Welfare Gain from Using Diagnosis Contingent Incentive Contracts to Improve Malaria Treatment*

Maria Dieci † Paul Gertler‡ Jonathan Kolstad§ Carlos Paramo¶

March 26, 2026

Abstract

We examine a novel diagnosis contingent incentive contract designed to improve the treatment of malaria. The contract provides incentives to use rapid diagnostic tests (RDTs) to diagnose patient malaria status combined with incentives to treat with antimalarial drugs (ACTs) if the patient tests positive but not if they test negative. Using data from a cluster randomized field experiment with 140 pharmacies in Kenya, we find that both patient subsidies and provider incentives significantly increased RDT testing uptake and reduced the purchase of unnecesary ACTs by those whose test negative. Patient incentives operate through lower prices, whereas provider incentives work through providers giving better information and advice. Using a model of patient choice, we estimate that diagnosis-contingent contracts increase social welfare substantially relative to program costs, where the primary gain in welfare comes from a reduction in the use of ACTs from patients who test negative and therefore do not need treatment. Finally, we use the experiment to estimate a structural model that allows us to explore counterfactual contract designs. We find that the biggest welfare gains per unit of program cost come from contracts that load all of the incentives in reducing the price of treatment contingent on testing positive for malaria due to biased patient beliefs. The optimal contract delivers 4 times the welfare gains of a policy that simply offers free testing while keeping expected costs to the planner below the retail price of tests.

*We thank William Dow, Lia Fernald, Amrita Ahuja, Stefano Bertozzi, Stephanie Bonds, Drew Cameron, Jessica Cohen, Ada Kwan, Eva Lyubich, Edward Miguel, Jacob Odipo, Bernhards Ogutu, and Prashant Yadav for helpful comments. Josh Morrow, Ruth Namai, Veronica Njeri, and Jessica Vernon from Maisha Meds where instrumental to design and implementation. Rita Cuckovich, Andrew Murithii, Salome Omondi, Carolyne Nekesa and Nettah Isavwa, provided excellent research assistance and field management. This study was approved by UC Berkeley Committee for the Protection of Human Subjects, Strathmore University Ethics Review Board in Kenya, and the Kenyan National Commission for Science, Tehcnology and Innovation. The study was registered at the AEA (AEARCTR-0004705). It was funded by grants from BMGF and USAID DIV. The authors declare that they have no material or financial interests in the results in this paper and that all errors are our own.

†mdieci@emory.edu; Emory University Rollins School of Public Health

‡gertler@berkeley.edu; University of California Berkeley, Haas School of Business and NBER

§kolstad@berkeley.edu; University of California Berkeley, Haas School of Business and NBER

¶carlos.paramo@stanford.edu; King Center on Global Development

1 Introduction

Widespread wasteful spending on low-value medical services represents a central challenge for healthcare systems, potentially exposing patients to unnecessary risks and introducing negative externalities such as antimicrobial resistance. Over- and under-treatment are ubiquitous worldwide with serious implications for both health care spending and health outcomes (Das and Hammer 2014; Whitehead, Dahlgren, and Evans 2001). For example, studies in both the US and China have found excessively high levels of antibiotic overuse.[1] Similarly, studies in sub-Saharan Africa have found high levels of antimalarial drug use among patients who are malaria negative and low levels among patients who are malaria positive.[2] Overuse is not limited to medications; providers have also been shown to increase their use of costly diagnostic tools, such as MRIs, when payment structures make it profitable, leading to more expensive and often unnecessary care (Baker 2010; Afendulis and Kessler 2007).

Much of the innovation in payment mechanisms over the last 50 years has been designed to curb unnecessary spending. However, these payment mechanisms often act as a blunt instrument; largely limited to either discouraging or encouraging services for everyone as opposed to increasing care for patients with high marginal value of treatment and discouraging care for low marginal value patients. The fact that payments only target patients with the average marginal benefit of treatment may help explain why providers often sell large amounts of medically unnecessary services. For example, insurance contracts that paid providers incentives for treatment encouraged low marginal value care (Manning et al. 1986). Fee-for-service and fixed-payment models also led to unnecessary expenditures on treatment (Gruber and Owings 1994; Baker 2010). On the patient side, unconditional subsidies that lower co-payments for malaria treatment encourage unnecessary care and discourage diagnostic tests (Cohen, Dupas, and Schaner 2015)[3].

Since diagnostic effort and testing are costly, payment mechanisms that pay for treatment regardless of diagnostic information or pay for bundled diagnosis and treatment packages can lead to the under-provision of diagnostic information. In these cases, an approach to prevent unnecessary treatment without also discouraging necessary treatment is to design payment mechanisms that generate diagnostic information that is then built into treatment incentives. We propose a novel "diagnosis-contingent" incentive structure that motivates di-

1. See Iizuka 2012; Currie, Lin, and Zhang 2011; Currie, Lin, and Meng 2014; Chen, Gertler, and Yang 2016; Fleming-Dutra et al. 2016; Daniels et al. 2019; Sulis et al. 2020; King et al. 2022.

2. See Ansah et al. 2010; Cohen et al. 2013; Cohen, Dupas, and Schaner 2015; O'Meara et al. 2016.

3. Note that there are cases where encouraging treatment alone might be socially optimal, for example, mass-distribution of deworming medications to children in high prevalence areas where the benefits of treatment are large and the costs are low, and where increasing diagnostic capacity can be prohibitive to program implementation (Miguel and Kremer 2004)

agnostic effort through both direct incentives for testing and indirect incentives for treatment contingent on a positive diagnosis. This design takes into account how incentives influence both diagnosis and treatment decisions, and has the potential to steer demand towards the optimal level of diagnosis to inform treatment choices.

We study "diagnosis-contingent" incentive contracts for malaria. The contract provides incentives to use rapid diagnostic tests (RDT) to determine whether malaria parasites are present, and then provides additional incentives to treat malaria using front-line antimalarial drugs only if the patient tests positive for malaria.[4] This contract design is better able to target treatment to malaria positive patients and away from malaria negative patients by incorporating information from the diagnosis stage into the incentives for the treatment stage.

Malaria is an important disease for studying clinical decision-making because it is a well-understood illness, has a high disease burden, and nearly all deaths and serious illness are preventable through effective and inexpensive medication (WHO 2021). Despite RDTs being cheap and readily available, relatively few patients presenting with malaria symptoms are diagnosed with RDTs or other parasitic tests prior to getting treated.[5] This may explain why large shares of large shares of malaria-negative patients receive antimalarial medication while some malaria-positive patients go untreated (Cohen et al. 2013; Cohen, Dupas, and Schaner 2015; O'Meara et al. 2016; Ansah et al. 2010).

We begin by presenting a model of consumer choice that considers how incentives paid directly to patients influence the demand for malaria testing. We then extend the model to consider how contracts that instead pay providers can increase testing uptake. If providers simply pass through discounts, the results are equivalent to demand-side incentives. However, this standard result ignores another key channel; that providers inform patients about their illness and advise them on treatment. Provider incentives can induce costly effort to inform and advise patients.

We use the model to derive a measure of social welfare that is the cost-effectiveness of increasing testing weighted by the welfare gain from a one unit increase. The social welfare measure provides a useful lens for studying the design of contracts. An immediate

4. RDTs are highly accurate tests that use a finger prick to confirm the presence or absence of malaria parasites in a patient's blood. Mfuh et al. 2019

5. The literature identifies several potential reasons as to why diagnostic testing is low. Patients may not demand tests because of (i) strong prior beliefs about their malaria status – i.e., a low perceived value of information from testing (Maffioli et al. 2019), (ii) the cost of the test is prohibitive (Cohen, Dupas, and Schaner 2015; O'Meara et al. 2018), and (iii) they do not want to wait for the diagnostic test result – i.e., impatience. Moreover, providers may not prescribe a test prior to treatment because (i) they have established practices of symptom-based diagnosis – i.e., established norms and habits (Mbonye et al. 2013), (ii) they are optimizing perceived patient preferences (Lopez, Sautmann, and Schaner 2022), and (iii) they have profit motives (Currie, Lin, and Meng 2014).

result is that diagnosis-contingent contracts are not necessarily always welfare-enhancing. Social welfare depends on the relative costs of treatment and testing, the beliefs of patients (subjective probability) that they are infected relative to the true positivity rate of malaria (disease), and any internalities and externalities of unnecessary treatment.

We use the model to study the optimal design of these contracts. We consider three margins: (1) whether to target incentives to providers or patients (2) how to distribute the incentives between diagnostic testing and treatment conditional on a positive test result, and (3) to determine the optimal level of the subsidy. We show that contracts that reduce the price of treatment conditional on a positive test result produce more social welfare than contracts that simply reduce the price of testing when patients beliefs are biased (i.e., overestimate their probability of being malaria-positive) and the cost of testing is low relative to the cost of treatment. In this case, patients respond more to the diagnosis-contingent treatment discount and the costs of the contracts are lower because the actual probability of being malaria-positive is low.

We test model predictions using a cluster-randomized field experiment (RCT) to estimate the effect of the diagnosis-contingent incentives on malaria testing and treatment decisions, as well as on welfare. We conducted the experiment in high malaria prevalence counties in Kenya, where more than 3.5 million people fall ill with malaria each year.[6] In this population, at least 60% of patients who purchase antimalarial drugs are malaria negative, which means that most of the treatment is an unnecessary expenditure and represents a high baseline level of drug overuse.

The field experiment randomized 140 pharmacies into either a status quo control group or one of three treatment groups.[7] The treatment groups are: (1) patient subsidies for RDT tests and for the anti-malarial drugs (ACTs) conditional on a positive test; (2) pharmacy performance incentives for RDT tests, and for prescribing ACTs conditional on a positive test; and (3) combined incentives (patient subsidies and pharmacy incentives) for RDT and ACTs for confirmed malaria-positive cases. The total value of the incentives was held constant in the three intervention arms. This design allowed us to evaluate the impact of the diagnosis-contingent incentive structure and the effect of targeting that incentive to the patient versus the provider.

Both, the experiment and the model, are important in assessing the optimal design of the contracts. The RCT allows us to empirically test the effectiveness of the specific contract

6. The study population lives near Lake Victoria and on the coast (Initiative 2021; Disease Control and Prevention 2018).

7. Over half of malaria patients in Kenya and across East Africa access treatment via pharmacies, often the preferred access point for primary care given pharmacies' convenience and reliable presence even in areas that are under-served by public health care clinics and hospitals (Musuva et al. 2017; Burton et al. 2011).

designs from the treatment arms, while the model allows us to study the optimal design of these contracts by producing counterfactual interventions.

We find that both patient subsidies and provider incentives are effective in increasing RDT uptake and in improving targeting treatment to malaria-positive patients. Patient subsidies increase the likelihood that patient takes a RDT test by 24 percentage points over a control group rate of 8 percent.[8] The impact of pharmacy incentives is 19 percentage points and not statistically indistinguishable from the estimate of patient subsidies. The incentives also led to a 12 percentage point decline in the use of ACTs. This decline in treatment demand is explained by patients who tested as a result of the incentives and found out that they were malaria negative since a negative test result reduces the propensity to purchase antimalarials from 100% to 5% in our sample. Therefore, these contracts are effective in reducing unnecessary treatments by malaria negative patients while continuing to offer access to appropriate care for malaria positive patients.

In terms of mechanisms, we find that patient subsidies resulted in significantly lower RDT prices. In contrast, none of the provider incentives were passed through to patients in terms of lower prices. Instead, provider incentives were associated with pharmacists providing more explanation of RDT results and counseling on treatment based on the test results. Both contracts have the same impact on ultimate demand for both testing and ACTs, but financial incentives seemed to work through an information and advice pathway when targeted to providers, whereas demand subsidies directly induced more patients to purchase RDTs by lowering prices. In both cases, the policies led to an increase in accurate illness status information leading to more appropriate use of ACTs.

