okeke et al., 1999]. Identifying ways to improve the quality of drugs– without raising prices– is important to improving the quality of healthcare, and ultimately health, in resource-poor countries.

Despite the potentially substantial threat that low quality drugs pose to global health, there is a lack of empirical evidence on low-quality drugs due to substantial methodological challenges. The antimalarial drug market in low-resource countries is typically composed of informal establishments with loosely enforced regulation. Official records may be incomplete or inaccurate, and building a sample frame to estimate an "average" market rate of drug quality is difficult. Chemical assays to determine the percentage of active ingredients– the gold standard to measure drug quality– are also time-consuming and expensive to conduct for a large number of samples. As a result of these challenges, the existing literature instead focuses on documenting that counterfeit and substandard drugs are found in a variety of countries and drug classes. Empirical work typically relies on small samples and non-random sampling methods, making it difficult to ascertain the average rate of drug quality in any given area and assess the extent of the problem. Furthermore, little is known about what factors explain variation in drug quality rates within a country, whether price acts as a signal of quality to consumers, or if shopkeepers are complicit in the sale of deceptively low quality drugs.

To address these limitations in the previous literature, we conduct a study with two primary objectives: 1) To estimate the average prevalence of low-quality antimalarial drugs; 2) To ascertain what transaction, customer, and vendor characteristics are associated with low-quality drugs. We purchase and test 879 dosages from 412 outlets in randomly sampled areas

of Uganda to estimate the average prevalence of counterfeit and substandard antimalarial drugs. We focus on artemether-lumefantrine (AL), the WHO recommended first-line treatment for uncomplicated malaria in Uganda. We concentrate on AL due to previous evidence of a high counterfeit rate for this drug and the significant health implications to consumers from ineffective treatment. We inspect and test all drugs using a handheld spectrometer, a device that is suitable for field studies and that can quickly test drugs for authenticity. This device also allows us to distinguish between two types of low-quality drugs: 1) counterfeit, yet medically effective drugs, and, 2) substandard medicines, those less likely to be medically effective. We link transaction data on the tested drugs with survey data from vendors at the outlets. This rich descriptive data allow us to examine supply and demand factors that are correlated with drug quality. We report results from survey modules on shop operations, beliefs towards low quality drugs, and a survey exercise which estimates whether providers know they sell low-quality drugs.

Contrary to the existing literature, we find that drug quality is relatively high in the study area. While we find that counterfeit drugs are relatively common at 17 percent, we estimate only 3.8 percent of drugs are substandard. While we cannot generalize beyond the study area, or to other classes of drugs, it appears that estimates in the previous literature are an upper bound on the rate of low-quality drugs. Our additional analysis of correlates leads to the three stylized facts regarding counterfeit and substandard drugs: 1) substandard low-quality drugs are typically diluted versions of high-quality drugs, as opposed to sugar pills; 2) substandard drugs are deceptive, meaning customers cannot distinguish substandard medicines from high-quality medicines; 3) a small number of vendors are complicit in the sale of low-quality drugs.

However, identifying outlets and vendors that are more likely to sell low-quality drugs, or signals of drug quality is difficult. Price is uncorrelated with whether a drug is substandard, suggesting that price is not a signal of quality. Instead, a significant predictor of quality is whether or not the vendor has combined different blister packs together, consistent with diluting high-quality dosages with low-quality tablets. In general, however, vendor behavior changes only slightly during low-quality drug sales. Observable characteristics of the establishment are generally not correlated with drug quality, and we reject the theory that degradation due to poor storage conditions is a cause of low-quality drugs. Our results suggest that mobile drug hawkers are how substandard medicines infiltrate supply chains, but that the relatively

high rate of counterfeit drugs suggest uneven and lax manufacturing standards. We end with a discussion of potential policy interventions, including results from the vendors themselves on what would likely work to further improve drug quality rates.

While our data are extensive, there are some limitations in our analysis. Results are specific to the study setting and context. Although our data are from randomly selected areas, we document important geographic heterogeneity. Thus, even studies based on random sampling may find different average rates. We also caution that while we account for potential confounding characteristics, our empirical analysis on correlates should not be interpreted causally. Finally, we are unable to measure seasonal time-varying factors, or the drug quality rates of other drug classifications. Theses limitations highlight the pressing need for ongoing, large-scale studies of drug quality in important markets for global health.

This paper is organized as follows. In Section 2 we provide relevant background on antimalarial treatment and the Ugandan operating environment. In Section 3 we provide the theoretical framework of our analysis, potential theories of how and why low-quality drugs end up in markets. We next turn to describing how our study addresses each of these theories in Section 4, and describe the data that we collect. In Section 5 we present descriptive results and in Section 6 we outline our empirical estimating equation. In Section 7 we present our results on the correlates of low-quality drugs. We then discuss important mechanisms to interpret our results, namely if providers are complicit in the sale of low-quality drugs in Section 8. In Section 9 we conclude.

2 Study Background

We first outline the market for antimalarial treatment in Uganda. We discuss the choices that individuals have for antimalarial treatment. We conclude with an outline of the laws and institutional characteristics relevant to the sale of low-quality drugs.

2.1 Healthcare in Uganda

All essential medicines, including antimalarial treatment, are available for free in public sector facilities. However, there are not enough facilities to meet the needs of the population. There are also concerns of corruption and "diversion" of free medicines from public sector outlets

for sale in the private sector. Furthermore, the public sector is characterized by long waiting times, poor service quality, and frequent stock-outs [Xu et al., 2006].

As a result of these problems in public sector distribution, the private sector is the first source of treatment for common diseases such as malaria [Konde-Lule et al., 2012]. The private sector is composed of nearly 17,000 drug shops and clinics, and 415 pharmacies [ACTwatch Group, PACE/Uganda & the Independent Evaluation Team, 2012].[1] Despite their key role in healthcare delivery, vendors may lack minimum medical qualifications. Stanback et al. [2011] estimates that up to 60 percent of drug vendors in Uganda operate without the regulated medical qualifications and licenses, potentially contributing to the low levels of healthcare quality. These averages are not unique to the Ugandan context, but rather reflect the larger problem of low healthcare quality in many developing countries. In addition to a lack of required training and licenses, vendors have been documented to overuse medicines (particularly antibiotics) and are observed to misdiagnose common ailments [Das et al., 2012, Das and Hammer, 2007, Leonard and Masatu, 2008].

2.2 Anti-malarial Treatment in Uganda

In Uganda, malaria represents the largest burden of illness and is the second-leading cause of death among children under 5 in Uganda despite being a largely treatable disease. It is endemic throughout 90 percent of the country all year round, although there are seasonal peaks. The recommended first-line treatment of uncomplicated malaria is artemether-lumefantrine (AL) [ACTwatch Group, PACE/Uganda & the Independent Evaluation Team, 2012]. A full adult dosage of AL consists of 24 tablets taken over 3 days, and is typically sold in one clearly labeled blister pack indicating dosages per day.[2] AL is part of a larger class of medicines known as ACTs (artemisinin-combination therapies), and is over 95 percent effective in both adults and children [Baird, 2005]. Quinine is another effective and commonly available medicine. However, official regulations recommend quinine to be reserved as a second-line treatment and for complicated malaria (a severe form of malaria that can result in death, and primarily

[1] The establishment types differ with respect to formality and size, but all are common sources of care for malaria. Drug shops and clinics are typically smaller establishments selling only over-the-counter medicines for common ailments, although clinics are more likely to have beds to treat patients and offer consultation services. In contrast, pharmacies carry a larger number of items, including prescription-only medicines.

[2] For children, a full dosage depends upon the age and weight of the child. While there are blister packs with fewer tablets for younger children, it is also common to cut packs.

affects children). Older treatments– such as sulphadoxine-pyrimethamine (SP) – are also widely available, and cheap, but are clinically ineffective due to parasitic resistance [ACTwatch Group, PACE/Uganda & the Independent Evaluation Team, 2012].

Work as recent as 2011 has found that AL is available at only 30 percent of outlets across six sub-Saharan African countries, including Uganda [O'Connell et al., 2011]. This low rate of availability may be due to either low supply, low demand, or both. Stock-outs are reportedly common in the private sector, potentially due to inefficient stock management and poorly functioning supply chains. On the demand side, this low level of utilization is potentially due to high prices, low levels of caregiver awareness, or to beliefs that low quality drugs are widespread. AL is a relatively newer treatment, and relatively expensive. In contrast to other sub-Saharan African countries, there is no regulation on prices in Uganda and few people have health insurance; thus, health care is paid nearly entirely out-of-pocket. Misconceptions regarding malaria transmission are common, and symptomatic diagnosis from caregivers is problematic [Nuwaha, 2002]. Low drug quality may also contribute to suboptimal utilization. If caregivers cannot tell the difference between high and low quality drugs, then standard economic theory predicts that households will be less willing to purchase drugs at all, potentially leading the market to collapse [Akerlof, 1970]. Potentially as a result of these issues, it is estimated that one-third of febrile ("symptomatic") children in Uganda do not receive first-line anti-malarial drugs [Uganda Bureau of Statistics, ICF International, 2012].

To increase the usage of essential medicines, there are large-scale manufacturing subsidies on select brands of AL under different global donor programs. While prices have fallen nearly by half since 2011, prices remain more than three times as high as the target price of the program (US$1).[3] These drugs are typically inspected for quality to ensure that minimum manufacturing standards are met. For example, the WHO pre-qualification program is intended to insure that the subsidized drugs that are distributed through governments and international agencies pass minimum quality standards.

[3]The effectiveness, and cost-effectiveness, of the subsidies at lowering prices is a subject of debate in global health. Studies generally find that the large-scale subsides of the AMF-m program were effective at lowering prices [Tougher et al., 2012].

2.3 Relevant Ugandan Laws and Other Institutional Background

As in many other markets of developing countries, the private sector for healthcare in Uganda is largely informal, with unevenly enforced penalties. There are currently weak legal penalties, regulations, and laws for dealing in fake drugs.[4] The government agency in charge of regulation is the Ugandan National Drug Authority (NDA). As part of its mandate, the NDA has undertaken initiatives to test imported drugs, license and regulate pharmacies and drug shops, contribute to consumer sensitization on the appropriate use of medicines, and raise awareness about the problem of counterfeit drugs. Criteria for licensing include verifying educational attainment, passing facility inspections, and paying relevant fees. A recent audit covering the period 2006-2010, however, indicates that the agency overall is understaffed and under-funded to achieve these objectives. For example, with respect to the guidelines to license all drug shops within the country, the audit finds that due to bureaucratic delays, many unlicensed outlets that had applied for a license had not yet received one. In fact, "no license had been issued within the stipulated timeframe" [Office of the Auditor General , 2010].

3 Theories of Low-Quality Drug Sales

In this section, we briefly outline various theories of how low quality drugs appear at outlets. We divide the various theories of low drug quality into three general categories: 1) demand-side factors, those factors that vary at the transaction level; 2) supply-side factors, those that vary at the outlet level; 3) institutional factors, including legal and social penalties, that vary at the either the outlet or market level. While some factors could arguably fit into several categories, our rough delineation will serve to guide the empirical analysis while keeping the unit of observation roughly the same within groups.

3.1 Demand-side Factors

The previous economics literature on low-quality drugs has focused on whether consumers are duped into purchasing low-quality drugs, or whether consumers can identify low-quality drugs

[4] A legal officer at Uganda National Bureau of Standards, another regulatory body in charge of counterfeit goods has been quoted as saying that under current laws, the maximum penalty is two years imprisonment and a nominal fine [Mubiri, 2011]. However some individuals have been sentenced in Uganda for longer periods and/or paid larger fines. Although the Ugandan Parliament has considered revising fines and jail sentences associated with convictions, no proposals have yet been passed.

through a price signal. Bate et al. [2011] find in non-random samples from 17 countries that the price of antimalarial drugs is weakly and positively correlated with drug quality, and that the subjective shop appearance is also correlated with drug quality. Bjorkman et al. [2012] similarly find that the counterfeit drug rate within a non-random sample of project villages is weakly positively correlated with price. These facts suggest that 1) price is a noisy signal of quality or 2) a low willingness-to-pay for quality of a drug, at least for the marginal consumer.

A complementary explanation for low drug quality is that certain types of customers are more likely to purchase low-quality drugs than others. Numerous studies have documented that misconceptions about malaria causes and symptoms are widespread, and low average education levels may make customers more vulnerable to fraud. Bjorkman et al. [2012] also show that the counterfeit rate is negatively correlated with the average village understanding of malaria transmission. In other words, shopkeepers take advantage of low information customers and differentially dispense substandard medicines to such groups. This theory can be generalized into a theory of either taste-based or statistical discrimination. Providers may simply not like individuals with certain characteristics, and selectively choose to give them low quality drugs. Alternatively, providers may make assumptions about demand characteristics on the basis of customer appearance or behavior. For example, providers may be able to guess an individual's knowledge or education level, or whether they are a local person, based upon gender or tribe of the shopper.