The diagnosis-contingent incentive contracts studied in the field experiment are very cost-effective from a social welfare perspective. We estimate that for each dollar spent on the diagnosis-contingent contract, patient and provider welfare increases by at least $1.32, implying a rate of return of at least 32% in terms of social welfare.[9] The welfare gains are mostly the result of reductions in expenditures on unnecessary drugs by malaria-negative patients. We also find that patient subsidies appear superior to provider incentives, although differences are not statistically significant.

Finally, we return to the model to evaluate the optimal design of diagnosis-contingent contracts. We derive a structural demand function that has as arguments the elements of the contract. We then use the data from the field experiment to identify and estimate the demand function, and conduct counterfactual analysis to compare the welfare effects

8. This result is consistent with what has been found in prior literature on consumer subsidies for RDTs and other health products.(Dupas 2014; Cohen, Dupas, and Schaner 2015)

9. These estimates are lower bounds since they exclude any potential externalities and internalities from unnecessary treatment (e.g. side effects, antimalarial resistance, etc...).

of alternative contract designs. We find that patients respond very strongly to conditional incentives that reduce the cost of treatment, suggesting that patients over-estimate their probability of having malaria. Contracts that load all of the incentives into reducing the cost of treatment conditional on being malaria positive dominate the mixed experimental contract and have a rate of return in terms of social welfare that is almost twice that of loading all of the incentives to reducing the price of testing (RDT). This result is driven by highly biased patient beliefs that they are malaria positive relative to the true probability.

In terms of optimal levels of incentives, we find that the contract that maximizes total welfare (at the lowest possible cost to the social planner) offers a diagnostic-contingent incentive to patients with expected costs to the social planner that are roughly equal to the retail price of tests. This contract delivers 4 times the welfare gains of a policy that simply offers free testing. In order to achieve these welfare gains, we estimate that diagnosis-contingent incentives for treatment should be higher than what we tested experimentally. We also find that in our context, diagnosis-contingent contracts always dominate unconditional price reductions regardless of the per-patient budget of the social planner.

This paper relates to the literature on performance pay that rewards providers either directly through bonuses tied to services provided or indirectly through supplements to capitation that have a similar structure.[10] Payments are typically made based on services provided, and do not explicitly take into account diagnostic information when setting copays for various services. This fits in with other literature on performance pay based on outcomes, including applications for medical doctors and teachers (Campbell et al. 2009; Prendergast 1999; Podgursky and Springer 2007). Financial incentives are well-established tools used around the world to promote a wide range of health behaviors.

This paper also contributes to the literature on demand-side incentives that operate based on the assumption that either price itself is a barrier to adopting a health behavior, or an incentive can nudge people to overcome other non-pecuniary barriers. Demand subsidies for preventive health visits, for example, incentivize utilization by lowering the cost of care to patients. In the US, insurance products that have modest copays or deductibles operate in the same way - by lowering the price patients pay. These examples, as well as other demand-side incentives for healthcare, highlight how these financing models typically operate - they reimburse a flat rate for services used through lower prices/copays (Arrow 1963; Pauly 1980; McGuire 2000; Cutler and Zeckhauser 2000). In addition, price experiments for health treatments have shown that people do not respond uniformly to prices, and instead the nature of the health decision and timing of the benefits affect demand elasticity (Dupas and

10. See, for example, Basinga et al. 2011; Gertler, Giovagnoli, and Martinez 2014; Ahmed et al. 2023; Yip et al. 2014; Peabody et al. 2014; Witter et al. 2012; Miller and Babiarz 2013.

Miguel 2017; Dow, White, and Bertozzi 2016; Dupas 2014; O'Meara et al. 2016).

While supply and demand incentives are typically studied independently, we compare them directly and put the two literatures in conversation with each other. In addition, we contribute to these two strands of literature by innovating in how health financing contracts are structured in terms of separating incentives for diagnosis and treatment, and linking the treatment incentive to information learned from diagnosis.

Third, this paper adds to the literature on how personalized health information and financial incentives can be combined to change health behavior. Information combined with financial incentives has shown more promise in encouraging health behavior adoption (Meredith et al. 2013; Ma et al. 2014; Dupas 2011). But, the quality of the information matters: general health information tends to be less effective in changing individual behaviors than individually tailored messages targeted at the key decision-makers (Gong 2015). Studies that have examined whether information provided by a malaria diagnostic test changes treatment-seeking behavior have found mixed results; information is effective in steering some patients towards appropriate treatment options, especially when coupled with an incentive, but many elect to ignore test results when making treatment decisions (Cohen, Dupas, and Schaner 2015; O'Meara et al. 2018).

In particular, Cohen, Dupas, and Schaner 2015 studies price subsidies for malarial diagnostic tests and treatments and finds that just under half of patients testing negative elect to receive treatment. The study population in Cohen, Dupas, and Schaner 2015 is at the household level, making important contributions by highlighting the challenges of balancing access to treatment and diagnosis. Our study population are pharmacy patients that have self-selected into seeking treatment, so access to treatment is a lower concern for this group relative to the risk of over-treatment. In addition to the study populations, there are two main differences between our studies that might explain the stark differences in results. First, our study introduces conditional ACT incentives that tie subsidy access to the test result, minimizing incentives for moral hazard from patients offered heavily discounted treatment even if negative. Second, our study took place a decade after their study, so patient trust for test results might have been influenced through learning and the widespread availability of RDTs after the COVID-19 pandemic.

This study leverages the two steps of the testing and treatment decision by providing a financial incentive for treatment conditional on the personalized health information provided by the test. We test the extent to which appropriate malaria treatment use is a result of information (RDT result) or an added financial incentive (for ACT), contributing to the long literature on the role of information and information asymmetries in health decision-making (Arrow 1963).

2 Model

In this section, we develop a simple model to elucidate the rationale for our diagnosis-contingent contract design. We begin by presenting summary statistics of the choices that motivate the model. Consider patients who inquire about malaria care in private pharmacies and must decide whether to test for malaria and purchase treatment. Figure B1 summarizes the conditional probabilities of buying diagnostic tests and treatment.

Among patients who seek malaria care in our treatment arms, 64% decide not to purchase any type of test. Since this population is conditional on inquiring about malaria care in the pharmacy, they likely have symptoms consistent with malaria, and thus, all of those who do not test purchase an antimalarial. Therefore, in the absence of diagnostic information, patients default to purchasing an antimalarial.

The second level of figure B1 shows the probability of testing positive conditional on testing using administrative data from pharmacies in our study. Here, we observe that among those that tested, only 24% are malaria positive. Therefore, the vast majority of patients who purchase an antimalarial without a test are doing so unnecessarily.

The third level presents summary statistics that describe whether the test results inform treatment choices. Almost all patients (97%) who test positive purchase an antimalarial. Among those who test negative, only 5% of patients ignore their test results and purchase an antimalarial. Therefore, patients follow the test results in making treatment decisions, implying that diagnostic information is useful in preventing unnecessary care.

2.1 Patient demand for RDTs

We allow for two sequential choices for patients presenting malaria symptoms. First, the patient decides whether or not to to test for malaria. Second, the patient must decide whether to purchase an antimalarial medication. Malaria suspect patients start with unknown malaria status $M \in \{m, m'\}$ with probability of malaria $\boldsymbol{P(m)}$, or some other health condition causing the symptoms with probability $\boldsymbol{P(m')}$, such that $\boldsymbol{P(m)} = 1 - \boldsymbol{P(m')}$.

The patient's realized utility depends on their true malaria status $M \in \{m, m'\}$ and whether they consume an antimalarial. If the patient is malaria positive and left untreated, they receive a disutility $-d_m$. Similarly, if the patient is malaria negative, the patient has disutility $d_{m'}$ from the non-malaria condition. Receiving unnecessary malaria treatment leads to a disutility represented by $-d_w$, in addition to the unnecessary expenditure on treatment. For instance, side effects from antimalarials, the true underlying condition staying untreated for longer, and malaria resistance concerns can affect the value of d_w. A healthy patient's utility is normalized to zero.

Patients choose whether to buy an ACT and pay price p_a to receive malaria treatment.

We write the true expected utility with and without ACT purchase as:

$$\mathbb{E}[U] = \begin{cases} \text{ACT} & -p_a - \boldsymbol{P(m')}(d_{m'} + d_w) \\ \text{No ACT} & -\boldsymbol{P(m')}d_{m'} - \boldsymbol{P(m)}d_m \end{cases} \tag{1}$$

Without the purchase of an RDT, true malaria status M is uncertain. If the patient buys an ACT, they will avoid the dis-utility of having untreated malaria, but they will have a probability $\boldsymbol{P(m')}$ of incurring a dis-utility for inappropriate treatment. If the patient does not buy the ACT, the patient has probability $\boldsymbol{P(m)}$ of incurring disutility d_m from having untreated malaria.

However, patients might not know the true value of $\{\boldsymbol{P(m)}, d_w\}$ and instead hold belief $\{P(m|i), d_{w,i}\}$ about their status and their dis-utility from unnecessary treatment. Hence, patients make decisions based on their choice utility which is a function of belief $\{P(m|i), d_{w,i}\}$. For instance, let $\mathbb{E}[U(\text{ACT})|i]$ and $\mathbb{E}[U(\text{No ACT})|i]$ be the patient's choice utility from receiving and skipping treatment respectively.

We first assume that $\mathbb{E}[U(\text{ACT})|i] > \mathbb{E}[U(\text{No ACT})|i]$ and $\mathbb{E}[U(\text{ACT})] > \mathbb{E}[U(\text{No ACT})]$. That is, observed patients will always choose to consume ACT and this will be optimal under uncertainty. This assumption implies that the disutility d_m is sufficiently large that a patient with malaria-presenting symptoms will always chose to receive an anti-malarial treatment when their status M is uncertain.[11] Therefore, the patient's perceived value of not testing is given by:

$$\mathbb{E}[U(\text{No RDT})|i] = -p_a - P(m'|i)(d_{m'} + d_{w,i}) \tag{2}$$

Patients, on the other hand, have the option to purchase an RDT test for price p_r which will inform them on their status M, that is, $P(m|\text{RDT positive}) = 1$ and $P(m|\text{RDT negative}) = 0$. We assume that buying an ACT and an RDT is feasible for the patients (i.e. $p_a + p_r <$ B for the patient's healthcare budget B).

We further assume that if a patient knows they are malaria positive after incurring cost p_r, then the patient will purchase an ACT. This follows from the assumptions that patient will purchase an ACT under uncertainty and that they can afford both a RDT and ACT. Both assumptions are consistent with what we observed in the data[12]. Moreover, since RDTs are assumed to remove the patient's uncertainty on their malaria status, if an RDT is negative, the patient should not purchase an antimalarial[13].

We can now write the patient's value of purchasing an RDT as:

$$\mathbb{E}[U(\text{RDT})] = -p_r - \boldsymbol{P(m)}p_a - \boldsymbol{P(m')}d_{m'} \tag{3}$$

11. Note that this assumption is consistent with our setting where patients have self-selected into seeking malaria care in a pharmacy. We validate this assumption empirically in figure B1.

12. The probability of buying an ACT if the patient tested positive is $> 95\%$ as seen in figure B1

13. Again, this is consistent with our data since the probability of buying an ACT after testing negative is 5%.

Because patients purchase an ACT after buying an RDT if they test positive, the patient's value of RDTs depends on the price of both products. RDTs do, however, guarantee that the patient will not incur disutility d_w, and allows for the possibility of avoiding unnecessary expenditures on antimalarial treatments. As before, patients decide whether to test based on their perceived value of testing $\mathbb{E}[U(\text{RDT})|i]$ which is a function of their belief about $\boldsymbol{P(m)}$.

Combining 2 and 3 together, patient i buys an RDT if and only if:

$$\begin{aligned} &\mathbb{E}[U(\text{RDT})|i] > \mathbb{E}[U(\text{No RDT})|i] \\ \iff &-p_r - P(m|i)p_a - P(m'|i)d_{m'} > -p_a - P(m'|i)(d_{m'} + d_{w,i}) \\ \iff &\boldsymbol{P(m'|i)(p_a + d_{w,i}) > p_r} \end{aligned} \tag{4}$$

Equation 4 describes the patient's optimality condition for the purchase of RDTs based on their choice utility. The decision to purchase an RDT depends on 4 key values: the patient's beliefs about their malaria status, the patient's perceived disutility of incorrectly receiving malaria treatment, and the prices of both RDTs and ACTs.