3.2 Supply-side Factors

Low quality drugs may alternatively be primarily driven by characteristics of vendors and supply chains. As part of their business decisions, vendors must also consider how to maximize profits based upon input costs. Vendors make choices of what types of drugs to stock, when to order, and from which supplier. There are several different suppliers that vendors could use to procure new stock at any given time; these suppliers range from other retail outlets to wholesale pharmacies to mobile drug hawkers. These suppliers may differ in their cost and quality of inventory. The informal nature of the market and lack of enforced regulation may make strategic behavior– such as selling cheaper, low quality drugs as high quality drugs– a more attractive alternative.

Theories of supply-side factors influencing drug quality generally imply that vendors know-

ingly dispensing substandard drugs to unsuspecting customers to maximize profits. However, the knowledge that vendors have regarding their inventory quality is unclear. Shopkeepers could, for example, unintentionally sell low-quality drugs that have degraded over time due to improper storage conditions. Exposure to heat, direct sunlight, or excessive humidity may cause active ingredients to degrade and become ineffective. Even if low drug quality is the result of purposeful human behavior, vendors themselves could still not be aware of the quality of their inventory. Counterfeiters are adept at mimicking licit packaging and providers themselves could lack the ability to distinguish high quality from low quality drugs. Considering potentially low business acumen, a plausible explanation is be that shopkeepers are deceived by suppliers and unwittingly dispense low-quality medicines to customers.

3.3 Other Institutional factors

Closely related to supply-side factors are characteristics of the firm's operating environment. Many firms, particularly in developing countries, operate informally, meaning they lack the regulatory credentials and licensing. Although these regulations and licenses are intended to reduce quackery, suggesting that regulation improves market quality, the opposite may just as well hold. Firms operating outside of the formal healthcare system may particularly wish to avoid the attention of law enforcement or regulators, causing them to be vigilant regarding their inventory quality.

An additional relevant factor may be the risks of legal or social penalties if low quality drugs were determined to be sold at an outlet. Laws as enforced make it difficult for legal penalties to be effective. However, customers may refuse to shop at the store, and the vendor's reputation may be harmed. In addition, it is possible that drug vendors would be socially ostracized, confronted, or potentially become victims of vigilante justice. If these social sanctions vary by area, then perceived penalties may explain varying drug quality rates.

Finally, recent work has demonstrated that competition decreases the likelihood that an establishment will sell low-quality drugs. Both Bennett and Yin [2014] and Bjorkman et al. [2012] design studies which randomly introduce competition into the medicine market; Bennett and Yin [2014] randomizes chain store entry into urban areas of India and Bjorkman et al. [2012] randomize a high-quality NGO competitor into monopolistic villages in Uganda. Both find that the increase in competition causes an increase in drug quality. Due to the large

literature of the effect of competition on quality, we analyze the effect of competition using our data in a companion paper, Fitzpatrick [2015a].[5]

4 Study Design

We collect a robust set of data to address these various theories of drug quality. Fieldwork took place from May-August 2013 and consisted of several rounds of data collection. First, the sample frame was constructed by doing a census of vendors within randomly selected areas. Second, two different covert shoppers visited each outlet and each purchased an antimalarial drug according to a randomly assigned script. Third, additional survey data were collected from the drug dispenser at each outlet. Figure 1 contains the project timeline.

4.1 Study Sites

Uganda is composed of 112 districts that are each divided into counties.[6] Each county is further divided into subcounties; each subcounty is divided into parishes; each parish is then divided into villages.[7] Within each selected study district, we randomly selected two rural and two urban subcounties.[8] We then randomly selected two parishes within each urban subcounty and three parishes within each rural subcounty in each of the five districts. In Kampala, the sample was selected comparably but with somewhat different definitions. Kampala does not have sub counties or counties, but rather 5 divisions are each divided into zones. We randomly selected 1-2 zones in each of Kampalas five divisions to roughly make the sample equally distributed within parish. In total, the study contains a total of 45 parishes, in which there are 142 villages with at least one drug outlet.[9] We conservatively estimate that the study area covers all drug vendors serving more than 200,000 people.

Study team members then conducted a census and created detailed maps of all drug outlets within study parishes with a corresponding physical description of the outside of the premises.

[5] For a summary of the literature on the effects of competition on quality in the U.S. context, see Gaynor [2006]

[6] The total number of districts at the time of data collection was 112.

[7] According to the 2002 census, the average size of a parish is 4,625 people; the average size of a sub-county is 25,289. The average size of villages was not reported [Uganda Bureau of Statistics, 2008].

[8] Bushenyi, Busia, Mbarara, Rukungiri, and Kampala (the capital) were chosen as study districts due to their size and proximity to borders and ports-of-entry to the country.

[9] Note this figure is less than 50 because there are fewer parishes in some subcounties.

We define drug outlet as "an immobile establishment that sells antimalarial drugs for profit."[10] Due to the informal nature of the market, finding all outlets can be difficult and plagued with measurement error. Many outlets lack outward signage, for example. Therefore, we cross-referenced information from local individuals (village chiefs,passers-by, and motorcycle taxi drivers) at each stage of data collection in order to ensure that all outlets were included in the study. We have evidence that these extensive data quality procedures were successful. For example, our study includes a small number of "other" types of outlets. These include individuals who sell antimalarial drugs out of their homes and shops that specialize in another market, such as hardware supplies, but also sell antimalarial drugs. The nature of this sampling strategy results in having a nearly complete picture of local antimalarial drug markets.

4.2 Drug Purchases and Testing

Following the recommendations of Newton et al. [2009], mystery shoppers were used to purchase medicines from these drug outlets.[11] The following protocol was used for purchasing drugs:

1. Buy the cheapest brand of AL offered.
2. If a full dose of AL is not available, buy quinine.
3. If a full dose of quinine is not available, then buy the next cheapest antimalarial available (typically SP).
4. If none of these is available, buy any other antimalarial.
5. If a full dose of any antimalarial is not available, do not buy anything.

At the conclusion of the fieldwork, all purchased drugs were inspected by research assistants. The recorded drug characteristics include brand, expiration date, number of tablets, and whether the drug had public sector markings ("diverted"). Drugs were then shipped to a laboratory at the University of Michigan for testing with a handheld Raman spectrometer,

[10] Note that this definition does not require that the establishment actually make a profit. Herbal shops are also excluded from the sample frame, as are charitable or public sector hospitals or pharmacies.

[11] Shoppers used a randomized script to either declare that they have malaria or ask the shopkeeper for a diagnosis; then they either asked for a specific product or for what the shopkeeper recommends. All shoppers did all scripts. There are statistically significant observed differences with respect to both price and quality, and these results are reported in Fitzpatrick [2015b].

the TruScan^{TM} RM.[12] Testing consists of comparing a purchased tablet with a separate, high-quality authentic tablet of the same brand. As part of testing we collected high-quality tablets from manufacturers and wholesalers in Uganda, and built a "spectral library" for the study. Each purchased tablet was tested at least once according to a strict protocol.[13] Our analysis is restricted to those dosages for which we were able to obtain a comparison high-quality tablet of the same brand (N=879). We aggregate to the transaction/sample level in order to standardize the unit of analysis.[14]

4.3 Definitions of Drug Quality

We define counterfeit as a purchased dosage ("sample") for which at least one tablet within the sample failed the spectrometry analysis. Counterfeit specifically refers to a tablet that has a different Raman spectrum than the authentic comparison tablet of the labeled brand. We identify the brand based upon the package labeling. Note that the majority of medicines studied are non-branded generics, typically local or national brands registered within Uganda. Our measure of counterfeit does not compare the chemical composition of a generic to an innovator brand, but rather a given brand to itself.

However, many brands are chemically similar. In addition, although tablets are intended to have uniform and consistent contents and manufacturing, some brands may have high within-brand variation. In practice a tablet that failed the comparison against its own high quality authentic tablet could potentially match against another brand within the library.[15] We define "substandard" as a subset of counterfeit medicines where the tablet's spectra could not be matched to any other high-quality authentic in the library. The ability to cross-check the authenticity against other brands is an advantage of creating a large spectral library, and testing a large number of brands with the same active ingredient. The logic behind creating this measure is that if we cannot find any chemical match to a given purchase, it is less likely

[12]This device is the suitable for field work and is currently being used by customs officials in Nigeria, Uganda, Ghana, and other countries to detect low quality medicines. A version of this device was also used in Bjorkman et al. [2012] and validated in Bate et al. [2009].

[13]The protocol is listed in the Online Appendix for the companion paper, Fitzpatrick [2015b].

[14]To protect respondent and manufacturer confidentiality, we do not publish quality results by brand or disaggregated results that could potentially identify outlets or locations.

[15]This is known as "low selectivity", because many drug compositions are similar. Among high-quality tablets of AL, for example, all brands were similar enough to at least one other brand to pass. The selectivity tests are non-transitive, because Brand A can be similar to Brand B, and Brand B can be similar to Brand C, but Brand A can be different from Brand C. The tests are also asymmetric.

to be medically effective– the measure of quality or primary importance to consumers.

One potential concern with our methodology is that the high fluorescence of the artemether can affect the validity of testing with the handheld spectrometer. Hajjou et al. [2013] and Bate and Hess [2010] perform validation exercises on a previous version of the TruScan^{TM} RM. Both studies confirm that there are a relatively high number of false-positives with respect to AL, implying the potential bias from our method of testing would likely lead to an *overestimate* of the rate of counterfeit and/or substandard drugs.[16] Since the bias of the test is of a known direction, we retest all failing tablets and a random subset of dosages with all passing tablets. This subset is 11 percent of all purchased tablets (N=2322). Results are in Appendix Table A1. Conditional on passing the first scan, the likelihood the same tablet passes the second scan is 98 percent; conditional on a tablet failing the first scan, the likelihood that the same tablet fails the second scan is 75 percent. Among this retested sample, the correlation of results across scans of the same tablet is 0.666. However, aggregating results to the transaction level reduces the amount of variability in results and improves internal consistency.[17] Conditional on a sample passing according to the first scan, the likelihood that the sample would maintain its classification in the second scan is 100 percent. Conditional on a sample failing according to the first scan, the likelihood that the sample still fails under subsequent testing is 72 percent.

Therefore, by using a conservative estimate of whether the tablet failed all of its scans, we are able to reduce the spectrometer bias towards counterfeit medicines. Furthermore, by aggregating failures to the transaction level, we are able to reduce the variability of results due to noise or error. We conclude that our testing methodology is valid and internally consistent. In terms of external validity, however, our results may represent an upper bound on the rate of low-quality antimalarial drugs, compared to chemical assays.

4.4 Other Data Sources

Later, we revisited the same outlets and conducted an extensive survey on their background and shop operations (N=452). The enumerator of the survey was a different person from the mystery shopper. Enumerators also recorded their observations and impressions of the shop. These surveys include 424 outlets where at least one dosage was purchased, and 415 outlets

[16] We are conducting additional analysis in order to estimate precisely the rates of Type I and Type II error with the updated version of the machine that was used.

[17] Recall a full adult dosage is 24 tablets.

with a complete survey and a purchase that was also part of the analysis sample (77 percent of all establishments). Among all purchases, we estimate that 85 percent of the individuals who dispensed the medicine also answered the survey.[18] There is no correlation between drug quality and survey completion (not shown).[19]

5 Descriptive Results

Before summarizing our key results on quality, we first present summaries of outlet characteristics, including vendor beliefs about fake drugs in their area. We frame our discussion in terms of the vendor and institutional characteristics that may affect the prevalence of low quality, and that we will later use as correlates in our empirical analysis. We then present our key result that drug quality is high in the study area, and that low-quality drugs are typically diluted versions of high-quality drugs.

5.1 Averages from Vendor Data

We summarize responses from the vendor survey in Table 1.[20] Fifty-three percent of the outlets are drug shops; 40 percent are clinics; the remainder are primarily pharmacies. Most vendors at these outlets are women; only 23 percent are men. Outlets are small, and have on average 2.31 employees; 63 percent of outlets have only 1 employee working. Profits are highly skewed, and the median profits are $77.13. We estimate that only 36 percent of vendors meet the legal education qualifications to dispense medicines, although 69 percent of drug shops report that their outlet had been visited by a representative of the NDA in the previous year. These figures demonstrate a gap between official and actual enforcement of regulation.

Despite the lack of formal qualifications, vendors are important sources of healthcare for their communities. On average vendors treat 21.8 customers per day (median value: 15), of which six are treated for malaria. Vendors generally are aware of proper medical protocols

[18] In order to determine whether the dispenser is the same as the respondent, we use the following methodology. Subsequent to both mystery shopping and the vendor survey, a member of the study team filled out a short survey with a physical description of the dispenser or respondent. We then take the difference in observations, controlling for study team member fixed effects, and estimate a probit regression where the left hand side is instances where we are sure there is or is not a match. We then use as a cutoff for a match 0.50 of the predicted values.