For a given patient distribution, we express demand for RDTs as:

$$D(p_a, p_r) = P(P(m'|i)(p_a + d_{w,i}) - p_r > 0) \tag{5}$$

where $P(m'|i)(p_a + d_{w,i})$ is a random variable reflecting individual $i's$ beliefs about malaria risk and disutility from receiving unnecessary ACT treatment when they are malaria negative.

Moreover, since the demand is a cumulative density function, this expression yields simple comparative statistics. Demand is increasing on the price of the ACT and the patient's beliefs about $\boldsymbol{P(m)}$ and d_w, and decreasing on the price of the RDT:

$$\frac{\partial D(p_a, p_r)}{\partial p_a} > 0 \quad \frac{\partial D(p_a, p_r)}{\partial p_r} < 0 \quad \frac{\partial D(p_a, p_r)}{\partial P(m'|i)} > 0 \quad \frac{\partial D(p_a, p_r)}{\partial d_{w,i}} > 0$$

2.2 Patient diagnosis contingent contracts

Our model of demand for diagnosis in hand we turn to contract design. A diagnosis contingent contract simply reduces the price of the ACT that the patient pays conditional on the patient testing and the outcome of that test. We express an ACT with discounted price as $p^*_{a|r} = (1 - \delta_a)p_a < p_a$ available only if the patient purchases an RDT through the contract's program. Parameter δ_a describes the relative size of the discount. The patient can continue to purchase the ACT without an RDT at market price p_a. We extend this by reducing the cost of the RDT test by δ percent, representing a direct subsidy for testing. The discounted price for RDTs is given by $p^*_r = (1 - \delta_r)p_r < p_r$ for discount rate δ_r and market price p_r. Under a diagnosis contingent contract, the patient's optimality condition becomes:

$$\begin{aligned} &p_a - P(m|i)p^*_{a|r} + P(m'|i)d_{w,i} > p^*_r \\ \iff &p_a - P(m|i)(1 - \delta_a)p_a + P(m'|i)d_{w,i} > p^*_r \\ \iff &(1 - P(m|i)(1 - \delta_a))p_a + P(m'|i)d_{w,i} > (1 - \delta_r)p_r \end{aligned} \tag{6}$$

When $\delta_r = \delta_a = 0$, this condition is identical to equation 4. When $\delta_r \in [0, 1]$ and $\delta_a \in [0, 1]$, the comparative statistics implied by equation 4 continue to hold. However, patient demand

for RDTs is now increasing on the discounts δ_r and δ_a:

$$\frac{\partial D(p_a, p_r, \delta_a, \delta_r)}{\partial \delta_r} > 0 \quad \frac{\partial D(p_a, p_r, \delta_a, \delta_r)}{\partial \delta_a} > 0$$

Diagnostic contingent contract structures that target prices paid by patients for both testing and treatment based on the outcome of the test, increase demand for malaria testing and prevent unnecessary treatments. When deciding whether to test, patients not only care about the price of the RDT, but also about the cost of treatment. Our proposed contract increases demand for testing through changes in the price of the both products. A lower price for RDTs increases demand for testing. Second, a conditional discount on ACTs indirectly increasing the value that testing by making the expected cost of testing positive lower.

Note that demand for ACTs is fully determined by the demand for RDTs in this model. Diagnosis contingent contracts reduce demand for malaria treatment on expectation since patients who purchase tests as a result of these interventions will only purchase treatment if they are actually positive. In particular, demand for ACTs is given by:

$$D_{ACT}(p_a, p_r, \delta_a, \delta_r) = 1 - \boldsymbol{P(m')} D(p_a, p_r, \delta_a, \delta_r) \tag{7}$$

Whether this reduction in treatment at the expense of more testing is welfare enhancing from the social planner's perspective is not immediately clear. On one hand, patients might fail to internalize externalities and internalities from unnecessary treatment (akin to the "internalities" discussed in (Baicker, Mullainathan, and Schwartzstein 2015)), or overestimate their probability of being malaria positive, in which case interventions that increase demand for testing could lead to increases in patient and social welfare. On the other hand, increasing demand for testing could prove distortionary if it leads to an increase in medical expenditures that is beyond the value from avoiding unnecessary treatments. Section 2.4 discusses the welfare implications of diagnosis contingent contracts in more detail.

2.3 Provider diagnosis contingent contracts

Providers counsel malaria suspect patients on the value of testing. For simplicity, assume that providers can signal the value of testing to the patient through $\theta \in \{0, 1\}$, whether the provider recommends to be tested or not. In particular, let provider advice update the patient's belief about d_w through function $d_{w,i}(\theta)$[14].

If a provider recommends a test $\theta = 1$, it is likely that the patient will interpret this as a signal that d_w is high, or equivalently, that the value of testing is high. If this is the case, then the provider's recommendation to test increases the demand for RDTs:

$$D(p_a, p_r|\theta = 1) - D(p_a, p_r|\theta = 0) > 0 \tag{8}$$

14. For tractability we load the effect of the provider's advice on patient belief about d_w since this can be interpreted as a constant shift in the patient's value for testing. Similar results follow if we allow provider advice to also impact patient's beliefs about their malaria status.

Provider's motivations to recommend testing are potentially twofold. On one hand, providers care about the patient's welfare and the potential for increasing malaria resistance with unnecessary treatments. On the other hand, providers might care about their financial incentives. For concreteness, let the provider's decision to recommend testing be given by:

$$\theta = 1\{W(d_w, d_{sw}, \boldsymbol{P(m)}, p_r, p_a, \delta_r, \delta_a) + \lambda f(\boldsymbol{\pi}, t) > e_d\} \tag{9}$$

Such that $W(d_w, d_{sw}, \boldsymbol{P(m)}, p_r, p_a, \delta_r, \delta_a)$ is a function that represents the provider's internalized patient's welfare and concerns about malaria resistance in their community, and $f(\boldsymbol{\pi}, t)$ represents the provider's financial incentives to recommend testing. Parameter d_{sw} denotes the marginal externality of a patient receiving unnecessary treatment while d_w denotes internalities to the patient from unnecessary treatment as discussed above. The provider's financial incentives are a function of the vector of markups $\boldsymbol{\pi}$ for all the malaria products sold in the pharmacy and a vector of any incentives included in the diagnosis contingent contracts (t). In particular:

$$\begin{aligned} f(\boldsymbol{\pi}, t) &= \mathbb{E}[\Pi|\theta = 1] - \mathbb{E}[\Pi|\theta = 0] \\ &= \Sigma_k(\pi_k + t_k)(P[k|\theta = 1] - P[k|\theta = 0]) \end{aligned} \tag{10}$$

Such that Π are the provider's expected profits from a malaria counseling interaction with the patient. These profits depend on the probability of patients deciding to buy product $k \in \{r, a, a|r\}$ conditional on their advice to the patient. The provider's incentives include direct transfers to the provider from the diagnosis contingent contract, and the markups of the pharmacy for the sale of the distinct available products.

Provider diagnosis contingent contracts change the financial incentive structure from the sale of the malaria products, encouraging providers to recommend testing. In particular, these contracts increase $\pi_k + t_k$ for $k \in \{r, a|r\}$. Since θ is likely to be positively correlated with the patient's beliefs about the value of testing, the probabilities that patients buy an RDT (r) or an ACT conditional on an RDT sale ($a|r$) should both be increasing on θ. In other words, $P[r|\theta = 1] - P[r|\theta = 0] > 0$ and $P[\{a|r\}|\theta = 1] - P[\{a|r\}|\theta = 0] > 0$. Therefore, provider diagnostic contingent contracts increase the provider's incentives to recommend testing $f(\boldsymbol{\pi}, t)$, and thus, potentially increase demand for testing by the patient.

2.4 Patient Welfare

We now turn to the analysis of the expected welfare effects of diagnosis contingent contracts, beginning with the effects of the contracts on patient welfare. Absent diagnosis-contingent contracts, expected patient welfare is given by:

$$\begin{aligned} \mathbb{E}(Welfare) &= P(RDT) * \mathbb{E}[U(RDT)] + P(RDT') * \mathbb{E}[U(NoRDT)] \\ &= -p_a[(1 - D(0)) + D(0)\boldsymbol{P(m)}] \\ &-p_r D(0) - d_{m'}\boldsymbol{P(m')} - d_w\boldsymbol{P(m')}(1 - D(0)) \end{aligned} \tag{11}$$

Equation 11 describes the welfare from a patient seeking malaria care implied by the model. First, define $\delta = \{\delta_a, \delta_r, t_r, t_{a|r}\}$ to be the vector of incentives from the contract including

patient discounts (δ_a, δ_r) and provider transfers $(t_{a|r}, t_r)$. Then the demand for an RDT in the absence of a contract is given by $P(RDT) = D(\delta = 0) = D(0)$.

On the other hand, when diagnosis contingent contracts are available with incentives $\delta = \hat{\delta}$, patient welfare becomes:

$$\begin{aligned}\mathbb{E}(Welfare) &= P(RDT) * \mathbb{E}[U(RDT)] + P(RDT') * \mathbb{E}[U(NoRDT)] \\ &= -p_a[(1 - D(\hat{\delta})) + D(\hat{\delta})\boldsymbol{P(m)}(1 - \delta_a))] \\ &-p_r D(\hat{\delta})(1 - \delta_r) - d_{m'}\boldsymbol{P(m')} - d_w\boldsymbol{P(m')}(1 - D(\hat{\delta}))\end{aligned} \tag{12}$$

In order to get the welfare effect from the introduction of a contract with incentives $\hat{\delta}$, we take the difference between equations 12 and 11. For simplicity, define $\Delta_D := D(\hat{\delta}) - D(0)$ as the contract's effect on the demand for RDTs. Welfare gains are thus:

$$\underbrace{p_a[\Delta_D\boldsymbol{P(m')} + \boldsymbol{P(m)}D(\hat{\delta})\delta_a]}_{\text{Unnecessary ACT expenditure}} - \underbrace{p_r[\Delta_D - D(\delta)\delta_r]}_{\text{RDT expenditure}} + \underbrace{d_w\boldsymbol{P(m')}\Delta_D}_{\substack{\text{Internalities} \\ \text{from unnecessary ACT}}} \tag{13}$$

Equation 13 is composed of three terms. The first term shows the effect on patient welfare due to changes in ACT expenditures. Note that since contracts have a positive effect on demand for RDTs ($\Delta_D > 0$), the first term is positive. In other words, patient welfare increases due to a reduction on ACT expenditures. The second term corresponds to the effect of contracts on patient welfare due to changes in RDT expenditures. The sign of this term is unclear, and will depend on whether the decrease in testing costs due to the contract outweigh increased demand for testing. Finally, the third term relates to an increase in patient welfare from avoiding unnecessary treatment. Contracts decrease unnecessary ACT uptake due to increased testing uptake. Note that the overall welfare effect to the patient is ambiguous, given that patient expenditures for testing could increase above their optimal level since patients can have incorrect beliefs about $\boldsymbol{P(m)}$ and their private values for d_w.

2.5 Social welfare

To study the social welfare effect of this contract, one needs to consider the costs of implementing the contract and the pharmacy welfare in addition to the patient welfare effects discussed above. First, the policy maker introduces incentives δ which are paid whenever the patient tests or purchases an ACT conditional on a positive test result. The expected costs to the social planner from implementing the contract net any externalities generated by the contract are given by:

$$D(\hat{\delta})[\delta_r p_r + t_r] + D(\hat{\delta})\boldsymbol{P(m)}[\delta_a p_a + t_{a|r}] - d_{sw}\boldsymbol{P(m')}\Delta_D \tag{14}$$

Note that while the social planner pays incentives δ_r and t_r for every patient who buys an RDT, the incentives associated with treatment (δ_a and $t_{a|r}$) are only paid for malaria positive patients. Moreover, the social planner internalizes any externalities generated by the contract.