[19] In addition, surveys with real customers were conducted at the same outlets (N=867 real customers at 350 outlets; 333 outlets with a complete survey also had a real customer survey).

[20] A summary of mystery shopper characteristics is in Appendix Table A2.

for malaria. Eighty-four percent of respondents to the survey correctly report AL as the first-line treatment, and 81 percent generally knew the correct protocol for taking AL. However, the advice given to mystery shoppers is, on average, slightly worse than responses given to enumerators during the survey. Among vendors making a drug recommendation, 75 percent recommended AL and 11 percent recommended another first-line (typically more expensive) treatment. In 9 percent of transactions, the vendor recommended SP, which is no longer an effective medicine. The distribution of recommendations is approximately the same at outlets where the respondent correctly reported the first-line treatment, indicating a gap between knowledge and practice.

There is a similar gap between practice and optimal public health practices. Despite the risks of increased drug resistance if individuals sick with malaria take partial dosages, vendors frequently sell partial dosages. Vendors report only 66 percent of customers buy a full dose, potentially reflecting the expense of first-line treatment. This relatively low fraction may also make it easier to dilute medicines with substandard tablets, as it provides a way for vendors to justify mixing packs. Similarly, despite the need to clinically diagnose malaria with a blood test (either through blood microscopy or rapid diagnostic testing– RDT), only 53 percent of establishments offer testing services, and the average cost of RDT testing is $1.09– approximately one-third the cost of first-line treatment. Conditional on having tests available, only 44 percent of mystery shoppers were advised to have the patient take a malaria test.

On the drug vendor survey, we included a module on counterfeit and low-quality drugs. We first asked vendors why other vendors would sell a fake drug. Of those answering the question, 85 percent suggested the reason was related to money, or increasing their profits; only 15 percent thought it was ignorance, and that vendors couldn't identify fake drugs in their stores. We then asked vendors what percentage of outlets in their parish and their district sold fake drugs. On average, vendors thought that 32 percent of outlets in their parish sold fake drugs; and 44 percent of outlets in their district sold fake drugs. Thirty-eight percent of respondents thought that they could identify a fake drug if they saw one, suggesting that identifying low-quality drugs is difficult for the majority of vendors.

We also asked vendors about their thoughts on what the social sanctions would be if a vendor was to be caught selling low quality drugs. Only 6 percent of vendors thought that there would be no consequences to selling a low quality drug. Nearly all respondents (95

percent) believed that the vendor would be reported to authorities, and 93 percent thought that customers would boycott their store. Ninety-two percent thought that the shop would be closed, and 88 percent believed that the stock would be confiscated. Seventy-seven percent of vendors thought that customers would do something violent (vigilante justice).

5.2 How available is AL? Averages from Mystery Shopping Data

In contrast to O'Connell et al. [2011], we find that complete drug stock-outs are relatively rare and that AL is readily available in most of the study area. In Table 2, we show that shoppers were able to purchase AL during 86 percent of the 933 successful purchases (N=806); of these, 796 were able to be tested (99 percent). This average translates into a purchase of AL at 92 percent of drug outlets, and suggests that the drug purchasing protocol was successfully implemented. While smaller fractions of the sample, quinine and SP are also commonly available and purchased drugs.

It is important to note that there are different characteristics associated with the different active ingredients.While AL is typically available, it is relatively expensive compared to the other treatments. The high price may deter utilization, although discounts are available for those who can successfully bargain. The average price paid for AL is $3.19, compared to $2.48 for quinine and $1.12 for SP. There is also a substantial amount of variation in prices, for each active ingredient. As measured by the coefficient of variation, prices vary substantially for each type of active ingredient, from 0.39 - 0.46.[21] Note that the type of drug purchased was decided through our study protocol. Therefore, we do not use the type of active ingredient as an outcome variable in our analysis. These differences extend beyond price. For example, quinine is highly likely to be sold as an incorrect dosage, a difference that may originate due to standard packaging. While AL is typically sold in one blister packs– allowing customers to recognize a full dosage more easily– quinine is typically sold as loose tablets in bulk-size bottles or in multiple blister packs. In addition, different drugs are more likely to be diverted from the public sector. While 7.9 percent of the AL sample and 6.3 percent of the SP sample had public sector markings,none of the quinine sample had public sector markings.[22]

[21] Sorensen [2000] finds in the US market that, for a given prescription drug, the highest price is 50 percent over the lowest price, and the coefficient of variation is 0.22.

[22] However, in the full sample of all purchases, 17.6 percent of the quinine sample had public sector markings.

5.3 Low Quality Drugs or Diluted High Quality Drugs?

Overall, we find that drug quality is relatively high in the study areas compared to previous studies, and that low-quality drugs are characterized by dilution of high-quality drugs. We estimate that average rate of counterfeit drugs is 17 percent, and the average rate of substandard drugs is 3.4 percent. However, the average rate of outlets ever selling counterfeit or substandard drugs is slightly higher. Counterfeit drugs were sold at 25 percent of outlets, and substandard drugs were sold at 5.6 percent of outlets.

The difference between the average rates at the transaction level and at the outlet level indicate that outlets tend to sell both high and low quality drugs. Similarly, tablets within a dosage are typically a mixture of both high and low quality. Within the group of samples with at least one counterfeit tablet, on average 61 percent of tablets were determined counterfeit; within the group of samples with at least one substandard tablet, on average 47 percent of tablets were substandard. We interpret this quality dilution as a strategic response to avoid detection. This interpretation is reminiscent of Salop [1977]'s seminal work on price variation: by increasing noise in the distribution of prices, providers make it difficult for consumers to distinguish high-priced sellers from low-priced sellers. This interpretation is bolstered by additional analysis of the distribution of failing tablets. The medical literature suggests that in the short-term, a patient who has malaria is likely to feel better so long as they consume at least 16 tablets.[23] The average dose response is consistent with the observed bimodal distribution of the number of substandard tablets within a failing dosage. We estimate that 57 percent of failing samples have enough *passing* tablets that the patient would likely still have their malarial episode temporarily cured. In other words, substandard medicines are more accurately thought of as diluted high-quality drugs than dosages of sugar pills.[24] We conclude that by selling both high and low quality drugs, vendors make it difficult for consumers to learn or distinguish true drug quality across or within transactions.

One caveat to interpreting these averages, and the subsequent results on correlates, is that we are only able to test drugs that were purchased during mystery shopping and part of the

[23] More precisely, the medical literature suggests that the final 8 tablets of the full adult dosage are no more effective at immediately improving health, but rather reduce the likelihood of a return infection of the same parasite over the next month by 18 percent [Van Vugt et al., 1999].

[24] Of course, that estimation is rough and is dependent on a number of assumptions: whether the individual actually does have malaria, the individual's malarial resistance level, diet, the presence of vomiting in the current malarial episode, and the order of which the tablets were consumed.

analysis sample. Because the focus of the study is AL, AL is more likely to be purchased. AL is also more likely to be part of the analysis sample; we were able to obtain a high-quality tablet for 99 percent of the AL sample. In contrast, only 53 percent of the quinine sample is also in the analysis sample and 81 percent of SP is also in the analysis sample.[25] Conditional on being in the analysis sample, we find substantial differences in terms of drug quality rates by active ingredients. In fact, SP and "other" antimalarials are never found in our sample to be either counterfeited or substandard. The substandard rate for quinine is particularly high, at 47.2 percent. These differentials suggest that the average rate of low-quality drugs may differ substantially by active ingredient, and potentially differ from other classifications of drugs.

5.4 Geographic Distribution

The average rates of drug quality, however, mask substantial heterogeneity within Uganda. Figure 2 presents bar charts showing how the average counterfeit and substandard rates substantially vary by district, by as much as 50 percent. In Figure 3 we display graphically in a bar chart for each of the 44 parishes the average counterfeit and substandard rate. A counterfeit drug was sold in 46 percent of villages, and in 59 percent of parishes. These figures demonstrate that previous studies based upon an average rate within a country may not be appropriately capturing the variation within a country. Likewise, studies that aggregate averages from a large number of countries may therefore be masking important variation in the average rate of a given country.

In our study, we find that population density is only a weak predictor of drug quality. We find that urban areas have higher rates of counterfeit drugs, but rural areas have slightly higher rates of substandard medicines.[26] In particular, 13.2 percent of transactions in rural areas were counterfeit, and 5 percent were substandard. In contrast in urban areas, 18 percent of transactions were counterfeit, but only 3 percent were substandard. The difference in rates at the transaction level is statistically significant for counterfeit drugs, but not substandard drugs. At the outlet level, 20 percent of rural outlets sold a counterfeit drug, and 7 percent of outlets sold a substandard drug. In urban areas 26 percent sold a counterfeit drug and

[25] Because quinine is often sold as loose tablets, we could not always identify the appropriate brand for testing.
[26] We classify "urban" as parishes located in the district town. All areas of Kampala are classified as "urban".

5 percent of outlets sold a substandard drug, although these differences are not statistically significant.

5.5 Comparison to Previous Studies

One challenge with comparing our results to the previous literature is that the bulk of existing studies have small samples or non-random sample selection criteria. For example, in the meta-analysis by Nayyar et al. [2012], only 5 studies with data from African countries collected data on artemether-lumefantrine, and all use convenience sampling. The average of those studies indicate a substandard rate of approximately 30 percent. When considering other types of antimalarial drugs, the view is similar. Only six of the twenty-eight studies cited in the meta-analysis rely on random sampling methods, and only three of those studies based on random sampling have more than 50 observations.

A primary concern with the previous reliance on small, or non-random sampling is that results may not be generalizable to a broader population. Small samples may not adequately represent the true mean of drug quality within a particular area. Researchers using non-random sampling methods may inadvertently target areas with lower drug quality on average. More recent work by Kaur et al. [2015] formalizes this intuition with a study in Nigeria. The authors find that convenience sampling yields a substantially higher rate of low-quality drugs than random sampling. The authors also find that there is little difference in the drug quality rates between mystery shopper (i.e., covert purchases) and overt shopper purchases, where providers gave informed consent to have their drugs tested for quality.

An additional difficulty is that the methodology for testing differs substantially from study to study; using equivalent definitions across different testing methodologies is not always possible. For example, Bjorkman et al. [2012] use an older version of the handheld spectrometer which allow for a measure of the counterfeit drug rate; they estimate that 21 percent of drugs in select areas of Uganda are counterfeit. However, the updated version of the machine that we use includes technology that allows for repeat testing on the same tablet. We are therefore able to cross-reference failing tablets to create a new measure of whether tablets are medically effective ("substandard"). Our results imply that substandard drugs are substantially less common than counterfeit drugs at 4 percent.

We cannot conclude, however, that Bjorkman et al. [2012] would have also found a low

prevalence of substandard drugs had they done additional testing for two main reasons. First, in 2012, after the data of Bjorkman et al. [2012] was collected and before this study took place, the National Drug Authority closed a national manufacturer of substandard antimalarial drugs. This event was highly publicized throughout the country [Mugisa, 2012]. Closing down a low-quality manufacturer would have a direct effect of eliminating a primary source of low-quality medicines, and an indirect effect of causing competing firms to increase quality to avoid closure. Second, the sampling methods between our studies differ substantially and have little overlap. Bjorkman et al. [2012] used drug outlets located in project villages in exclusively rural areas, and 55 percent were monopolists. Our results are based upon random sampling methods from both urban and rural areas, and only six percent of vendors are monopolists.

In conclusion, the average findings reported in previous literature are potentially upper bounds on the true low drug quality rate. However, due to the nature of the testing methodology, definitions used, sampling strategy, and other ongoing events in the market, it is potentially misleading to make comparisons to any specific previous study. Nevertheless, by our estimation, the rate of ineffective medicines is less than 4 percent in the study area, indicating that low quality drugs are not widespread in all areas of developing countries.

6 Empirical Analysis on Correlates

In our empirical analysis, we first test what transaction-level correlates are associated with drug quality. Second, we analyze the outlet-level correlates of low quality drugs. Whether the drug quality rate for outlet i (transaction i) is correlated with a given characteristic X is analyzed through the following regression:

$$Quality_{ip} = \beta_0 + \beta_1 X_i + + \beta_2 Z_i + \gamma_p + \mu_{ip} \tag{1}$$

where $Quality$ is measured in two primary ways: whether or not at least one tablet within the dosage is substandard, and whether or not at least one tablet within the dosage is counterfeit. X is a vector of characteristics that are potentially linked with drug quality, and Z is a vector of control variables to account for potential confounding factors. Our outcome variables of drug quality are binary, and we use a linear probability model to estimate marginal effects; however, results are robust to using a conditional logit specification. Because there is sub-

stantial variation across and within districts, we include a parish fixed effect, γ, to account for any parish-specific characteristics. The fixed effects typically boost statistical power, although in relevant regressions we show a district fixed effect specification that has more degrees of freedom. We estimate robust standard errors, and cluster at the outlet level in regressions where the unit of observation is a transaction.[27]

7 Results: Correlates of Low Quality Drugs

7.1 Price and Transaction-Level Correlates

We begin by testing the correlation between quality and characteristics of a given transaction that would be observable to customers. These characteristics include price, the number of separate blister packs, whether the dosage had public sector markings (diverted), and whether there was an expiration date marked. Results are in Table 3.

We find that while counterfeit drugs are priced \$0.01 lower than non-counterfeit drugs, there is no correlation between price and whether the drug is substandard. Assuming patients care primarily about the medical efficacy of drugs, the lack of a price signal for substandard drugs is consistent with consumer deception. In contrast, other characteristics of the transaction are potentially stronger, albeit noisy signals of quality. For example, the number of blister packs– loosely thought of as "chances" to dilute the dosage – are negatively and strongly correlated with both counterfeit and substandard drugs.[28] Each additional blister pack in the purchase is associated with a 3.3 percentage point higher likelihood of being sold a counterfeit drug and a 2.8 percentage point higher likelihood of being sold a substandard drug. Contrary to expectations, drugs lacking an expiration date are 12 percentage points *less* likely to be counterfeit; there is no relationship between having an expiration date and whether the drug is substandard. Diverted medicines are 26.6 percentage points more likely to be counterfeit and diverted drugs are 9.3 percentage points more likely to be substandard. Results are robust to controlling for fixed effects for drug type and mystery shopper, indicating that shopper characteristics are not mediating factors in these relationships.

The negative correlation between drug quality and whether or not the drug is diverted

[27] In order to increase statistical power, we use the full set of non-missing observations available for each group of outcomes. As a result, the sample size may differ between tables.

[28] Loose tablets were each considered a separate blister pack to have the same interpretation.

is surprising. Medicines in the public sector are quality-assured through the Global Fund pre-qualification program for manufacturers. There are two interpretations of the negative correlation. One interpretation is that there is corruption and a lack of appropriate control measures as part of the pre-qualification program. Alternatively, adept counterfeiters may strategically try to make substandard medicines look like diverted drugs. Our data do not allow us to distinguish between these potential explanations.

We now turn to identifying whether vendor behavior during the transaction is associated with drug quality. We consider two sets of independent variables in Table 4: objective characteristics (in all columns) and subjective characteristics (in columns 3-4). Overall, few measures of vendor behavior are correlated with drug quality. Among the set of objective characteristics, only one– asking any questions about the patient– is statistically correlated with drug quality. Shoppers who reported that the dispenser asked any questions about the patient were 9 percentage points less likely to dispense a counterfeit drug, and 6 percentage points less likely to dispense a substandard drug, suggesting that dispensers who give advice may feel a stronger reputational incentive to dispense a high-quality drug. Similarly only two subjective characteristics are correlated with drug quality: vendors rated as "very unfriendly" by mystery shoppers are *less* likely to sell both counterfeit and substandard drugs, while vendors working in outlets rated as "below average quality" by mystery shoppers are more likely to sell substandard drugs. These correlations are robust to including the price paid for the drug, suggesting that these factors are not reflective of the vendor's potential profits from the transaction.

We next turn to examining whether certain groups of customers are more likely to receive low-quality drugs, as in either taste-based or statistical discrimination. We begin by examining characteristics of the shoppers who purchased the drug during mystery shopping.[29] Results in Panel A of Table 5 show that female shoppers are no more likely to buy both counterfeit and substandard drugs. Similarly, shopper tribe is not correlated with drug quality.

In Panel B, we examine whether the vendor's response to the tribe or gender of the shopper depend upon the gender or tribe of the vendor. This analysis slightly reduces our sample size because we are restricted to the sample where the person who dispensed the medicine is the same person who answered the survey (N=603). We find that shoppers of the same

[29]There are 16 different shoppers. Shoppers were not randomly assigned to shops, although shoppers were not assigned to shops in a systematic manner. Summary statistics of shoppers and the matched sample of vendors are in A2.

tribe as the vendor are 8.5 percentage points *more* likely to be sold a counterfeit drug and 4 percentage points more likely to be sold a substandard drug, although the point estimates differ statistically from zero only when controlling for a district fixed effect. Being of the same sex as the vendor is unrelated to quality. These results suggest that although drug quality rates may vary within a market, quality differences are not associated with racial or gender discrimination.

A more subtle factor associated with consumer demand characteristics may be when customers tend to shop. For example, anecdotal evidence suggests that farmers tend to go to trading centers during the day, when it is too hot to work, or after dark. Thus, we test whether the time or day a shopper visited an outlet changes the quality of the resulting sale. Although there is no linkage between time of day and drug quality, substandard drugs are 3.9 percentage points less likely to be sold over the weekend than during the week (not shown). These results are robust to controlling for drug type and mystery shopper fixed effects. Prices are also \$0.18 higher on weekends than weekdays (not shown). Systematic quality fluctuations may indicate that certain types of customers shop different days of the week, and that vendors respond to these demand fluctuations.

7.2 Supply-Side Explanations: Analysis of Drug Quality at the Outlet Level

We now change our unit of analysis to be the outlet level, to test if low quality "outlets" have specific characteristics. There is a large policy emphasis on establishment type, including legal requirements for operation and training of staff and differences in costs of licensing. Therefore, we begin by testing whether establishment type is correlated with drug quality. We find little evidence that establishment type is correlated with drug quality. Results are in Table 6. In Columns 1 and 2, we use the full sample of all outlets with purchases; in Columns 3 and 4 we replicate the analysis with the sample of all purchases at which a survey was also completed (excluding missing values) in order to look at additional outlet characteristics. Establishment type is not related to quality. However, the lack of observed quality differences with respect to establishment type is not simply semantic. For example, establishment type is strongly correlated with the price paid. Compared to pharmacies, clinics charge on average \$0.72 higher and drug shops charge \$0.36 higher per dosage. These results are robust to using the mystery-shopper establishment classification and also the self-reported establishment type on

the vendor survey.[30] Compared to pharmacies, clinics are 41 percentage points less likely to have a qualified person working there; drug shops are 45 percentage points less likely. Clinics and drug shops are more likely to have a license on display in view of the enumerator. Pharmacies also have a substantially larger selection of antimalarial drugs, nearly 7 more antimalarial drug options than clinics or drug shops. Although establishment type does signal to customers specific characteristics of the outlet, establishment type does not help customers distinguish high quality drugs from low quality drugs.

Although the simple establishment type classification is not predictive, some more specific characteristics related to the physical appearance of the store may be related to drug quality. We next test whether various facility characteristics are related to the quality rate. These are similar to the characteristics that regulators use in conducting their evaluations of outlets. For example, if vendors do not keep careful records of inventory, then drugs may be left on the shelves for a long time; drugs stored in direct sunlight may be more likely to chemically degrade.

We find no evidence that drug quality is related to poor storage conditions. Many coefficients are not different from zero. Those that are significant go in the "wrong" direction: there is actually a positive relationship between drug quality and some measures of poor storage conditions. Results are in Table 7. Outlets that were observed to either have no system for recording inventory are 31 percentage points more likely to sell counterfeit drugs and no more likely to sell substandard drugs compared to outlets which kept records. A store that is classified as "very crowded" is 4.6 percentage points less likely to sell a substandard drug, but no less likely to sell a counterfeit drug. Stores with a concrete floor (as opposed to a dirt floor) are 16.4 percentage points less likely to sell a counterfeit drug but no less likely to sell a substandard drug. Similarly, outlets where inventory that is stored on the floor–potentially more susceptible to insects, rodents, or ground moisture– are 4.9 percentage points less likely to sell a substandard drug. Although there may be potential issues of multicollinearity, point estimates are approximately the same when characteristics are examined separately in a simple regression; results are also robust to including controls for establishment type. These coefficients are all inconsistent with a story where poor storage conditions indirectly cause low

[30]Price and quality are also not related to whether the mystery shopper classification was the same as the self-reported establishment classification.

quality drugs.

On the drug vendor survey, we also asked about the vendors' typical supply chain processes. Results are in Table 8. Although not always significant, there is a suggestive relationship between outlets who buy their drugs from mobile providers/promoters and drug quality. Outlets who purchase from drug promoters are 12.4 percentage points more likely to sell a substandard drug (p-value = 0.104). This result is robust to including establishment type fixed effects. Furthermore, the lack of statistical significance may reflect a lack of statistical power as opposed to the lack of a relationship. In Columns 5 and 6, we control for district fixed effects as opposed to parish fixed effects, freeing up additional degrees of freedom. In the specification with district fixed effects, outlets purchasing from drug hawkers are 11.8 percentage points more likely to sell a substandard drug, significant at the 10 percent level. These results suggest that substandard drugs enter the supply chains through mobile drug hawkers/promoters.

7.3 Institutional Factors

Table 9 examines whether other characteristics of the establishment, or the operating environment are correlated with drug quality. These results show that monthly profits, and whether the outlet ever reported charging a consultation fee are significantly correlated with drug quality. Less profitable establishments are more likely to sell substandard drugs, but less likely to sell counterfeit drugs. This pattern is consistent with the interpretation that outlets faced with financial difficulty may seek lower-cost inputs, such as drugs provided by drug hawkers. Outlets that charge a consultation fee are nearly 5 percentage points less likely to ever sell substandard drugs. This correlation may reflect a reputation incentive; diagnosing the patient's illness, and charging the patient accordingly, may increase the reputation incentive to provide high quality treatment. In contrast, other characteristics are unrelated with drug quality. For example, whether the outlet has a name posted outside– a potential measure of whether the outlet attempts to avoid regulation– is uncorrelated with drug quality, as is whether the outlet has been inspected by a regulator or government official in the past 6 months.

8 Discussion

One key institutional detail relevant to the interpretation of results is whether vendors are complicit in the sale of low-quality drugs. If shopkeepers knowingly distribute low-quality drugs, then providing education or training is unlikely to improve market quality. Instead, limited resources should be spent on increased monitoring, potentially at drug factories or points of entry, or alternatively enforcement of laws and anti-corruption measures. If providers are also being duped by suppliers or are unaware of degrading drug inventory, then raids and shop closures – the current policy approach – may be unfair, expensive to maintain, and limit healthcare access. We asses whether providers know their quality in two ways. First, we present results from a randomized survey exercise aimed at estimating the fraction in the sample that have knowingly sold a low-quality drug. Second, we put those results in context with existing literature and the pattern of results that we found in our study. We conclude with a discussion of potential policy proposals that would be effective at reducing low quality drug rates, and present evaluations of potential policies from the vendors themselves.

8.1 Do Drug Vendors Know Quality? Results from A Survey Experiment

One way to determine if vendors know the quality of the drugs that they are selling is to simply ask them through direct elicitation. However, direct elicitation may lead to systematic response bias. Vendors may be unwilling to report to an enumerator their culpability in a socially undesirable and illegal activity. Thus, it is expected that direct elicitation would underestimate the fraction of the sample who knowingly To overcome these potential response biases, we conduct a survey exercise known typically as "list randomization" in economics or the "item-count technique" (ICT) in political science [Kuklinski et al., 1997, Karlan and Zinman, 2012].

8.1.1 Details of the Survey Experiment

The list randomization works as follows. Survey respondents are randomly assigned into "Treatment" and "Control" groups. Respondents in the Control group are shown a list of activities and asked to report how many of the activities they had done. Respondents in the Treatment group are shown the same list of activities, plus one sensitive activity, and asked

to report how many activities that they have done. The difference in mean responses between the two groups identifies the proportion of the sample estimated to have done the sensitive activity, without revealing the behavior of any individual respondent. Thus, respondents in theory should feel free to report sensitive behavior truthfully.

Due to its promise in estimating difficult-to-measure population parameters accurately, list randomization has become the subject of substantial recent research in survey methodology and political science. While this technique has been used successfully to measure a wide variety of sensitive activities ranging from voting to racial prejudices, existing work also highlights some limitations.[31] Studies in which list randomization has "failed"–meaning failed to find a larger reporting rate of a sensitive activity than through direct elicitation– typically attribute failure to a poor choice of non-sensitive activities [Droitcour et al., 1991]. If non-sensitive activities are not prevalent enough in the population, or display enough variation, then respondents may not believe their responses are anonymous. Alternatively, if non-sensitive activities are too common, then there is a risk of a ceiling effect, where respondents reporting to have done all activities would in fact reveal their participation in a sensitive activity if they truthfully reported. Additionally, recent research has suggested that list randomization may have a downward bias due to cognitive difficulties in adding up responses, particularly in situations with low levels of human capital [Su, 2015]. Finally, compared to directly asking respondents about sensitive activities, list randomization is less statistically efficient. As potential solutions, **?** , and Blair and Imai [2012] all develop new estimators to improve statistical power and test statistically for design and ceiling effects.