Second, the profits to the pharmacy or clinic offering malaria care to the patient can change as a result of the introduction of these contracts. Taking into account the pharmacy

welfare is both important for studying social welfare effects and the sustainability of the policy. The effect of the contracts on pharmacy profits are ambiguous. This is because pharmacy welfare will depend on the profit margins for treatment and testing, in addition to the effects on patient choice probabilities. The effect on the pharmacy profits from these contracts is given by:

$$E[\Pi|\delta = \hat{\delta}] - E[\Pi|\delta = 0] = \Delta_D[\pi_r - \pi_a \boldsymbol{P(m')}] + D(\hat{\delta})[t_r + \boldsymbol{P(m)} t_{a|r}] \tag{15}$$

Where Π is the expected profits from a malaria-suspect patient encounter, $\pi_a = p_a - c_a$ is the profit margin for treatment, and $\pi_r = p_r - c_r$ is the profit margin for testing[15]. The marginal costs of testing and treatment are given by c_a and c_r. The first term in equation 15 relates to the change in profits from the introduction of the contract, whereas the second term relates to transfers from the policy maker to the pharmacy as part of the contract. In the results section, we find evidence that profits to pharmacies either stay unchanged or increase in our setting.

With the effects of these policies on patient welfare, pharmacy profits, and program costs in hand, we can now derive an equation for total social welfare effects from diagnosis-contingent contracts by adding up the distinct components. Therefore, the effect on social welfare is given by:

$$\begin{aligned} &\Delta_D[\boldsymbol{P(m')} p_a - p_r + d_w \boldsymbol{P(m')} + \pi_r - \pi_a \boldsymbol{P(m')} + d_{sw} \boldsymbol{P(m')}] \\ &= \Delta_D[\boldsymbol{P(m')} c_a - c_r + (d_w + d_{sw}) \boldsymbol{P(m')}] \end{aligned} \tag{16}$$

Note that the sign of the effect on social welfare of diagnosis-contingent contracts is ambiguous. Since the effect on demand for testing is positive, social welfare effects will depend on the relative magnitudes of marginal costs of testing and treatment — c_a and c_r — as well as the positivity rate of malaria-suspect patients — $\boldsymbol{P(m)}$. In fact, equation 16 implies the following result:

Theorem 1 *Let* $\boldsymbol{P(m')} > \frac{c_r}{d_w + d_{sw} + c_a}$*, then a diagnosis-contingent contract is welfare-improving.*

Condition in proposition 1 has an intuitive interpretation. First, if the probability of patients having malaria is very high ($\boldsymbol{P(m')} \simeq 0$), these contracts are unlikely to increase welfare since most patients will end up consuming antimalarial regardless of whether they purchase a test or not. Second, the higher the cost of testing relative to treatment ($c_r >> c_a$), the less likely that a contract that steers demand towards testing will increase social welfare. Third, if unnecessary treatments lead to costly externalities and internalities (d_{sw}, d_w), these contracts are likely to increase social welfare. Fourth, the higher the cost of treatment (c_a), the more likely that contracts that avoid unnecessary health expenditures will be welfare-improving[16].

15. Without loss of generality, assume that $\pi_a = \pi_{a|r}$; where differences in profitability across the a and $a|r$ pseudo-products are captured by transfer $t_{a|r}$.

16. While this condition relates to whether these contracts are welfare-improving if enforced, it does not

2.6 Observed lower bound on social welfare effects

Note that the value of d_w, which relates to internalities and externalities from over-treatment, is difficult to measure and unlikely to be directly observed. However, one can bound the social welfare effects of these contracts presented in equation 16 by focusing on market characteristics that are more commonly observable such as marginal costs and positivity rates. This is because unnecessary treatments do not carry any benefits ex-post so the term that includes $d_w + d_{sw}$ is either null or positive. In particular, a bound on social welfare effects from these contracts is given by:

$$\Delta_D[\boldsymbol{P}(\boldsymbol{m'})c_a - c_r] \tag{17}$$

This lower bound on social welfare includes the change in patient expenditures, profits for the pharmacy, and costs of introducing such contracts for the policy maker.

If marginal costs are unobserved and only data on product prices is available, an alternative lower bound can be constructed. This is because by revealed choice, the change in profits for pharmacies that accept these contracts is either positive or null, so one can focus solely on the components that relate to changes in patient expenditures, pharmacy transfers from the contract, and contract costs, leading to the following lower bound:

$$\Delta_D[\boldsymbol{P}(\boldsymbol{m'})p_a - p_r] \tag{18}$$

2.7 Effect of a diagnosis contingent contract on Social Welare

We now discuss which contract characteristics are more desirable for a policy maker interested in increasing social welfare. In order to evaluate different policy designs from the policy maker's perspective we need to introduce an additional metric. Given a fixed budget, a policy maker is not just interested in increasing welfare, but rather doing so in a way that maximizes the impact of each dollar spent on implementing the policy. In other words, a policy maker is concerned about the cost-effectiveness of these contracts. Following our conceptual framework, the cost effectiveness of contract $\hat{\delta}$ that aims to increase social welfare is given by the ratio of equations 16 and the fiscal cost component of equation 14:

$$\underbrace{\frac{\Delta_D}{D(\hat{\delta})[\delta_r p_r + t_r] + D(\hat{\delta})\boldsymbol{P}(\boldsymbol{m})[\delta_a p_a + t_{a|r}]}}_{\text{RDT uptake cost-effectiveness}} \underbrace{[\boldsymbol{P}(\boldsymbol{m'})c_a - c_r + (d_w + d_{sw})\boldsymbol{P}(\boldsymbol{m'})]}_{\text{Social welfare weight}} \tag{19}$$

Note that the cost-effectiveness of these contracts given by equation 19 can be decomposed into two parts. The first component is a standard cost-effectiveness measure for a policy that aims to influence the uptake of a product: the per-dollar impact of the policy on the demand for testing. The second component is a social welfare weight, which relates to whether increasing demand for testing is socially desirable.

guarantee that pharmacies will be willing to participate in the program. If these contracts lead to social gains, incentives to the pharmacy ($t_{a|r}$ and t_r) rather than the patient (δ_a and δ_r) can be a useful tool to align incentives and increase participation.

Our modified cost-effectiveness measure (relative to the standard measure in the health policy literature) highlights the fact that a policy maker aiming to increase the uptake of testing could be overestimating or underestimating the cost-effectiveness of a program from a social welfare perspective. For example, the policy could inadvertently reduce social welfare if testing is too costly relative to its social value. Alternatively, the policy could be generating social gains far and beyond the standard measure due to the prevention of unnecessary expenditures and the reduction of externalities from unnecessary treatments. In what follows we assume that the condition in proposition 1 holds, implying that increasing the demand for testing is socially desirable.

With our cost-effectiveness measure in hand, we now turn to the question of which contract characteristics increase cost-effectiveness. In particular, should the policy maker target incentives to test directly or should they target incentives to treat for patients with a positive test result? Note that since contract characteristics δ only appear in the first component of our cost-effectiveness measure, we can focus our attention on which contract characteristics increase demand for testing the most per dollar amount.

We find that diagnosis-contingent ACT price-reductions can be more cost-effective than RDT discounts under some circumstances. To see why this is the case, consider the marginal effect on patient demand from changes to discounted prices for ACTs and RDTs, holding the market price constant. For exposition purposes, assume that all patients have the same perceived probability of being malaria positive $P(m|i) = A$. This leads to:

$$\frac{\partial D(\delta)}{\partial p_r^*} = -D'(\delta) \tag{20}$$

$$\frac{\partial D(\delta)}{\partial p_{a|r}^*} = -P(m|i)D'(\delta) = -A * D'(\delta) \tag{21}$$

As observed in equations 20 and 21, patients are more responsive to changes in the price of the RDT than changes in the price of the ACT by a factor of $1/A \in [1, \text{inf})$. However, the real cost of the program is influenced by the true probability that a given patient is malaria positive — $\boldsymbol{P(m)}$. For an RDT discount, the per-patient program cost will be on expectation $(p_r - p_r^*)D(p_a, p_r, p_r^*, p_{a|r}^*)$, while for a conditional ACT discount the expected cost per patient will be $\boldsymbol{P(m)}(p_a - p_{a|r}^*)D(p_a, p_r, p_r^*, p_{a|r}^*)$. This is because the program only pays for the ACT subsidy when a patient is malaria positive.

Hence, if the true probability of being malaria positive is sufficiently low relative to the patients' beliefs, conditional ACT subsidies can be more cost-effective. To see why, note that as the belief $A \to 1$, the marginal effect of a conditional ACT discount approaches the marginal effect of RDT discounts. However, as $\boldsymbol{P(m)} \to 0$, the expected per patient cost of the conditional ACT discount will approach zero.

3 Experimental Design

The study randomized 140 pharmacies into 3 intervention groups and a control group. The three treatment arms are (Table 1):

1. Patient subsidy group (T1): Clients who sought care for suspected malaria cases at these pharmacies paid a subsidized price for RDTs (90% subsidy, a $0.10 USD copay) and a subsidized price for ACTs (80% subsidy, a $0.30 USD copay) conditional on a confirmed positive malaria diagnosis. The prices were advertised in large posters in prominent spots in the pharmacy.

2. Pharmacy incentive group (T2): Pharmacy owners received an incentive to sell RDTs ($0.90 USD), and an additional incentive to prescribe ACTs to malaria-positive patients ($0.80 USD). Pharmacy attendants received a $0.30 USD incentive for recording transaction information in the malaria case management platform and completing the sale of incentivized products. Pharmacies were free to set prices charged to patients.

3. Combined group (T3): Clients were eligible for discounted rapid tests (60% subsidy, a $0.40 USD copay) and discounted ACTs conditional on a positive test result (60% subsidy, a $0.60 USD copay). Pharmacy owners received an incentive to sell rapid tests ($0.20 USD), and an additional incentive to prescribe ACTs to malaria-positive patients ($0.10 USD). Pharmacy attendants received a $0.30 USD incentive for recording transaction information in the malaria case management platform and completing the sale of incentivized products. Pharmacies were free to set prices charged to patients.

The total value of the incentive was held fixed at 200 Kes (~$2 USD in 2021 exchange rates) across all treatment arms.[17]

The pharmacies that participated in the study were existing users of Maisha Meds's digital sales management platform. Maisha Meds is a Kisumu-based healthcare social enterprise that provides sales and inventory management support to small pharmacies and clinics throughout Kenya. The platform records all pharmacy transactions and product stock. The incentive interventions were integrated into Maisha Meds's digital platform and managed centrally by the Maisha team. Subsidy and incentive amounts were automatically calculated based on the products that are being bought/sold and verified by implementation staff independent of the pharmacies prior to disbursement to ensure implementation fidelity.

Pharmacy staff received training on the importance of diagnostic testing (all arms), proper RDT administration, and use of the malaria case management tool. Stocks of RDTs

17. The incentive amount is consistent with prior literature, was determined after a pilot phase, and was calibrated to ensure pharmacy profitability would not be adversely affected, compared to the status quo.

and ACTs were provided on consignment through Maisha Meds in the intervention arms, while in the control group they managed their own stock.

3.1 Sample Enrollment

The sample consists of for-profit pharmacies and the clients that present with malaria symptoms located in the thirteen counties in the malaria endemic and epidemic areas of Kenya's western regions. These pharmacies manage their own stock and sales of diagnostic tests and medications, and set their own prices.

To be eligible to participate in the study, pharmacies needed to be part of the Maisha Meds network and active users of the Maisha Meds digital sales and inventory management platform. Additionally, they had to be licensed pharmacies that were registered with Kenya's Pharmacy and Poisons Board. They also had to be willing to be randomized to one of the study arms, manage their sales through the digital tool, and to offer incentives (either supply- or demand-side) for malaria testing and treatment if assigned to one of the intervention arms.

All eligible pharmacies were mapped. Those located at least 0.5 km from other potential study participants were invited to participate.[18] Using these criteria 175 pharmacies were identify as eligible and were invited to participate in the study, of which 140 accepted.[19]

These 140 pharmacies were randomly assigned to one of the four arms in waves, stratified on average monthly malaria product sales volumes (above/below median), urban/rural, and location of pharmacy in lake endemic county. Figure B3 shows the geographic span of the experiment across the target regions in Kenya and the final selection of pharmacies.

3.2 Data

See Appendix Table B6 for study timeline and a description of the primary sources of data. The study was initially planned to begin in June 2020, but was delayed due to COVID-19. The pharmacy onboarding, standardized patient visits, and control group testing activities were all done in person following appropriate COVID-19 precautions.[20]

We use the following data sources for analysis:

1. *Baseline data:*

18. The average distance between study sites is 6.24 km (range of 0.5 km to 46.2 km).

19. Appendix Table B8 reports balance on baseline variables between pharmacies accepted (in sample) and those that declined (refusals). Facilities that declined to participate had been using the digital sales platform for longer than facilities in the sample frame. No other meaningful imbalances were found.

20. The research and implementation teams followed Kenyan and UC Berkeley CPHS guidelines for conducting research while keeping study staff, implementation staff, and study subjects safe from COVID-19. All personnel and pharmacy staff were required to wear masks, maintain 1 meter distance from each other, and sanitize hands frequently. The research and implementation teams provided adequate PPE and hand sanitizer for all study and implementation personnel. Pharmacies were required by the Kenyan government to have all staff wearing masks, and have hand washing stations for staff and pharmacy clients, and pharmacies in our sample were compliant with these requirements during the study period.

(a) Pharmacy owner survey: survey about number of staff, pharmacy business operations, patient volumes, pharmacy characteristics, costs and revenues, and knowledge of malaria case management.

(b) Pharmacy staff survey: survey about malaria case management knowledge, worker motivation, and use of the digital platform used to manage sales and inventory.

(c) Sales data: continuously collected transaction data including prices and quantities of products purchased (RDTs, ACTs and malaria products) for sites that used the POS platform prior to study launch.

2. *Administrative data:*

(a) Sales data: continuously collected transaction data including prices and quantities of products purchased, location, date, and time of sale, and pharmacy staff who made the sale for over 50,000 malaria-related patient encounters between June 2021 - February 2022.[21]

(b) Malaria case management data: continuously collected transaction data on all rapid test and treatment purchases made through incentive program, including information on age/gender of patient, rapid test result, prices and quantities of medications purchased, location, date, and time of sale. Over 8,000 malaria transactions logged between June 2021 - February 2022.

3. *Standardized Patient Survey:* We use data from standardized mystery patient (SP) visits to measure the appropriateness of the care delivered using the same clinical case scenario. We trained individuals (SPs) to present an identical standardized illness case scenario as real walk-in clients to providers. During encounters with providers, SPs portrayed real patients presenting a standardized, pre-scripted acute adult malaria case. The SPs were confirmed to be malaria-negative based on malaria microscopy tests administered by a reliable, high-quality laboratory before and after the month of field work. SPs and field work supervisors also monitored any potential symptoms throughout field work; all were otherwise healthy. By using trained SPs portraying the same illness case to generate the care data, we avoid bias from selection on patient illness type and severity that is inherent in care data collected using other common methods such as patient exit interviews, direct clinical observation, or health records (Peabody et al. 2000).

21. Prices observed in the data are retail prices set by pharmacists in the digital tool.

4. *Testing subsample data:* data on test positivity from testing of random subset of 230 pharmacy clients at control group sites to obtain test positivity rate in a sample unaffected by the interventions (8.5 clients tested/site, 28 sites participated). Additional test positivity data from administrative records from 10 control group pharmacies that kept records of tests conducted (N=2547) on-site between January-February 2022.

3.3 Estimation Methods

All primary analyses are conducted at the patient level.[22] For all binary outcomes, we report marginal treatment effects from adjusted linear probability models using the following specification:

$$Pr(Y_{ip}) = \beta_0 + \beta_1 T_{1ip} + \beta_2 T_{2ip} + \beta_3 T_{3ip} + \lambda_s + \tau_t + \boldsymbol{X}_p + \epsilon_{ip} \tag{22}$$

where Y_{ip} is a malaria testing or treatment outcome, T_{jip} are treatment assignment indicators for each intervention j for individual i seeking care at pharmacy p, with the control group as the reference category, λ_s are strata fixed effects, τ_t are calendar month fixed effects, and ϵ_{ip} is the error term. We include variables that had significant imbalance with the control group at the $\leq 10\%$ level at baseline (Appendix Table B7) as covariates in this adjusted model ($\boldsymbol{X_p}$), as specified in the pre-analysis plan. The β terms the marginal treatment effect of each intervention relative to the control group, as percentage point changes. Results of unadjusted models (excluding $\boldsymbol{X_p}$), and using logistic regression, are consistent with findings from the adjusted models, and can be made available upon request.

4 Sample Balance Across Study Arms

Table B7 reports the experimental balance checks at baseline, and shows that randomization was balanced across a large set of pre-specified covariates. Out of 42 tests conducted, 5 are significant at the ≤ 10 percent level. When we conduct a joint test for orthogonality using a multinomial logit model with treatment assignment as the categorical outcome, we find that the χ^2-test produces a p-value of 0.46. This suggests that these covariates are not jointly predictive of group assignment. In the adjusted models, we control for covariates that were unbalanced at baseline from comparisons with the control group.

5 Results

5.1 RDT Use

Table 2 (Column 1) reports the estimated effects of the incentive interventions on RDT use. As a reference point, only 8% of patients who sought care for malaria-related symptoms in

22. The analyses specified in this section were pre-registered in a pre-analysis plan (AEARCTR-0004705). We discuss any deviations from the pre-analysis plan where relevant.

control group pharmacies purchased a rapid diagnostic test prior to obtaining treatment, which is consistent with trends found across the full pharmacy sample prior to the start of the experiment (Appendix Figure B4) as well as with other existing research on rapid diagnostic test use in pharmacy settings across East Africa.[23]

The contracts increased RDT use substantially, consistent with the comparative statics derived in sections 2.2 and 2.3. We find large and statistically significant effects in all three arms; patient discounts resulted in a 23 percentage point increase in RDT uptake, while the pharmacy incentives and the combination of patient discounts and pharmacy incentives resulted in 19 and 22 percentage point increases, respectively. However, the differences across arms were not statistically significant from each other. Pharmacy-administered incentives to either patients or providers lead to more people being tested for malaria prior to receiving treatment.

5.2 ACT Use

Table 2 column 2 presents results on the impact of incentives on overall ACT use. Despite the very limited diagnostic tests sold in the control group, the vast majority of control group patients who sought care for suspected malaria purchased ACTs (89%) or another antimalarial medication (4%).[24] Based on the malaria prevalence rate of 24% in our study sample across patients who tested, as much as 76% of malaria-suspect patients purchase ACTs unnecessarily[25]. Therefore, there are large levels of medication waste.

We find that the incentives caused a decrease of between 8-15 percentage point decline (column 2) on ACT uptake. Again, the three arms are not statistically distinguishable from each other. This result is consistent with the model implications. Since contracts are effective at increasing testing rates, more patients learn their true malaria status. Note that the model implies that $\Delta ACT = -\boldsymbol{P(m')}\Delta RDT$. With a value of $\boldsymbol{P(m')} = 0.76$, the range of decrease in ACT demand implied by the model (14-18 percentage points) roughly overlaps with the range estimated through linear regression (note that the small non-compliance to test results can attenuate this predicted reduction). Therefore, the results are consistent with the demand function for treatment presented in section 2.2,

We conclude that the reduction in ACT demand is the result of a reduction in unnecessary

23. For example, in Cohen et al. 2013; Cohen, Dupas, and Schaner 2015; O'Meara et al. 2016; Ansah et al. 2010.

24. Column 3 of Table 2 shows (zero) program impact on non-ACT antimalarial medication sales, so we do not find evidence of substitution to other medications.

25. We also conducted a random testing activity where we found a prevalence rate of 0.40 among 35 individuals that purchased antimalarials or malaria tests. This estimate's CI includes the prevalence rate conditional on testing. Our preferred measure of P(m) = 0.24 uses the administrative data on test results since this is more precisely estimated due to a larger sample of patients and study sites, however, robustness checks are included assuming $P(m) = 0.40$.

antimalarial consumption. Figure B1 shows that the propensity to purchase treatment drops from 100% to 5% after the patient receives a negative test result in our sample (while 97% of malaria positive patients continue to purchase antimalarials). Since we find no evidence that patient volumes changed due to the study interventions, we can conclude that the observed decline in ACT demand is the result of malaria negative patients reducing unnecessary antimalarial consumption thanks to the higher incentives to test and find their true status.

5.3 Mechanisms

As exemplified in the model, there are two potential mechanisms through which diagnosis contingent contracts could impact demand for testing: changes in prices to the patient, or changes in the counseling of the provider. We investigate mechanisms using administrative data on prices along with data from an audit study that uses standardized patients (SPs) to measure the content of the care visit using the same clinical case scenario, with results in Table 3.[26] SPs have an advantage over client exit surveys or administrative transaction data in our setting as they avoid bias from selection on patient characteristics and malaria status (real or perceived). The SP data, by capturing data on an unselected patient sample - where the only variation is by the experimental design - allows us for cleaner identification of mechanisms.

We trained individuals (SPs) to present an identical standardized illness case scenario as real walk-in clients to providers. SPs followed a uniform script for how to present a suspected malaria case in a pharmacy setting: SPs were instructed to complain of fever, headache and joint pains in their opening statement and then provided additional information about their illness episode and health history if the pharmacist followed up with additional questions. SPs conducted a total of 411 visits across 137 facilities in the study sample, with three different SPs visiting each facility. SP visits provided a unique opportunity to assess the implementation fidelity and quality of care of the patient-provider interaction at study pharmacies.

First, we report intervention effects on patient prices, as reported in the administrative data (Table 3, columns 1-3)[27]. We find significant evidence of price pass-through for testing and antimalarials in T1 and T3, but no evidence of price reductions in T2. In column 1, we see that there is price pass-through in the patient discount arm (T1) and mixed arm (T3), but no evidence of pass-through in the supply-side incentive arm (T2). Column 2 shows the price of treatment conditional on testing. Again, there is only evidence of price reductions

26. SPs have been used to measure quality of care extensively. For example see: Peabody et al. 2000; Das et al. 2012; Das et al. 2016; Mohanan et al. 2015; Kwan et al. 2018; Kwan et al. 2019; Kwan 2022; Das et al. 2022; Boone et al. 2023.

27. Price data from the point of sale system is normalized to match dose-level prices for testing and AL-combination treatments. For details, see appendix 8.1.

in T1 and T3. Column 4 shows that treatment prices outside the loyalty program (outside-option to the patient) are invariant to treatment assignment. Column 4 shows the price of treatment without a test across the three treatment arms. As expected, there is no effect from the interventions. Taken together, the results on prices suggests that in the patient discount arm (and to a lesser extent, the combined arm), the increase in testing uptake and improvements in treatment targeting can be explained by reduced patient prices on rapid diagnostic tests and conditional discounts on treatment.

But this price mechanism does not appear to explain why we find similar effects on testing and treatment decisions in the provider incentive arm (T2). We test an alternative mechanism: provider counseling. Columns 4 of Table 3 presents results on pharmacist advice and counseling behavior, using data from SP exit surveys. Specifically, we analyze SP reports of whether the pharmacist comprehensively explained their test result and treatment regimens, a measure of quality of counseling. We find that only 31% of the SPs at the control group sites report receiving comprehensive information on tests and / or treatment options, and the diagnosis-contingent incentive contracts significantly improve this. We see that the improvements in counseling are driven entirely by the provider-side incentives (column 4). When pharmacists are incentivized directly, they are 14-16 percentage points more likely than control group pharmacists to clearly explain treatment options to SPs. This suggests that when incentivized directly, pharmacists change their behavior and provide more comprehensive counseling on testing and treatment options to suspected malaria patients.

Taken together, these results suggest that price-pass through is likely to explain the demand-side treatment effects, and the information/counseling channel is likely to explain the supply-side treatment effects we find in Table 2.