We conduct list randomization to limit these concerns in several ways. First, enumerators were trained extensively on implementing this module over one half-day of the three day enumerator training, and survey protocol was regularly reviewed over the course of fieldwork. Training included reminding the enumerators that respondents were not to announce which activities that they had or had not done. Second, we illustrated how their responses would be confidential through a practice round. Third, we varied the order in which the questions were shown, and also varied the non-sensitive activities in the fake-drug treatment in order to limit out potential design effects. Finally, we ask about multiple sensitive activities to compare the prevalence of selling a fake drug against other known problems in the market as a check on

[31] For a review of the literature on list randomization, see Droitcour et al. [1991] or Wolter and Laier [2014].

whether the methodology worked as intended. The three different sensitive activities: paying a bribe to the regulator (the National Drug Authority, or NDA); selling antibiotics when they knew it was unnecessary; knowingly selling a fake drug.[32] Additional details, sensitivity analysis, and the results of a test of successful randomization are in the Online Appendix.

8.1.2 Results of List Randomization

We begin by following the standard in the literature and reporting mean differences. Comparing simple mean differences suggests that 14 percent of the sample has ever paid a bribe to an NDA agent (p=0.17), 13 percent has ever sold an antibiotic when it was unnecessary (p=0.20), and 3 percent has ever sold a fake drug (p=0.73). None of these differences are statistically significant, suggesting low statistical power.

Next, we follow Holbrook and Krosnick [2010] and test whether certain characteristics interacted with the treatment dummy are significantly correlated with reporting more items. Correlates that are significant when interacted with the treatment dummy are characteristics of those individuals who report engaging in either selling a fake drug, paying a bribe, or selling unnecessary antibiotics. Results in 10 show that there is a significant correlation between having a substandard drug sold at the outlet and the vendor reporting knowingly selling a fake drug. In contrast, there is no effect of having a substandard drug sold at the outlet and the vendor reporting paying a bribe and selling unnecessary drugs. Combined with the (insignificant) point estimate, these results suggest that there is a small percentage of vendors who knowingly sell substandard drugs at their outlet.

The accumulation of evidence, including this paper and Fitzpatrick [2015b], supports this view that vendors who sell low-quality drugs are aware that their drugs are of low quality. Of course, if vendors did not know drug quality, then we would still expect some characteristics to be significantly related by chance alone. However, the most likely mechanism associated with this result– drug degradation on shelves– does not hold up empirically. Furthermore, there are suggestive relationships between quality and selling through drug hawkers and promoters, as well as the prevalence of diverted medicines (i.e., other illicit behavior). Finally, the empirical average of 5.6 percent of outlets which sold a substandard drug to a mystery shopper is within

[32]Note that in implementing this exercise, respondents were presented with lists to read in both English and local languages. The English word used was "fake", but each of the local languages have the same word for counterfeit/fake/substandard.

the 95 percent confidence interval from the list randomization exercise.

These results are consistent with the results of Bennett and Yin [2014] and Bjorkman et al. [2012]. Both of these randomized studies find that competition improves quality– interpreted as vendors strategically responding to competition in order to increase profits. All together, these facts suggest that substandard drugs are caused by a small set of vendors who knowingly dilute their quality to deceive customers. This interpretation is also consistent with the figure we report of 85 percent of vendors suggesting that fake drugs are sold in order to increase profits, because they are lower cost inputs. It is also consistent with vendors reporting that a high fraction of other vendors sell fake drugs.

8.2 How to Improve Drug Quality?

We asked what potential policy interventions the vendors themselves thought would be effective at reducing low-quality drug rates. From a policy perspective, implementing fines and standardizing jail penalties for being convicted of selling low-quality drugs should have an important deterrence effect. Moreover, there is an ongoing debate over penalties for selling fake drugs among lawmakers in Uganda. For those convicted of selling low-quality drugs, the median fine recommended by vendors is 1 million UGX ($386), and the average jail term of nearly 9 years, with a median jail term of 3 years. These penalty recommendations are uncorrelated with quality (not shown).

Respondents were also asked to rank the potential policies on a scale of 1 ("definitely will not work") to 4 ("definitely will work"). The most popular policy intervention that was thought to be most effective was to increase training programs for vendors and/or pharmacists to recognize fake drugs; 97 percent of the sample thought that would be an effective strategy. Overall, 95 percent thought that increased inspections at borders and points of entry would be effective; 91 percent thought that increased inspections at outlets would be effective. Only 87 percent thought that increased fines and 85 percent thought that increased jail sentences would be effective. The relatively least popular policy was consumer education campaigns, although 80 percent of respondents still thought that such a policy would be effective.

9 Conclusion

Recent empirical evidence suggests that low quality healthcare is prevalent in many developing countries. One example of the problem of low quality is the problem of fake, or counterfeit antimalarial drugs. Improving the quality of antimalarial drugs is important to encourage appropriate and effective treatment, reduce medical complications, and improve health. In addition, if counterfeit drugs increase drug resistance of malaria, then reducing their prevalence becomes a public health concern. However, there is little evidence on the prevalence of low quality medicine in Uganda, and little guidance for policymakers on where to target anti-counterfeit policy with scarce resources.

In this paper, we contribute to the existing literature on low quality drugs. We first estimate the average rate, and are among the few to do so with a large, representative sample. We find that the rate of counterfeit drugs is approximately 17 percent, but the rate of substandard drugs– those less likely to be medically effective– is approximately 4 percent. Second, we provide important insight into the mechanisms through which low quality drugs infiltrate markets. By combining unique and detailed datasets on both supply and demand characteristics, we provide new empirical analysis on correlates of low-quality drugs. We find measures of regulation are uncorrelated with drug quality rates, potentially reflective of ineffective institutions in resource-poor environments. However, competition may be a market-based manner in which to improve drug quality. We find that poor storage conditions, and subsequent degradation, are an unlikely cause of substandard medicines. In contrast, mobile drug hawkers appear to be one way substandard medicines enter the supply chain.

We find that customers would have a difficult time discerning low-quality drugs on the basis of price or other characteristics. Vendor behavior changes only slightly when they are selling a low-quality drug. We find low-quality drugs are typically dilutions of high-quality dosages. It is relatively uncommon for vendors to sell a complete dosage of ineffective tablets. In dosages with substandard drugs, on average only half of tablets fail a handheld spectrometer test. We interpret this dilution as strategic behavior: dilution increases noise, and makes it more difficult for customers to learn about true drug quality from personal experience. We find that men are more likely to be sold substandard medicines.

Our final result is that a small percentage of vendors are complicit in the sale of low-quality

drugs. The point estimate on knowingly selling a fake drug is similar to the empirical average, although not significantly different from zero. In addition, other circumstantial evidence–such as responses from the vendors themselves on why an outlet would sell low-quality drugs–suggest that at least some subset of providers have knowledge of their drug quality.

The relatively high rates of drug quality we observe are contrary to recent research based upon small, non-random samples. Ex-ante, it is unclear whether the existing literature represents an upper or lower bound on low-quality drug rates. For example, in rural areas, vendors may personally know their clients and similarly the lack of competition may make reputation motives to sell high-quality drugs very salient. On the other hand, in urban areas, regulation may be higher, and the larger number of people may mask accountability. Our data indicate that although the rate may differ substantially between different geographic areas of the same country, previous work likely overestimates the rate of low-quality drugs.

We fully acknowledge that drug quality rates could be different in different study areas, or among different drug classifications. In the study area, malaria represents the number one burden of illness; results may in particular differ in areas where malaria is less common, or where the local population is less familiar with this illness. Regardless, the fact that our overall results indicate that substandard and counterfeit drugs are substantially less of a problem than previously estimated has important policy implications. While a four percent rate of substandard drugs is not necessarily an "acceptable" rate of drug quality, policymakers must make decisions regarding scarce resources. Our results should help them evaluate where to devote money, manpower, and expertise, in order to maximize their impact among their communities. As such, policymakers may choose to further reduce drug quality rates or instead to focus on other problems of healthcare quality that plague developing countries. Our results also highlight that future research on drug quality should incorporate technology and random sampling methods to maximize accuracy, impact, and relevance to the local population.

References

ACTwatch Group, PACE/Uganda & the Independent Evaluation Team. Outlet Survey Report 2010 (Round 3) for the Baseline Independent Evaluation of the Affordable Medicines Facility (AMFm), Uganda, ACTwatch/PACE/PSI. Technical report, Oct. 2012.

G. Akerlof. The Market for "Lemons": Quality Uncertainty and the Market Mechanism. The Quarterly Journal of Economics, 84(3):488–500, 1970.

J. K. Baird. Effectiveness of Antimalarial Drugs. The New England Journal of Medicine, 352 (15):1565–1577, Apr. 2005.

R. Bate and K. Hess. Anti-malarial drug quality in Lagos and Accra - a comparison of various quality assessments. Malaria Journal, 9(1):157, 2010.

R. Bate, R. Tren, K. Hess, L. Mooney, and K. Porter. Pilot Study Comparing Technologies to Test for Substandard Drugs in Field Settings. African Journal of Pharmacy and Pharmacology, 3(4):165–170, Apr. 2009.

R. Bate, G. Z. Jin, and A. Mathur. Does price reveal poor-quality drugs? Evidence from 17 countries. Journal of Health Economics, 30(6):1150–1163, Dec. 2011.

D. Bennett and W. Yin. The Market for High-Quality Medicine. NBER Working Paper 20091, pages 1–59, Apr. 2014.

M. Bjorkman, J. Svensson, and D. Yanagizawa-Drott. Can Good Products Drive Out Bad? Experimental Evidence from Local Markets for Antimalarial Medicine in Uganda. Stockholm University Working Paper, pages 1–47, Aug. 2012.

G. Blair and K. Imai. Statistical Analysis of List Experiments. Political Analysis, 20(1):47–77, Jan. 2012.

J. Das and J. Hammer. Money for nothing: The dire straits of medical practice in Delhi, India. Journal of Development Economics, 83(1):1–36, May 2007.

J. Das, A. Holla, V. Das, M. Mohanan, D. Tabak, and B. Chan. In Urban and Rural India, A Standardized Patient Study Showed Low Levels of Provider Training and Huge Quality Gaps. Health Affairs, 31(12):2774–2784, Nov. 2012.

J. Droitcour, R. A. Caspar, M. Hubbard, T. L. Parsley, W. Visscher, and T. M. Ezzati. The Item Count Technique as a Method of Indirect Questioning: A Review of Its Development and a Case Study Application. In P. B. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz, and S. Sudman, editors, Measurement Errors in Surveys, pages 185–210. Wiley, New York, Sept. 1991.

A. Fitzpatrick. Does Public Healthcare Provision Spill Over Onto the Private Sector? Evidence from the Ugandan Antimalarial Market. University of Michigan mimeo, pages 1–19, Aug. 2015a.

A. Fitzpatrick. Do Informed Consumers Reduce the Price and Prevalence of Counterfeit Drugs? Evidence from the Antimalarial Market. University of Michigan mimeo, pages 1–66, Apr. 2015b.

M. Gaynor. What Do We Know About Competition and Quality in Health Care Markets? NBER Working Paper 12301, pages 1–42, June 2006.

M. Hajjou, Y. Qin, S. Bradby, D. Bempong, and P. Lukulay. Assessment of the performance of a handheld Raman device for potential use as a screening tool in evaluating medicines quality. Journal of Pharmaceutical and Biomedical Analysis, 74:47–55, Feb. 2013.

A. L. Holbrook and J. A. Krosnick. Social desirability bias in voter turnout reports: Tests using the item count technique. Public Opinion Quarterly, 74(1):37–67, Mar. 2010.

D. Karlan and J. Zinman. List Randomization for Sensitive Behavior: An Application for Measuring Use of Loan Proceeds. Journal of Development Economics, 98(1):71–75, May 2012.

H. Kaur, E. L. Allan, I. Mamadu, Z. Hall, O. Ibe, M. El Sherbiny, A. v. Wyk, S. Yeung, I. Swamidoss, M. D. Green, P. Dwivedi, M. J. Culzoni, S. Clarke, D. Schellenberg, F. M. Fernández, and O. Onwujekwe. Quality of Artemisinin-Based Combination Formulations for Malaria Treatment: Prevalence and Risk Factors for Poor Quality Medicines in Public Facilities and Private Sector Drug Outlets in Enugu, Nigeria. PLoS ONE, 10(5):e0125577, May 2015.

J. Konde-Lule, S. Gitta, A. Lindfors, S. Okuonzi, V. O. Onama, and B. C. Forsberge. Private And Public Health Care In Rural Areas of Uganda. BMC International Health and Human Rights, 10(29):1–8, June 2012.