6 Welfare

In this section, we estimate the welfare impact of introducing diagnosis contingent contracts using the framework presented at the beginning of the paper. The measure derived in equation 16 quantifies the impact on social welfare per unit of cost for the social planner. Before proceeding to the welfare analysis, we will present evidence that validates key assumptions made in our conceptual framework.

6.1 Are diagnosis contingent contracts welfare enhancing?

Within the model's framework, whether preventing unnecessary treatment by encouraging higher testing is socially optimal depends on the relative costs of testing and treatment, the probability of being malaria positive, and the internalities and externalities from unnecessary treatment (as given by the condition from theorem 1). Diagnosis contingent incentive contracts have the potential of increasing total social welfare because it is possible for pa-

tients to make suboptimal decisions (from either a social or individual perspective) if they fail to internalize the externalities and internalities of unnecessary treatment, or if they hold incorrect beliefs about their malaria status.

We find evidence that the condition in Theorem 1 holds in our setting, suggesting that diagnosis-contingent contracts are expected to increase social welfare. While we do not observe d_w and d_{sw}, we do observe proxies for costs of testing and treatment, as well as the positivity rate $\boldsymbol{P(m)}$. For a given patient in our data, the total cost of treatment (c_a) is on average \$1.60USD. On the other hand, the total cost of testing is \$0.60USD (accounting for multiple units being sold per-patient on average). The probability of a patient being malaria negative, conditional on testing, is $\boldsymbol{P(m')} = 0.76$. Therefore, regardless of the value of $d_w > 0$ and $d_{sw} > 0$, these contracts are expected to increase welfare in our setting ($\boldsymbol{P(m')} > 0.36$).

In the model section, we showed that one can use the expenditures of the patient of each product category instead of the costs to create a lower bound on welfare. This is because, by revealed choice, expected profits of participating in the contracts should be positive for pharmacies, allowing us to focus on patient welfare to bound total effects. Indeed, table B9 shows that the contracts either had null effects on profits or positive effects across the three treatment arms, while reducing total patient expenditures across the board. The average expenditures on treatment by patients is roughly 265KES, while the mean expenditure on tests is 120KES[28]. Therefore, even with this weaker test one would expect total welfare to increase if contracts that encourage testing are introduced ($\boldsymbol{P(m')} > 0.45$), consistent with the reduced form evidence on the effect on profits and patient expenditures in table B9.

6.2 Social welfare effects

In this section we estimate the per-unit-of-cost effect of these interventions on social welfare using the measure derived in equation 19. Table B10 presents the inputs used for estimation. To estimate average per patient costs of the program, we assume that marginal operational costs of these incentive contracts are negligible, allowing us to focus on the direct costs of implementing the price reductions and incentives. The contracts are implemented through a digital system, so the main operational cost are stock deliveries to these pharmacies. With economies of scale, the marginal cost of delivery at the patient level is likely negligible relative to the cost of the product itself when the contracts take advantage of existing supply-chain infrastructure, such as in our context. Moreover, we use administrative data on marginal costs of treatment and testing for pharmacies and the positivity rate from the random testing to estimate the welfare weight used to transform the cost-effectiveness measure into social

28. Note that in table 3 we instead present analysis on per-unit prices rather than expenditures; see appendix 8.1 for details on how these alternative units relate to each other

welfare effects.

For ease of interpretation, we report the increase in patient and provider welfare per unit of cost for the social planner, where a value above 1 implies a positive rate of return from the intervention. In other words, we report equation 16 plus 1 since patients and pharmacies benefit from transfers directly, which can be interpreted as a benefit-cost ratio. Table 5, panel A, displays the results. Column 3 shows the lower bound on the social welfare weight implied by the average cost (of holding stock) and the mean positivity rates of patients in the sample using administrative data. For each patient that shifts demand towards testing, we expect a social welfare increase of \$0.64 on expectation. Column 4 shows the cost-effectiveness of each treatment arm in terms of RDT demand increase. With patient subsidies, 0.65 additional patients purchase an RDT per dollar cost for the policy maker. Pharmacy incentives are slightly less cost-effective, with only a 0.51 patient increase per dollar. Finally, the combined treatment leads to a 0.56 increase in patients purchasing RDTs per dollar.

Column 5 shows the cost-effectiveness in terms of patient and provider welfare gains. The patient subsidy arm results in a \$1.41 increase on the welfare of patients and pharmacies per dollar spent by the policy maker, implying a rate of return of 41%. The other two arms are also cost-effective, with effects of \$1.32 and \$1.36 for the pharmacy incentive and the combined treatment arms respectively. We find that all of these welfare effects are significantly larger than 1, implying that all the policies have a positive rate of return. The increase in welfare is the result of the prevention of unnecessary treatment costs due to increased testing. Note that it could have been possible for these interventions to reduce social welfare if the increase in testing had raised overall malaria treatment expenditures. Therefore, it is reassuring that in this context the social welfare weight is positive and that shifting demand towards testing is socially desirable.

While the social welfare results presented here offer a lower bound on the effect of these interventions, it is likely that this lower bound is close to its true value. The marginal externality of a patient receiving unnecessary treatment is likely to be small in our setting given that the interventions were implemented in a small subset of the overall population. Similarly, internalities from unnecessary treatment such as patient's concerns about malaria resistance in their communities and side effects from unnecessary treatment are unlikely to be large relative to the changes in health expenditures.

7 Optimal contract design

Our framework, combined with the random variation from the experiment, allows us to study how contract design impacts its effectiveness in increasing social welfare. There are two margins that we consider: whether to target the patient or the provider's incentives,

and whether to give direct incentives to test or diagnosis-contingent discounts on treatment.

In order to evaluate the relative effectiveness of each contract component we need to estimate how the demand for testing responds to them. Note that equation 6 implies that demand for RDTs can be represented as follows:

$$RDT_i = 1\{P(m'|i)d_{w,i}(\theta) - (p_r - p_a) - P(m|i)p_{a|r} > 0\} \tag{23}$$

This representation highlights that patients care about the cost of testing relative to treatment $(p_r - p_a)$ and that provider advice shifts the demand for testing through $d_{w,i}(\theta)$. Moreover, the ratio of the coefficients of $(p_r - p_a)$ and $p_{a|r}$ in the demand curve allows you to recover $P(m|i)$.

To answer the question of whether to target the patient or the provider's incentives, we need to recover the marginal effect of the provider incentives on test demand. Note that in the model this effect operates through changes in advice and information θ. Therefore, to recover the marginal effect of provider incentives one needs to use a proxy for the provider's incentives to recommend. Let $\Delta P(k) = P(k|\theta = 1) - P(k|\theta = 0)$. Then, the provider's financial incentives can be rewritten as:

$$\begin{aligned} f(p,c,t) &= \Delta P(r)[(\pi_r + t_r) + \boldsymbol{P(m)}(\pi_{a|r} + t_{a|r})] + (\pi_a)\Delta P(a) \\ &= \Delta P(r) g(p,c,t) + (\pi_a)\Delta P(a) \end{aligned} \tag{24}$$

Note that the first component, $g(p,c,t) = (\pi_r + t_r) + \boldsymbol{P(m)}(\pi_{a|r} + t_{a|r})$, is only a function of the contract characteristics for which we have experimental variation. We argue that $g(p,c,t)$ is an appropriate proxy for the provider's financial incentives once one controls for the non-experimental markup π_a using our proposed instruments discussed below. Appendix 8.1 provides additional details for why this proxy is appropriate even when θ is unobserved.

For tractability, assume that the effect of advice on demand is linear on the provider's financial incentives:

$$P(m'|i)d_{w,i}(\theta) = \mathbb{E}[P(m'|i)d_{w,i}(\theta)|g(p,c,t)] + \epsilon_i = \beta_0 + \beta_1 g(p,c,t) + \epsilon_i \tag{25}$$

Therefore, demand is given by:

$$RDT_i = 1\{\beta_0 + \beta_1 g(p,c,t) - (p_r - p_a) - P(m|i)p_{a|r} > -\epsilon_i\} \tag{26}$$

Hence, for a patient-provider unit, one can recover estimates of $\frac{\partial D(.)}{\partial p_r}$, $\frac{\partial D(.)}{\partial p_{a|r}}$, and $\frac{\partial D(.)}{\partial g}$ through the following linear-probability equation:

$$P(RDT_i) = \hat{\beta}_0 + \hat{\beta}_1 g(\pi,t) + \alpha(p_r - p_a) + \alpha_{a|r} p_{a|r} + e_i \tag{27}$$

To identify causal estimates of the parameters of interest we need exogenous variation on $g(\pi,t)$, $p_r - p_a$ and $p_{a|r}$. Note that the experiment generates exogenous variation on provider's incentives for tests $g(\pi,t)$, testing price p_r, and diagnosis conditional ACT price $p_{a|r}$ by randomizing access to patient and provider incentives. However, the experiment does not generate independent variation between p_r and $p_{a|r}$ since T2 and T3 bundle RDT incentives and diagnosis contingent incentives for ACTs. Nevertheless, in the administrative data we observe the marginal costs to the pharmacy for treatment and testing outside of the

loyalty program, allowing us to instrument for p_a and p_r. Therefore, by instrumenting for $g(\pi, t)$, $(p_r - p_a)$, and $p_{a|r}$ through treatment assignment indicators, and marginal costs for malaria treatment and testing, one can recover causal estimates of the three marginal effects of interest.

We include two additional control variables in our estimation procedure. First, one needs to include π_a as a control variable since it is a component of $f(\pi, t)$ and it is mechanically correlated with c_a, which is one of our instruments. In addition, we include an indicator for whether prices were filled-in with arm averages whenever price data was unavailable from the POS system.

7.1 Empirical model results

Table 4 shows our preferred estimates from equation 27. Column 1 shows the main estimates, using treatment assignment indicators, and marginal costs of treatment and testing as instruments for $(p_r - p_a)$, $p_{a|r}$, and $g(\pi, t)$. Demand for tests responds to the three contract characteristics. First, strikingly, demand for tests has a very similar response to both direct RDT discounts and diagnosis contingent ACT discounts, implying that patients have very upwardly biased beliefs about their malaria status. We find that $\alpha = -0.23$ and $\alpha_{a|r} = -0.21$, both significant at the 99% and 95% levels respectively. This suggests that the average patient's belief that they are malaria positive is $\mathbb{E}[P(m|i)] = 0.22/0.23 \approx 95\%$. This is consistent with the patient overestimating their probability of having the disease relative to the true probability (24% using administrative data on RDT results). The p-value for the test of whether the implied patient $P(m|i)$ is equal or lower to the observed $P(m) = 0.24$ is 0.02, so we can reject that the implied patient beliefs are unbiased[29]. If patients had accurate beliefs about their malaria status, demand for testing should only respond to diagnosis contingent incentives by 24% of the price elasticity[30]. Appendix table C13 displays sensitivity checks to our model estimates where we assume a higher true positivity rate $P(m)$ of 0.40, 0.64 (welfare break-even point), and 0.90.

Demand for tests also responds to the financial incentives of the provider, conditional on price. The coefficient for $g(\pi, t)$ is equal to 0.40, suggesting that on the margin, demand responds strongly to the financial incentives of the provider. This result is consistent with the reduced form analysis in table 2, where demand for testing in treatment 2 increased absent any discounts to the patient. As observed in the mechanisms results in table 3, this seems to be driven by changes in the information shared by the provider to the patient.

29. We use Bias-Corrected and Accelerated (BCa) bootstrap to estimate this p-value.

30. Note that in our stylized model we do not include patient risk-aversion. Risk aversion could be an alternative channel contributing to the strong response to diagnosis-contingent contracts, although patients would need to be extremely risk-averse to explain this pattern in the absence of unbiased beliefs.

7.2 Counterfactual contract design

Now we turn to the question of which contract design is more effective at increasing social welfare. Table 5, Panels B and C, show the implied welfare effects from counterfactual contracts using the parameters estimated in Table 4. In our counterfactual contracts we vary whether the incentives are loaded on direct RDT discounts (incentives), diagnosis-contingent ACT discounts (incentives), patient discounts, or provider incentives. We use the fitted values from the constrained linear probability model estimated in Table 4 to obtain the predicted treatment effects on demand from these contracts.