J. H. Kuklinski, M. D. Cobb, and M. Gilens. Racial Attitudes and the "New South". The Journal of Politics, 59(2):323–349, May 1997.

K. Leonard and M. Masatu. Using the Hawthorne Effect to Examine the Gap Between a Doctor's Best Possible Practice and Actual Performance. Journal of Development Economics, 93(2):226–243, May 2008.

R. Mubiri. NDA Closes Down 100 Drug shops for Illegal Operations. Uganda Radio Network, pages 1–4, June 2011.

A. Mugisa. Uganda bans Indian drug supplier. New Vision, pages 1–4, Aug. 2012.

G. Nayyar, J. G. Bremand, P. N. Newton, and J. Herrington. Poor-quality Antimalarial Drugs In Southeast Asia And sub-Saharan Africa. The Lancet Infectious Diseases, 12(6):488–496, May 2012.

P. N. Newton, S. J. Lee, C. Goodman, F. M. Fernández, S. Yeung, S. Phanouvong, H. Kaur, A. A. Amin, C. J. M. Whitty, G. O. Kokwaro, N. Lindegårdh, P. Lukulay, L. J. White, N. P. J. Day, M. D. Green, and N. J. White. Guidelines for Field Surveys of the Quality of Medicines: A Proposal. PLoS Medicine, 6(3):e52, 2009.

F. Nuwaha. People's Perception Oof Malaria In Mbarara, Uganda. Tropical Medicine and International Health, 7(5):1–9, Apr. 2002.

K. A. O'Connell, H. Gatakaa, S. Poyer, J. Njogu, I. Evance, E. Munroe, T. Solomon, C. Goodman, K. Hanson, C. Zinsou, L. Akulayi, J. Raharinjatovo, E. Arogundade, P. Buyungo, F. Mpasela, C. B. Adjibabi, J. A. Agbango, B. F. Ramarosandratana, B. Coker, D. Rubahika, B. Hamainza, S. Chapman, T. Shewchuk, and D. Chavasse. Got ACTs? Availability, Price, Market Share And Provider Knowledge Of Anti-malarial Medicines In Public And Private Sector Outlets In Six Malaria-Endemic Countries. Malaria Journal, 10(1):326, Oct. 2011.

Office of the Auditor General . Value for Money Audit Report on the Regulation of Medicines in Uganda by National Drug Authority. Technical report, Dec. 2010.

I. Okeke, A. Lamikanra, and R. Edelman. Socioeconomic and Behavioral Factors Leading to Acquired Bacterial Resistance to Antibiotics in Developing Countries. Emerging Infectious Diseases, 5(1):18–27, Jan. 1999.

S. Salop. The Noisy Monopolist: Imperfect Information, Price Dispersion and Price Discrimination. Review of Economic Studies, 44(3):393–406, Oct. 1977.

A. T. Sorensen. Equilibrium Price Dispersion in Retail Markets for Prescription Drugs. Journal of Political Economy, 108(4):833–850, Aug. 2000.

J. Stanback, M. Ottaviani, M. Bekiita, O. Nakayiza, and A. K. Mbonye. Injected with Controversy: Sales and Administration of Injectable Contraceptives in Uganda. International Perspectives on Sexual and Reproductive Health, 37(01):024–029, Mar. 2011.

Y. Su. Direct Questioning or List-Based Questioning: Evidence from a Survey Experiment on Intravenous Infusion Use and Smoking in China. Harvard University Working Paper, pages 1–43, Apr. 2015.

S. Tougher, Y. Ye, J. Amuasi, I. Kourgueni, R. Thomson, C. Goodman, A. Mann, K. Hanson, and ACTwatch Group, PACE/Uganda & the Independent Evaluation Team. Effect of the Affordable Medicines Facility-malaria (AMFm) on the availability, price, and market share of quality-assured artemisinin-based combination therapies in seven countries: a before-and-after analysis of outlet survey data. The Lancet, 380(9857):1916–1926, Dec. 2012.

Uganda Bureau of Statistics. 2002 Uganda Population and Housing Census. Technical report, Oct. 2008.

Uganda Bureau of Statistics, ICF International. Uganda Demographic and Health Survey 2011 [FR264]. Technical report, 2012.

M. Van Vugt, P. Wilairatana, B. Gemperli, I. Gathmann, L. Phaipun, A. Brockman, C. Luxemberger, N. J. White, F. Nosten, and S. Looareesuwan. Efficacy Of Six Doses of Artemether-Lumefantrine (Benflumetol) in Multidrug-Resistant Plasmodium Falciparum Malaria. American Journal of Tropical Medicine and Hygiene, 60(6):936–942, Dec. 1999.

F. Wolter and B. Laier. The Effectiveness of the Item Count Technique in Eliciting Valid Answers to Sensitive Questions. An Evaluation in the Context of Self-Reported Delinquency. Survey Research Methods, 8(3):153–168, Dec. 2014.

K. Xu, D. B. Evans, P. Kadama, J. Nabyonga, P. O. Ogwal, P. Nabukhonzo, and A. M. Aguilar. Understanding The Impact Of Eliminating User Fees: Utilization And Catastrophic Health Expenditures In Uganda. Social Science & Medicine, 62(4):866–876, Feb. 2006.

Figure 1: Project Timeline

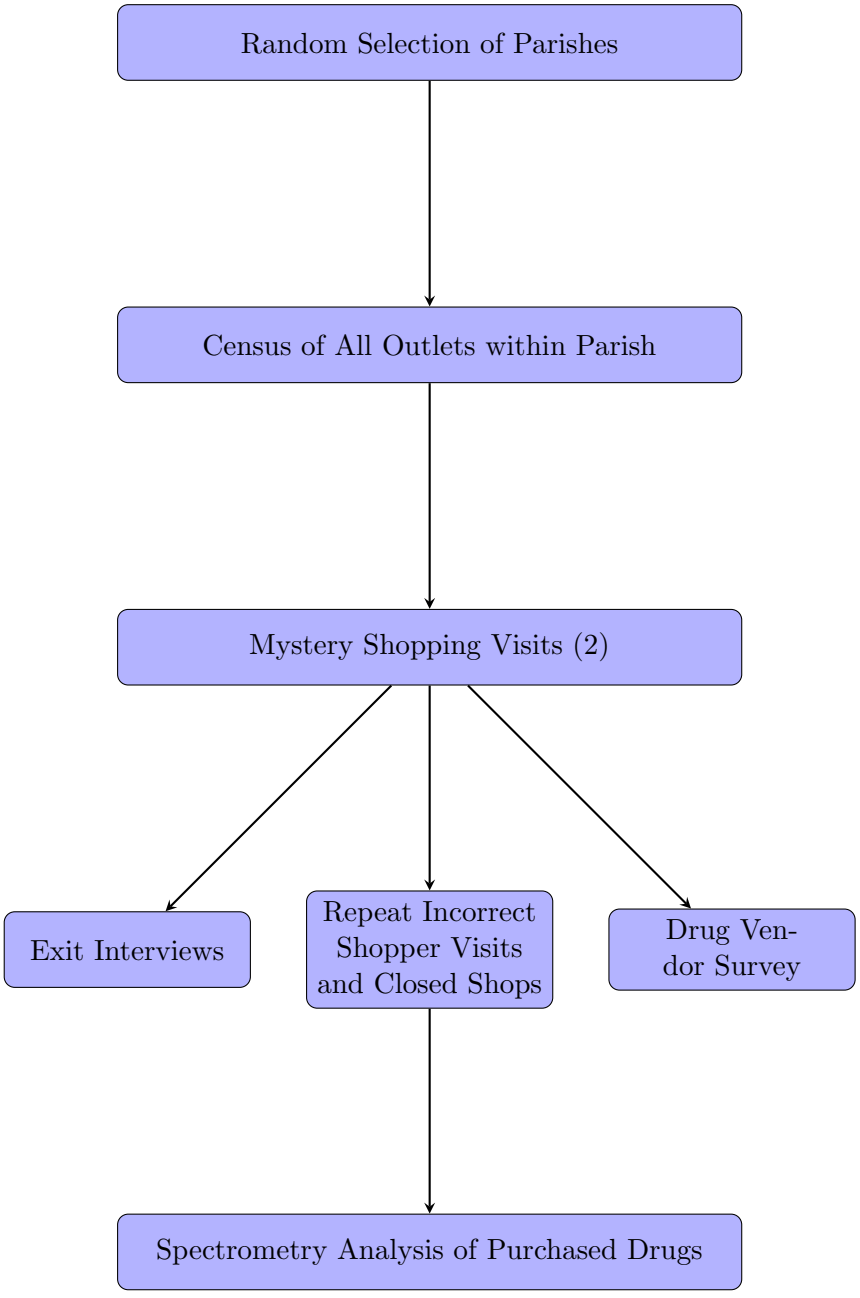
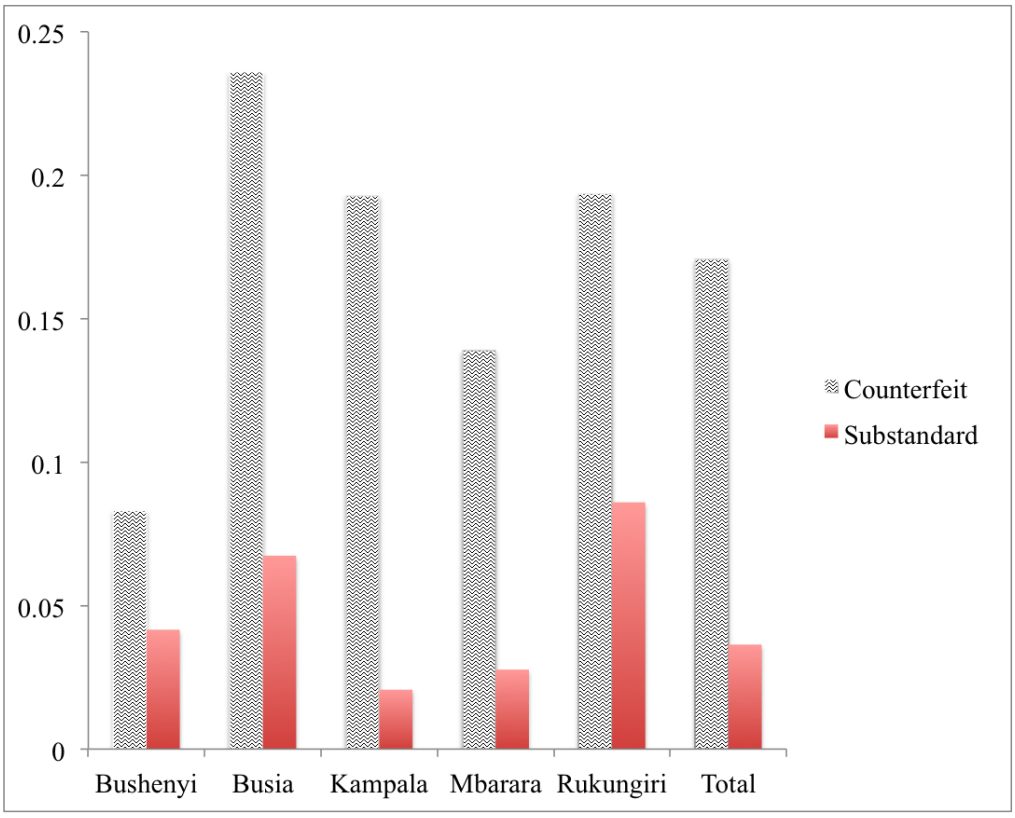
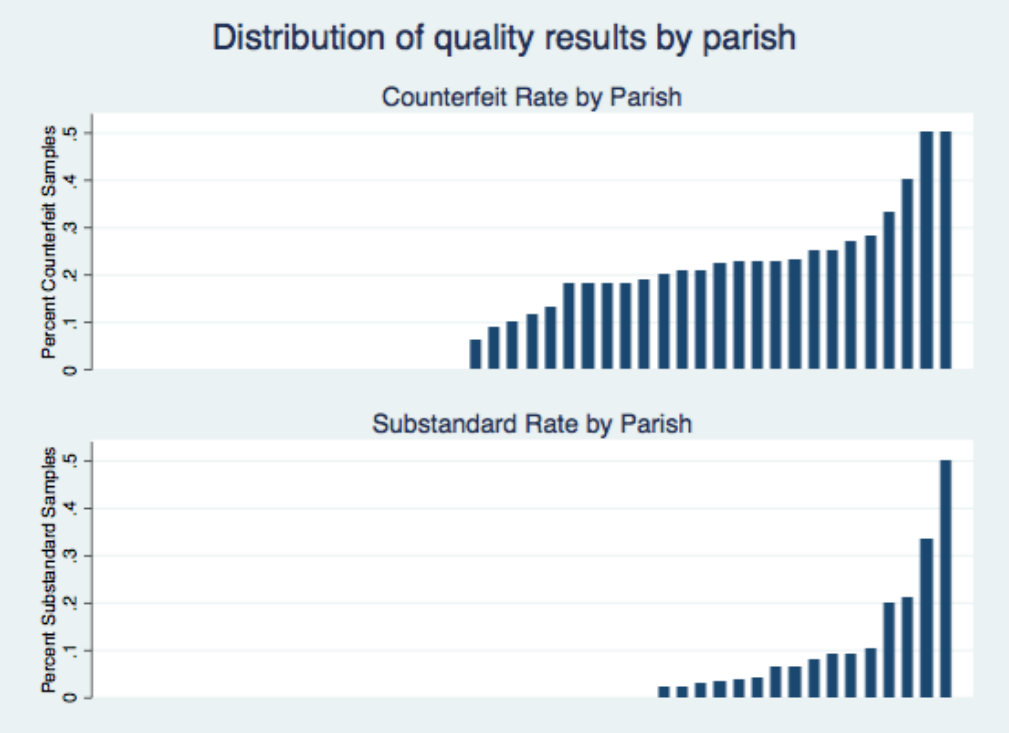


Figure 2: Quality Results by District



Notes: Above are the number of transactions classified as either counterfeit or substandard, averaged across each of the five study districts. Counterfeit refers to whether at least one tablet within the dosage failed the handheld spectrometry test and is estimated in the analysis sample. Substandard refers to whether at least one tablet within the dosage failed the handheld spectrometry test and could not be found to match another brand in the library. Substandard is also only estimated in the analysis sample.

Figure 3: Drug Quality by Parish



Notes: Above are the number of transactions classified as either counterfeit or substandard, averaged across each parish. Counterfeit refers to whether at least one tablet within the dosage failed the handheld spectrometry test and is estimated in the analysis sample. Substandard refers to whether at least one tablet within the dosage failed the handheld spectrometry test and could not be found to match another brand in the library. Substandard is also only estimated in the analysis sample. There was no counterfeit drug sold in forty percent of all parishes, and there was no substandard drug sold in 64 percent of parishes.

Table 1: Vendor Data Summary

Panel A: Vendor and Outlet Characteristics	
Drug Shop	0.53
Clinic	0.40
Pharmacy and Other	0.07
Male Respondent	0.23
Number of employees	2.32
Qualified Person	0.36
Visited by NDA in past 6 months	0.69
Median Profits, USD	77.13
Number of customers, previous day	21.80
Knows first-line treatment	0.84
Percentage of customers who buy full dose	0.66
Sell malaria tests	0.53
Panel B: Why would a vendor sell a fake drug	
Money or Profits	0.85
Ignorance	0.15
Panel C: Beliefs on fake drugs	
Percent outlets fake drugs in parish	0.32
Percent outlets fake drugs in district	0.44
Could identify fake drugs by sight	0.38
Panel D: Customer reaction to selling fake drugs	
Nothing would happen	0.06
Reported to authorities	0.95
Customers would boycott	0.93
Shop would be closed	0.92
Stock would be confiscated	0.88
Customers would do something violent	0.77

Notes: Data are taken from the vendor survey (N=452). Qualified person is a generated variable based upon responses of education level received by the respondent and the type of establishment, based upon the official regulations. Visited by NDA in past six months refers to whether the respondent's outlet had been visited by a representative of the National Drug Authority, the regulator of drug shops, and is conditional on self-reported establishment type being drug shop. Knowing first-line treatment is a dummy variable based upon whether the respondent correctly identified the first-line treatment for malaria. Responses in Panel B are categorized responses to an open-ended question. Could identify fake drugs by sight is a dummy variable of the respondent's answer to whether they thought they could identify a fake drug if they saw it, excluding responses of "don't know". Responses tallied in Panel D are not mutually exclusive.

Table 2: Mystery Shopper Data Summary

Variables	AL (1)	Quinine (2)	SP (3)
Analysis Sample	0.988	0.529	0.810
Number	796	18	64
Number of Different Brands	6	2	3
Correct Dosage	0.948	0.167	0.953
Number Blister Packs	1.383	3.556	1.203
Average Price Paid	3.185	2.607	1.177
Bargained	0.599	0.556	0.531
Coefficient of Variation of Price Paid	0.458	0.398	0.452
Diverted Drug	0.079	0.000	0.063
Counterfeit	0.186	0.111	0.000
Fraction Tablets Counterfeit	0.611	0.544	0.000
Substandard	0.038	0.472	0.000
Fraction Tablets Substandard	0.468	0.417	0.000

Notes: Above is a summary of all data collected by mystery shoppers. 'AL' refers to artemether-lumefantrine. 'Quinine' refers to quinine sulphate. 'SP' refers to sulphadoxine-pyrimethamine. Table excludes 'Other', or other brands and active ingredients, including those that could not be identified. Number of different brands excludes where the brand was unknown. The correct dosage is the full adult dosage. Number blister packs includes any loose tablets sold as their own pack. Average price paid is in US dollars. The exchange rate at the time of data collection was $1US = 2593 UGX. Bargained refers to whether the covert shopper was successfully able to reduce the price through bargaining. Diverted drug means a drug with government markings. Analysis Sample indicates that the drug was able to be part of the analysis sample, and able to be tested. This required obtaining an authentic sample of the brand. Counterfeit refers to whether at least one tablet within the dosage failed the handheld spectrometry test and is estimated in the analysis sample. Substandard refers to whether at least one tablet within the dosage failed the handheld spectrometry test and could not be found to match another brand in the library. Substandard is only estimated in the analysis sample. Both "Fraction of Tablets" variables are conditional on any tablet in the dosage being counterfeit or substandard, respectively.

Table 3: Transaction Correlates

Variables	Counterfeit (1)	Substandard (2)	Counterfeit (3)	Substandard (4)	Counterfeit (5)	Substandard (6)
Number of Blister Packs	0.032*	0.028**	0.036	0.028*	0.034	0.027*
	(0.019)	(0.014)	(0.022)	(0.014)	(0.022)	(0.014)
No expiration date	−0.119**	−0.006	−0.045	0.013	−0.051	0.006
	(0.047)	(0.030)	(0.054)	(0.034)	(0.054)	(0.037)
Diverted Drug	0.267***	0.098**	0.248***	0.094*	0.220***	0.092*
	(0.071)	(0.049)	(0.071)	(0.049)	(0.070)	(0.048)
Price Paid USD	−0.014**	−0.001	−0.024***	−0.003	−0.024***	−0.003
	(0.006)	(0.004)	(0.009)	(0.004)	(0.008)	(0.004)
Constant	0.151***	0.021	0.188***	0.029	0.138**	−0.013
	(0.039)	(0.023)	(0.043)	(0.023)	(0.057)	(0.028)
Drug Type Fixed Effects			X	X	X	X
Mystery Shopper Fixed Effects					X	X
Observations	879	879	879	879	879	879
R-squared	0.093	0.152	0.104	0.154	0.130	0.168

Notes: Sample is all purchases that could be tested for quality that had non-missing values for all variables. Robust standard errors in parentheses. Above are OLS estimates from a linear probability model. "Counterfeit" refers to a purchased dosage in which at least one tablet failed the handheld spectrometry test; "Substandard" refers to a purchased dosage in which at at least one tablet failed the handheld spectrometry test and also did not match any other brands within the library. The outcome variable in odd columns is whether or not at least one tablet within the dosage was counterfeit; the outcome variable in even columns is whether or not at least one tablet within the dosage was substandard. Number of blister packs is the number of separate packages containing tablets; loose tablets are each counted as their own pack. No expiration date is a dummy variable indicating whether or not an expiration date was visible on the package. Diverted drug means that the drug had public sector markets. Price paid is the final transaction price, and is in US dollars. The exchange rate used is \$1US = 2593 UGX. Drug type fixed effect is a set of dummy variables for the drug's active ingredient classification. All regressions control for parish fixed effects and the script used to purchase the dosage. $***p < 0.01, **p < 0.05, *p < 0.1$

Table 4: Vendor Behavior and Subjective Impressions

Variables	Counterfeit (1)	Substandard (2)	Counterfeit (3)	Substandard (4)
Successfully Bargained	0.014	−0.017	0.013	−0.018
	(0.032)	(0.016)	(0.032)	(0.016)
Gave correct instructions	0.028	0.004	0.024	0.001
	(0.033)	(0.012)	(0.034)	(0.012)
Express doubts	0.028	0.016	0.03	0.015
	(0.037)	(0.014)	(0.037)	(0.015)
Ask any questions	−0.092∗	−0.057∗∗	−0.075	−0.060∗∗
	(0.049)	(0.027)	(0.047)	(0.027)
Picked from back			0.02	0.01
			(0.051)	(0.025)
Very Friendly			0.029	−0.022
			(0.052)	(0.028)
Very Unfriendly			−0.100∗∗	−0.045∗∗
			(0.041)	(0.020)
Above Average Quality			−0.057	0.016
			(0.035)	(0.014)
Below Average Quality			0.073	0.078∗∗
			(0.049)	(0.034)
Constant	0.161∗∗∗	0.062∗∗	0.170∗∗∗	0.052∗
	(0.059)	(0.031)	(0.062)	(0.031)
Mystery Shopper Fixed Effects	X	X	X	X
Drug Type Fixed Effects	X	X	X	X
Observations	869	869	869	869
R-squared	0.105	0.139	0.114	0.153

Notes: Above are OLS estimates from a linear probability model. Sample is all purchases of antimalarial drugs that could accurately be tested for quality and excludes missing values. Data are reported by the mystery shopper immediately following the transaction. "Counterfeit" refers to a purchased dosage in which at least one tablet failed the handheld spectrometry test; "Substandard" refers to a purchased dosage in which at at least one tablet failed the handheld spectrometry test and also did not match any other brands within the library. "Successfully bargained" refers to whether the shopper successfully negotiated a price discount. "Drug picked from back of outlet" refers to whether the vendor went to the back of the outlet, or otherwise left the sight of the shopper, to pick the medicine. All regressions control for parish fixed effects and the script used to purchase the dosage. $***p < 0.01, **p < 0.05, *p < 0.1$

Table 5: Mystery Shopper Characteristics

	Parish Fixed Effect		District Fixed Effect	
	Counterfeit	Substandard	Counterfeit	Substandard
Panel A: Full Sample	(1)	(2)	(3)	(4)
Minority Tribe	0.027	−0.027	0.015	−0.02
	(0.039)	(0.018)	(0.037)	(0.015)
Female Shopper	−0.045	−0.018	−0.04	−0.02
	(0.028)	(0.013)	(0.026)	(0.013)
Constant	0.181***	0.090***	0.182***	0.083***
	(0.033)	(0.021)	(0.031)	(0.019)
Observations	879	879	879	879
R-squared	0.055	0.109	0.014	0.030
Panel B: Matched Sample	(1)	(2)	(3)	(4)
Same Tribe	0.066	0.035	0.087**	0.036
	(0.048)	(0.027)	(0.044)	(0.024)
Same Sex	0.004	−0.006	0.005	−0.013
	(0.032)	(0.015)	(0.030)	(0.015)
Constant	0.129***	0.056***	0.126***	0.051**
	(0.037)	(0.021)	(0.036)	(0.020)
Observations	603	603	603	603
R-squared	0.078	0.161	0.025	0.032

Notes: Above are OLS estimates from a linear probability model. The first two columns control for a parish fixed effect, and the second two columns control for a district fixed effect. Panel A uses all purchases that could be tested for quality and Panel B uses purchases from outlets where the dispenser is the same person who completed the survey. Same Tribe and Minority Tribe refer to the tribe of the shopper. Minority tribe is specific to each district's most prevalent ethnic group. Same sex indicates that the gender of vendor is the same as the gender of the shopper. Counterfeit refers to a purchased dosage in which at least one tablet failed the handheld spectrometry test; Substandard refers to a purchased dosage in which at at least one tablet failed the handheld spectrometry test and also did not match any other brands within the library. Robust standard errors in parentheses, clustered at the outlet level. All regressions control for parish fixed effects and the script used to purchase the dosage.$***p < 0.01, **p < 0.05, *p < 0.1$

Table 6: Drug Quality and Establishment Type

Variables	Sold Counter-feit (1)	Sold Sub-standard (2)	Sold Counter-feit (3)	Sold Sub-standard (4)	Average Shop Price (5)	Qualified Person (6)	License on Display (7)	#Antimalarials Stocked (8)
Clinic	0.075	−0.012	0.069	−0.026	0.718***	−0.409***	0.261***	−6.781***
	(0.091)	(0.048)	(0.099)	(0.058)	(0.184)	(0.095)	(0.065)	(0.696)
Drug shop	0.021	−0.036	0.032	−0.039	0.357**	−0.448***	0.461***	−6.872***
	(0.095)	(0.051)	(0.103)	(0.061)	(0.160)	(0.105)	(0.083)	(0.792)
Constant	0.210**	0.081*	0.212**	0.088	2.545***	0.739***	0.015	11.753***
	(0.086)	(0.046)	(0.093)	(0.056)	(0.139)	(0.090)	(0.062)	(0.693)
Observations	459	459	405	405	405	405	405	405
R-squared	0.078	0.083	0.09	0.073	0.259	0.2	0.189	0.438
P-Value of Joint F-test	0.496	0.677	0.693	0.807	0.000	0.000	0.000	0.000