First, we report the fit of the model's predictions in Table B11. To do so, we compare the welfare effects that were estimated directly in table 5, Panel A, against those implied by fitting the empirical model in Panel B of that same table at the mean values of each arm. In order to calculate the costs of each contract, we estimate $D(\delta)$ by adding together the control group's mean and the implied treatment effect from the fitted values from table 4. Overall, we find that the fitted values from the empirical model match very closely those estimated directly for both the patient and the provider's contracts. For the provider contract, the fitted models overestimates the treatment effect on RDT demand by a small margin, well within the confidence interval of the estimates in table 2. With both contracts, implied welfare effects very closely match across approaches with no significant differences.

Second, in Panel B of Table 5, we compare three patient contracts: a contract that loads the incentive on a RDT discount that brings the price down to 0USD, a contract with a diagnosis-contingent ACT discount that brings the diagnosis-contingent price to 0USD, and the experimental contract from T1, which is a combination of both types of discounts[31]. Column 1 shows that the three contracts have similar expected effects on testing demand in the range of 23-30 pp. This is because we estimated α and $\alpha_{a|r}$ to be approximately equal to each other. However, expected costs per patient vary widely between the three types of contracts, ranging from \$0.09 to \$0.44 per patient on expectation. This is because RDT discounts are paid every time a patient decides to purchase a test, leading to higher costs on expectation. On the other hand, contracts with only diagnosis-contingent discounts incur a cost only when the patient purchases a test and is malaria positive.

We find that diagnosis-contingent contracts are very cost effective relative to direct RDT discounts. Columns 3 and 5 show the cost-effectiveness and welfare estimates. In terms of the traditional cost-effectiveness measure, for each dollar of program cost the RDT discounts increase testing by 0.74 additional patients, while diagnosis-contingent ACT discounts increases demand by 2.79 patients per dollar. In terms of welfare gains, diagnosis contingent

31. Note that the size of the incentives varies by contract due to differences in prices across product categories

contracts increase patient and pharmacy welfare by \$2.79 for each dollar spent by the policy maker on the contracts. Direct RDT subsidies and the experimental contract are less effective, with a welfare effect of \$1.47 and \$1.44 respectively. The increase in welfare from an ACT discount only relative to the experimental design is almost significant at the 90% level (p=0.105).

Analogously, in Panel C we compare three contract designs: one that loads provider incentives on RDT sales, one that loads incentives on diagnosis -contingent ACT sales, and the experimental provider contract from T2. A key difference from the patient's contract counterfactuals is that we assume that providers have unbiased beliefs about the probability that the patient has malaria (conditional on testing). Nevertheless, it is still ambiguous whether diagnosis-contingent contracts are more cost-effective than direct RDT discounts since this will be determined by the patient's response to the provider's recommendation (β_1) relative to the contract costs.

Patient demand for testing varies significantly depending on the provider's incentive design. RDT incentives increase demand by 41 p.p., while diagnosis-contingent contracts only increase demand for testing by 11 p.p. – since providers understand the low probability that these incentives materialize. The program costs also vary meaningfully across contract designs. RDT incentives are more costly on expectation (\$0.49 per malaria-suspect patient case) since they are paid whenever a patient tests and the effect on demand from these incentives is large. Diagnosis-contingent contracts, on the other hand, are low-cost on expectation (\$0.05 per malaria-suspect patient case) since the demand response is small and the incentives are only paid when a malaria positive patient tests. We find that the diagnosis-contingent provider contract is more cost-effective than the RDT incentive contract (2.14 versus 0.83 additional tests sold per dollar spent by the policy-maker). Similarly, diagnosis contingent contracts have a larger effect on social welfare, increasing patient and pharmacy welfare by \$2.37 per unit of cost relative to \$1.53 for the RDT discount. The increase in welfare for ACT-only incentives to the provider is highly significant relative to either the experimental or RDT-only contracts.

Therefore, regardless of whether the policy targets the patient or the provider, diagnosis-contingent contracts are more cost-effective. Overall, the most cost-effective contract design are diagnosis-contingent discounts for the patient, since patients hold very biased beliefs about their malaria status.

Another key object of interest is the optimal level of the subsidy. Note that the welfare weight in equation 19 is constant across individuals[32] and accounts for the costs to the social planner. As long as this weight remains positive, the marginal value of shifting a

32. One could relax this assumption by allowing $P(m)$ and d_w to vary across individuals.

patient towards testing is positive. Consequently, under the assumption of linearity, total welfare is maximized when every patient shifts to testing (see Figure 1). Figure 1 illustrates that the simulated ACT-only incentive contract (Table 5) yields an expected social welfare gain of over $8,000 USD in our study sample. If we are willing to extrapolate beyond the support of the magnitude of the experimental subsidies, a higher subsidy level could achieve gains higher than $30,000 USD. In particular, the contract that maximizes total welfare only requires an expected cost per patient of just over $\$1USD$. While this contract would promise an incentive of over $\$4.2USD$ to patients who are positive, the low positivity rate in this sample of 0.24 would keep expected costs lower to the social planner. It is important to note that this exercise relies on heavy extrapolation; however it is strongly indicative that a larger subsidy size than those experimentally tested would increase total welfare. Moreover, while larger subsidy sizes would imply negative prices for ACTs, one can think of such contracts as insurance contracts for having a malaria infection rather than subsidies for the product itself. Note that the welfare gains from diagnosis-contingent ACT incentives always dominate RDT discounts and mixtures of the two contract types (e.g. T1), highlighting that conditional incentives are much more cost-effective than unconditional incentives to test[33].

Finally, in figure 2 we show that different positivity rate or patient belief environments significantly impact the cost-effectiveness of these policies. The top sub-figure highlights that the cost-effectiveness of these contracts is decreasing on the true positivity rate of the population being targeted. When $P(m) = 0.64$ the policy breaks even and leads to no welfare gains. On the other hand, when $P(m) = 90$, diagnosis-contingent contracts lead to a decrease in social welfare. Intuitively, this is because at such high positivity rates, encouraging testing leads to unnecessary medical expenditures. The bottom sub-figure shows the sensitivity of the relative cost-effectiveness of diagnosis-contingent contracts and RDT discounts to patient beliefs. When patients overestimate their probability of having malaria (as in our context), diagnosis contingent contracts are more cost-effective. When patients have unbiased beliefs, offering RDT discounts is as cost-effective as diagnosis contingent incentives. However, if patients underestimated their probability of having malaria instead, simple RDT discounts would be more effective at increasing social welfare.

33. It is important to note that we assume no costs associated with raising or delivering program funds; accounting for such costs would lower the optimal demand level. Moreover, monitoring costs to prevent fraud could potentially increase with contract incentives near the optimal level. Furthermore, because the three experimental interventions do not permit identification of curvature in the demand curve at different points, we cannot pinpoint the precise optimal level without extrapolating via linearity or another functional form. However, given the positive social welfare weight and strong response measured at observed levels, we can infer that the optimal subsidy level lies beyond the experimental levels tested. In figure C1, we show that our results are robust to a positivity rate of $P(m) = 0.4$, albeit cost-effectiveness will decrease with higher positivity rates.

8 Discussion

This paper examines the effects of a novel diagnosis-contingent contract structure to improve malaria case management in a cluster-randomized control trial in Kenya. The experimental treatments provided financial incentives to patients, pharmacists, or both for RDTs and ACTs conditional on testing positive for malaria and were implemented in private sector pharmacies in thirteen malaria-prone counties. By tying financial incentives for treatment to diagnostic outcome, we propose a flexible innovation in how payment contracts for health services could be structured to emphasize quality of care rather than service volume.

This paper contributes to the literature on performance-based financing mechanisms by examining the behavioral channels through which provider incentives impact healthcare quality. It also innovates in health financing contracts, proposing differential payment structures based on diagnostic information. Additionally, it explores how combination of personalized health information and financial incentives influence health behavior. And finally, it adds to our understanding of how incentives targeted at the demand-side or the supply-side can affect decision-making.

We find encouraging results of the demand- and supply-side incentives on both testing and treatment targeting. Overall, the incentives interventions increased RDT use substantially in a setting with very low baseline testing levels. On average, patients who sought care in treatment pharmacies were 23 percentage points more likely to receive a formal malaria diagnosis prior to purchasing treatment for suspected malaria. This represents a more than 300% increase over the control group. Incentive interventions were also effective encouraging appropriate use of antimalarials. We find an overall 12 percentage point decrease in the use of ACTs as a result of the treatment, and this is due to malaria negative patients opting out of purchasing unnecessary antimalarials. For patients who test positive, we find that they are appropriately nudged to take ACTs, consistent with their diagnostic test result. Interestingly, we find statistically indistinguishable effects of the demand-side and supply-side treatment arms, suggesting that incentives yield similar outcomes whether they are provided directly to patients or they are provided to pharmacists.

We explore mechanisms through which the incentive interventions worked in order to contextualize the main findings. We find that the patient subsidies resulted in significantly lower prices being paid by patients. However, we find no evidence of pass-through of the RDT incentive in either of the two supply-side arms, and no evidence of price pass-through on ACT prices in any of the three treatment arms. Instead, we find evidence that in the supply-side incentive arms, pharmacists explained diagnosis and treatment options more comprehensively to their patients. Improved, individualized health information appears to

be the channel through which the supply-side incentives resulted in the overall changes in RDT and ACT use seen in the main results. In sum, the demand subsidies induced more patients to purchase RDTs that provided accurate illness status information, which led to more appropriate use of ACTs. And, the supply-side incentives led pharmacists to provide more detailed diagnosis counseling and treatment recommendations, yielding similar overall effects on malaria case management.

Finally, we find that diagnosis contingent contracts are highly effective at increasing social welfare since they reduce unnecessary treatment by a larger amount than the cost of the contract to the social planner. The rate of return of the diagnosis contingent contracts we tested is of at least 32%. Counterfactual simulations suggest that there is room to increase the cost-effectiveness of these contracts significantly relative to the ones we tested empirically due to biased patient beliefs about their probability of being malaria positive. The highest gains can be achieved when the social planner focuses solely on treatment incentives conditional on testing, rather than direct incentives to test. In fact, the optimal contract loads all the incentives in diagnosis-contingent incentives for treatment at a higher level than what we experimentally tested. Taken together, our results imply that diagnosis-contingent contracts may have the potential to reduce medical waste and curb spending while better targeting health care resources to areas of proven need.

References

Afendulis, Christopher C, and Daniel P Kessler. 2007. "Tradeoffs from integrating diagnosis and treatment in markets for health care." *American Economic Review* 97 (3): 1013–1020.

Ahmed, Tashrik, Aneesa Arur, Damien De Walque, and Gil Shapira. 2023. "Incentivizing quantity and quality of care: evidence from an impact evaluation of performance-based financing in the health sector in Tajikistan." *Economic Development and Cultural Change* 71 (2): 000–000.

Ansah, Evelyn K, Solomon Narh-Bana, Michael Epokor, Samson Akanpigbiam, Alberta Amu Quartey, John Gyapong, and Christopher JM Whitty. 2010. "Rapid testing for malaria in settings where microscopy is available and peripheral clinics where only presumptive treatment is available: a randomised controlled trial in Ghana." *Bmj* 340.

Arrow, Kenneth J. 1963. "Uncertainty and the welfare economics of medical care." *The American Economic Review* 53 (5): 941–973.

Baicker, Katherine, Sendhil Mullainathan, and Joshua Schwartzstein. 2015. "Behavioral hazard in health insurance." *The Quarterly Journal of Economics* 130 (4): 1623–1667.

Baker, Laurence C. 2010. "Acquisition of MRI equipment by doctors drives up imaging use and spending." *Health Affairs* 29 (12): 2252–2259.