Notes: Sample in first two columns is all outlets where a drug was purchased and could be tested for quality. Sample in remaining columns is all outlets at which there was a survey completed, and excludes missing values/don't know/don't recall. "Clinic"/"Drug shop" is recorded from the census; the omitted category is "Pharmacy/Other". "Sold Counterfeit" refers to whether a drug classified as counterfeit was ever sold from that outlet during mystery shopping. "Sold Substandard" refers to whether a drug classified as substandard was ever sold from that outlet during mystery shopping. "Average Shop Price" is the average price paid during mystery shopping among all purchases. "Qualified person" is whether the respondent was estimated to meet the legal qualifications to dispense medicines. It is assessed from responses based upon years of experience, education, and type of establishment. "License on Display" was whether the enumerator marked that during the vendor survey there was a license on display. "No. Antimalarials in Stock" refers to the number of antimalarials listed on the drug inventory portion of the drug vendor survey. "P-Value of Joint F-test" refers to the p-value of the F-test that Clinic and Drug Shop are jointly zero. All regressions include parish fixed effects. Robust standard errors in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$

Table 7: Facility Observations

Variables	Counterfeit Sold (1)	Substandard Sold (2)	Counterfeit Sold (3)	Substandard Sold (4)
No System	0.344***	−0.025	0.347***	−0.022
	(0.101)	(0.025)	(0.102)	(0.026)
All inventory protected from sunlight	0.055	0.034*	0.046	0.032
	(0.069)	(0.018)	(0.070)	(0.020)
Crowded store	−0.023	−0.038	−0.019	−0.040
	(0.062)	(0.025)	(0.062)	(0.025)
Concrete floor	−0.162*	−0.017	−0.180**	−0.009
	(0.092)	(0.046)	(0.092)	(0.047)
Establishment was very clean	−0.017	−0.018	−0.008	−0.023
	(0.057)	(0.030)	(0.058)	(0.030)
Inventory on floor	−0.117	−0.056*	−0.129	−0.056*
	(0.083)	(0.029)	(0.086)	(0.030)
Drug Shop			0.048	−0.041
			(0.111)	(0.064)
Clinic			0.089	−0.031
			(0.106)	(0.058)
Constant	0.353***	0.060	0.314**	0.09
	(0.113)	(0.048)	(0.145)	(0.070)
Observations	408	408	408	408
R-squared	0.128	0.084	0.131	0.085

Notes: Sample is all outlets at which there was a survey completed, and excludes missing values/don't know/don't recall. Data are from the observations of the enumerator after the conclusion of the survey. "Manual or no record-keeping ssystem" refers to whether the enumerator observed a manual inventory and record-keeping system, or no system. This variable is a zero for hybrid systems or computerized record-keeping. "All inventory protected from sunlight" is a dummy variable indicating that the entire stock of inventory was not exposed to direct sunlight. "Crowded store" is whether the enumerator judged that the store had adequate space to display its inventory. "Concrete floor" is a variable indicating whether the floor was concrete, as opposed to a dirt floor. "Establishment was very clean" is a subjective measure by the enumerator about the cleanliness of the outlet. "Inventory stored on floor" was whether the enumerator noted that there were boxes of inventory stored on the floor of the outlet, as opposed to shelves or cabinets. All regressions include parish fixed effects. Robust standard errors in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$

Table 8: Supply Networks

	Counterfeit Sold	Substandard Sold	Counterfeit Sold	Substandard Sold	Counterfeit Sold	Substandard Sold
Ever bought stock from:	(1)	(2)	(3)	(4)	(5)	(6)
Drug Promoter	0.077	0.122	0.079	0.122	0.078	0.117∗
	(0.109)	(0.075)	(0.108)	(0.075)	(0.093)	(0.069)
Wholesale Pharmacy	−0.101	−0.048	−0.102	−0.042	−0.001	−0.009
	(0.141)	(0.067)	(0.143)	(0.071)	(0.125)	(0.063)
Retail Pharmacy	0.102	−0.028	0.102	−0.023	0.087	−0.032
	(0.063)	(0.023)	(0.064)	(0.024)	(0.061)	(0.024)
Drug Shop			0.011	−0.025	0.018	−0.035
			(0.097)	(0.061)	(0.086)	(0.052)
Clinic			−0.015	−0.045	0.015	−0.041
			(0.092)	(0.059)	(0.082)	(0.053)
Constant	0.333∗∗	0.098	0.333∗∗	0.122	0.224	0.097
	(0.141)	(0.066)	(0.160)	(0.079)	(0.144)	(0.072)
Parish FE	X	X	X	X		
District FE					X	X
Observations	415	415	415	415	415	415
R-squared	0.092	0.09	0.092	0.093	0.016	0.031

Notes: Data are taken from survey responses to the drug vendor survey where a purchase was made that could be tested (N=416). "Sold Counterfeit" refers to whether a drug classified as counterfeit was ever sold from that outlet during mystery shopping. "Sold Substandard" refers to whether a drug classified as substandard was ever sold from that outlet during mystery shopping. Each variable is a dummy variable indicating whether the respondent ever purchased stock from a given type of wholesaler. Multiple responses were allowed. Regressions control for either parish or district fixed effects. Robust standard errors in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$

Table 9: Other Outlet Characteristics

Variables	Counterfeit Sold (1)	Substandard Sold (2)	Counterfeit Sold (3)	Substandard Sold (4)
No name visible	0.058	0.003	−0.027	−0.025
	(0.068)	(0.039)	(0.056)	(0.032)
Any agency visit	0.017	−0.049	0.039	−0.04
	(0.068)	(0.045)	(0.059)	(0.040)
Profits, USD	0.018∗	−0.009∗	0.019∗∗	−0.008∗
	(0.009)	(0.005)	(0.008)	(0.005)
Business training program	−0.035	0.012	−0.016	0.021
	(0.054)	(0.031)	(0.049)	(0.028)
Any debt	0.086	0.005	0.084	0.002
	(0.063)	(0.033)	(0.059)	(0.030)
Number of employees	0.012	0.002	0.018	0.004
	(0.014)	(0.007)	(0.012)	(0.006)
Have beds	−0.040	−0.026	−0.037	−0.049
	(0.085)	(0.045)	(0.076)	(0.040)
Test for malaria	−0.068	0.049	−0.086	0.048
	(0.063)	(0.034)	(0.057)	(0.031)
Charge consultation fee	0.050	−0.047∗	0.049	−0.049∗
	(0.080)	(0.029)	(0.075)	(0.026)
Constant	0.115	0.125	0.052	0.122∗
	(0.132)	(0.082)	(0.113)	(0.073)
Establishment Type Fixed Effect	X	X	X	X
Parish Fixed Effect	X	X		
District Fixed Effect			X	X
Observations	391	391	391	391
R-squared	0.11	0.097	0.043	0.041

Notes: Sample is all outlets with a completed survey. "Counterfeit Sold" is whether a counterfeit drug was sold at that outlet during mystery shopping. "Substandard Sold" is whether a substandard drug was sold from that outlet. "No name visible" indicates whether there is a printed sign. "Any agency visit" is whether any regulator, including the NDA, had visited over the past 6 months. "Profits" are self-reported monthly profits, reported in 1000 USD. The exchange rate is $US1 = 2593 UGX. "Any debt" indicates the establishment has debt; "Debt question non-response" indicates vendor non-response. "Have beds" indicates whether the establishment has beds to treat patients; "Test for malaria" indicates whether testing is available at the outlet; "Charge consultation fee" is whether the respondent reported that the outlet ever charged a consultation fee to diagnose illnesses. Robust standard errors in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$

Table 10: Effects of Demographics on List Randomization Results

Variables	Sold a Fake Drug (1)	Paid a Bribe (2)	Sold Extra Antibiotics (3)
Treatment List	−0.688	−0.759	−0.484
	(0.645)	(0.616)	(0.765)
Substandard Sold	−0.439∗	0.183	−0.014
	(0.241)	(0.264)	(0.334)
Substandard Drug Sold*Treatment List	0.940∗∗∗	−0.15	0.051
	(0.335)	(0.492)	(0.427)
List Randomization Form Dummy	−0.375∗∗∗	−0.144	−0.189
	(0.132)	(0.375)	(0.442)
List Randomization Form Dummy*Treatment List	0.897∗	0.963∗	0.294
	(0.496)	(0.528)	(0.597)
District =Busia	−0.614	−1.067∗∗	−0.743
	(0.379)	(0.440)	(0.542)
District=Kampala	0.366	−0.783∗	−0.604
	(0.343)	(0.415)	(0.516)
District=Mbarara	−0.077	−0.710∗	−0.435
	(0.302)	(0.371)	(0.486)
District=Rukungiri	−0.811∗∗	−0.866∗∗∗	−0.405
	(0.350)	(0.243)	(0.340)
Busia*Treatment List	1.110	1.433∗∗	0.557
	(0.691)	(0.691)	(0.808)
Kampala*Treatment List	0.479	0.821	0.686
	(0.657)	(0.637)	(0.782)
Mbarara*Treatment List	0.458	0.796	0.773
	(0.617)	(0.576)	(0.736)
Rukungiri*Treatment List	0.216	−0.298	0.038
	(0.513)	(0.437)	(0.564)
Constant	2.452∗∗∗	3.654∗∗∗	2.874∗∗∗
	(0.330)	(0.404)	(0.504)
Observations	412	412	412
R-squared	0.173	0.058	0.034

Notes: SNotes: Robust standard errors in parentheses. Above are OLS estimates from a linear probability model of a list randomization exercise. The dependent variable in all columns is the number of activities that the respondent has reported doing. The treatment list group was shown the same list of non-sensitive activities as the control group, plus one sensitive activity. The sensitive activity in Column 1 is "ever sold a fake drug"; the sensitive activity in Column 2 is "ever paid a bribe to a regulator (NDA)"; the sensitive activity in Column 3 is "ever sold antibiotics to a customer when they knew it wasn't needed". "Treatment List" indicates whether the respondent was randomly assigned to see the longer list of activities. $***p < 0.01, **p < 0.05, *p < 0.1$

A Table Appendix

Table A1: Spectrometry Testing Transition Matrix

Panel A: Tablet Level (N=2322)		
	Second Scan Outcome	
First Scan Outcome	Pass	Fail
Pass	0.297	0.007
	N = 689	N=16
Fail	0.177	0.519
	N = 411	N=1206

Panel B: Transaction Level (N=879)		
	Second Scan Result	
First Scan Result	Non-Counterfeit	Counterfeit
Non-Counterfeit	0.76	0.00
	N=665	N=0
Counterfeit	0.07	0.18
	N=59	N=155

Notes: In total, there were 879 purchases of antimalarial drugs which could be tested using the handheld spectrometer. A full adult dosage of artemether-lumefantrine contains 24 tablets, a full adult dosage of sulphadoxine-pyrimethamine contains 3 tablets, and a full dosage of quinine contains 30 tablets. There were 23,083 tablets scanned with the handheld spectrometer, and 2,322 were scanned at least twice. In Panel A, the sample is restricted to all tablets that were scanned twice. Each cell represents the marginal distribution/probability of passing or failing the first scan, and then the second scan. In Panel B, the sample is all purchases. Each cell represents the marginal distribution of results based upon the first scan of each tablet, and the second scans of each tablet (if multiple scans were performed).

Table A2: Mystery Shopper and Matched Data Summary

Panel A: Shopper Characteristics (N=16)	Average
Age	34.14
Number of Visits	56.44
Female	0.50
Tribe:	
Banyankole	0.438
Bakiga	0.188
Konzo	0.188
Baganda/Other	0.188

Panel B: Matched Sample (N=603)	Average
Dispenser is a Female	0.783
Shopper is Same Tribe as Dispenser	0.186
Shopper is Same Sex as Dispenser	0.537
Shopper is a Female Shopper	0.572
Shopper was a Minority Tribe	0.504
In Kampala	0.380

Notes: Panel A are self-reported data by the 16 mystery shoppers. Banyankole and Bakiga are the predominant tribes of the Western region. In instances where shoppers were of more than 1 tribe, the mother's tribe was used. Number of visits refer to the number of times that a script was recited to a vendor. Panel B contains data at the transaction level from the matched sample where the dispenser was the individual who answered the vendor survey (N=603). This sample is 85% of all transactions. Minority tribe refers to whether the dispenser is of a minority tribe in that district. Note that shoppers were not randomly assigned to each shop.