Basinga, Paulin, Paul J Gertler, Agnes Binagwaho, Agnes LB Soucat, Jennifer Sturdy, and Christel MJ Vermeersch. 2011. "Effect on maternal and child health services in Rwanda of payment to primary health-care providers for performance: an impact evaluation" [in en]. *The Lancet* 377, no. 9775 (April): 1421–1428. ISSN: 0140-6736, accessed January 26, 2020. https://doi.org/10.1016/S0140-6736(11)60177-3. http://www.sciencedirect.com/science/article/pii/S0140673611601773.

Boone, Claire E, Paul J Gertler, Grace Makana Barasa, Joshua Gruber, and Ada Kwan. 2023. "Can a private sector engagement intervention that prioritizes pro-poor strategies improve healthcare access and quality? A randomized field experiment in Kenya." *Health Policy and Planning* 38 (9): 1006–1016.

Burton, Deron C, Brendan Flannery, Bernard Onyango, Charles Larson, Jane Alaii, Xingyou Zhang, Mary J Hamel, Robert F Breiman, and Daniel R Feikin. 2011. "Healthcare-seeking behaviour for common infectious disease-related illnesses in rural Kenya: a community-based house-to-house survey." *Journal of health, population, and nutrition* 29 (1): 61.

Campbell, Stephen M, David Reeves, Evangelos Kontopantelis, Bonnie Sibbald, and Martin Roland. 2009. "Effects of pay for performance on the quality of primary care in England." *New England Journal of Medicine* 361 (4): 368–378.

Chen, Brian K, Paul J Gertler, and Chun-Yuh Yang. 2016. "Physician ownership of complementary medical services." *Journal of Public Economics* 144:27–39.

Cohen, Jessica, Pascaline Dupas, and Simone Schaner. 2015. "Price Subsidies, Diagnostic Tests, and Targeting of Malaria Treatment: Evidence from a Randomized Controlled Trial" [in en]. *American Economic Review* 105, no. 2 (February): 609–645. ISSN: 0002-8282, accessed November 21, 2019. https://doi.org/10.1257/aer.20130267. https://www.aeaweb.org/articles?id=10.1257/aer.20130267.

Cohen, Jessica L, Prashant Yadav, Corrina Moucheraud, Sarah Alphs, Peter S Larson, Jean Arkedis, Julius Massaga, and Oliver Sabot. 2013. "Do price subsidies on artemisinin combination therapy for malaria increase household use?: evidence from a repeated cross-sectional study in remote regions of Tanzania." *PloS one* 8 (7).

Currie, Janet, Wanchuan Lin, and Juanjuan Meng. 2014. "Addressing antibiotic abuse in China: an experimental audit study." *Journal of development economics* 110:39–51. ISSN: 0304-3878.

Currie, Janet, Wanchuan Lin, and Wei Zhang. 2011. "Patient knowledge and antibiotic abuse: Evidence from an audit study in China." *Journal of health economics* 30 (5): 933–949.

Cutler, David M, and Richard J Zeckhauser. 2000. "The anatomy of health insurance." In *Handbook of health economics,* 1:563–643. Elsevier.

Daniels, Benjamin, Ada Kwan, Madhukar Pai, and Jishnu Das. 2019. "Lessons on the quality of tuberculosis diagnosis from standardized patients in China, India, Kenya, and South Africa." *Journal of clinical tuberculosis and other mycobacterial diseases* 16:100109.

Das, Jishnu, and Jeffrey Hammer. 2014. "Quality of primary care in low-income countries: facts and economics." *Annu. Rev. Econ.* 6 (1): 525–553.

Das, Jishnu, Alaka Holla, Veena Das, Manoj Mohanan, Diana Tabak, and Brian Chan. 2012. "In urban and rural India, a standardized patient study showed low levels of provider training and huge quality gaps." *Health affairs* 31 (12): 2774–2784.

Das, Jishnu, Alaka Holla, Aakash Mohpal, and Karthik Muralidharan. 2016. "Quality and accountability in health care delivery: audit-study evidence from primary care in India." *American Economic Review* 106 (12): 3765–99.

Das, Veena, Benjamin Daniels, Ada Kwan, Vaibhav Saria, Ranendra Das, Madhukar Pai, and Jishnu Das. 2022. "Simulated patients and their reality: an inquiry into theory and method." *Social Science & Medicine* 300:114571.

Disease Control, Centers for, and Prevention. 2018. *CDC Malaria Support in Kenya.* https://www.cdc.gov/malaria/malaria_worldwide/cdc_activities/kenya.html.

Dow, William H, Justin S White, and Stefano Bertozzi. 2016. "Incentivizing health care utilization and health outcomes." In *World Scientific Handbook of Global Health Economics and Public Policy: Volume 1: The Economics of Health and Health Systems,* 421–451. World Scientific.

Dupas, P., and E. Miguel. 2017. "Chapter 1 - Impacts and Determinants of Health Levels in Low-Income Countries" [in en]. In *Handbook of Economic Field Experiments,* edited by Abhijit Vinayak Banerjee and Esther Duflo, 2:3–93. Handbook of Economic Field Experiments. North-Holland, January. Accessed January 26, 2020. https://doi.org/10.1016/bs.hefe.2016.09.003. http://www.sciencedirect.com/science/article/pii/S2214658X16300113.

Dupas, Pascaline. 2011. "Health behavior in developing countries." *Annu. Rev. Econ.* 3 (1): 425–449. ISSN: 1941-1383.

———. 2014. "Short-run subsidies and long-run adoption of new health products: Evidence from a field experiment." Publisher: Wiley Online Library, *Econometrica* 82 (1): 197–228. ISSN: 0012-9682.

Fleming-Dutra, Katherine E, Adam L Hersh, Daniel J Shapiro, Monina Bartoces, Eva A Enns, Thomas M File, Jonathan A Finkelstein, Jeffrey S Gerber, David Y Hyun, Jeffrey A Linder, et al. 2016. "Prevalence of inappropriate antibiotic prescriptions among US ambulatory care visits, 2010-2011." *Jama* 315 (17): 1864–1873.

Gertler, Paul, Paula Giovagnoli, and Sebastian Martinez. 2014. *Rewarding Provider Performance to Enable a Healthy Start to Life: Evidence from Argentina's Plan Nacer.* Accessed November 25, 2019. https://openknowledge.worldbank.org/bitstream/handle/10986/18801/WPS6884.txt.

Gong, Erick. 2015. "HIV testing and risky sexual behaviour." *The Economic Journal* 125 (582): 32–60.

Gruber, Jonathan, and Maria Owings. 1994. *Physician financial incentives and cesarean section delivery.*

Iizuka, Toshiaki. 2012. "Physician agency and adoption of generic pharmaceuticals." *American Economic Review* 102 (6): 2826–58.

Initiative, U.S. Presidents Malaria. 2021. *U.S. Presidents Malaria Initiative Kenya Malaria Operational Plan FY 2020.*

King, Jessica, Timothy Powell-Jackson, James Hargreaves, Christina Makungu, and Catherine Goodman. 2022. "Pushy Patients Or Pushy Providers? Effect Of Patient Knowledge On Antibiotic Prescribing In Tanzania: Study examines the effect of patient knowledge on antibiotic prescribing in Tanzania." *Health Affairs* 41 (6): 911–920.

Kwan, Ada. 2022. *Implementing quality of care measures: lessons from a standardized patient study in seven provinces of China,* 6.

Kwan, Ada, Benjamin Daniels, Sofi Bergkvist, Veena Das, Madhukar Pai, and Jishnu Das. 2019. "Use of standardised patients for healthcare quality research in low-and middle-income countries." *BMJ global health* 4 (5): e001669.

Kwan, Ada, Benjamin Daniels, Vaibhav Saria, Srinath Satyanarayana, Ramnath Subbaraman, Andrew McDowell, Sofi Bergkvist, Ranendra K Das, Veena Das, Jishnu Das, et al. 2018. "Variations in the quality of tuberculosis care in urban India: a cross-sectional, standardized patient study in two cities." *PLoS medicine* 15 (9): e1002653.

Lopez, Carolina, Anja Sautmann, and Simone Schaner. 2022. "Does patient demand contribute to the overuse of prescription drugs?" *American Economic Journal: Applied Economics* 14 (1): 225–260.

Ma, Xiaochen, Zhongqiang Zhou, Hongmei Yi, Xiaopeng Pang, Yaojiang Shi, Qianyun Chen, Mirjam E Meltzer, Saskia Le Cessie, Mingguang He, and Scott Rozelle. 2014. "Effect of providing free glasses on childrens educational outcomes in China: cluster randomized controlled trial." *Bmj* 349:g5740. ISSN: 1756-1833.

Maffioli, Elisa M, Wendy OMeara, Elizabeth Turner, and Manoj Mohanan. 2019. "Can Individuals Beliefs Help Us Understand Non-Adherence to Malaria Test Results - Evidence from Rural Kenya." *Economic Research Initiatives at Duke (ERID) Working Paper,* no. 243.

Manning, Willard G, Howard L. Bailit, Bernadette Benjamin, and Joseph P. Newhouse. 1986. "The demand for dental care." *RAND Health Insurance experiment series.*

Mbonye, Anthony K, Sham Lal, Bonnie Cundill, Kristian Schultz Hansen, Siân Clarke, and Pascal Magnussen. 2013. "Treatment of fevers prior to introducing rapid diagnostic tests for malaria in registered drug shops in Uganda." *Malaria journal* 12:1–10.

McGuire, Thomas G. 2000. "Physician agency." *Handbook of health economics* 1:461–536.

Meredith, Jennifer, Jonathan Robinson, Sarah Walker, and Bruce Wydick. 2013. "Keeping the doctor away: experimental evidence on investment in preventative health products." *Journal of Development Economics* 105:196–210. ISSN: 0304-3878.

Mfuh, Kenji O., Olivia A. Achonduh-Atijegbe, Obase N. Bekindaka, Livo F. Esemu, Calixt D. Mbakop, Krupa Gandhi, Rose G. F. Leke, Diane W. Taylor, and Vivek R. Nerurkar. 2019. "A comparison of thick-film microscopy, rapid diagnostic test, and polymerase chain reaction for accurate diagnosis of Plasmodium falciparum malaria." *Malaria Journal* 18, no. 1 (March): 73. ISSN: 1475-2875, accessed November 5, 2021. https://doi.org/10.1186/s12936-019-2711-4. https://doi.org/10.1186/s12936-019-2711-4.

Miguel, Edward, and Michael Kremer. 2004. "Worms: identifying impacts on education and health in the presence of treatment externalities." *Econometrica* 72 (1): 159–217.

Miller, Grant, and Kimberly Singer Babiarz. 2013. "Pay-for-performance incentives in low-and middle-income country health programs."

Mohanan, Manoj, Marcos Vera-Hernández, Veena Das, Soledad Giardili, Jeremy D Goldhaber-Fiebert, Tracy L Rabin, Sunil S Raj, Jeremy I Schwartz, and Aparna Seth. 2015. "The know-do gap in quality of health care for childhood diarrhea and pneumonia in rural India." *JAMA pediatrics* 169 (4): 349–357.

Musuva, Anne, Waqo Ejersa, Rebecca Kiptui, Dorothy Memusi, and Edward Abwao. 2017. "The malaria testing and treatment landscape in Kenya: results from a nationally representative survey among the public and private sector in 2016." *Malaria journal* 16 (1): 1–13.

O'Meara, Wendy Prudhomme, Diana Menya, Jeremiah Laktabai, Alyssa Platt, Indrani Saran, Elisa Maffioli, Joseph Kipkoech, Manoj Mohanan, and Elizabeth L. Turner. 2018. "Improving rational use of ACTs through diagnosis-dependent subsidies: Evidence from a cluster-randomized controlled trial in western Kenya" [in en]. *PLOS Medicine* 15, no. 7 (July): e1002607. ISSN: 1549-1676, accessed November 21, 2019. https://doi.org/10.1371/journal.pmed.1002607. https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1002607.

O'Meara, Wendy Prudhomme, Manoj Mohanan, Jeremiah Laktabai, Adriane Lesser, Alyssa Platt, Elisa Maffioli, Elizabeth L. Turner, and Diana Menya. 2016. "Assessing the independent and combined effects of subsidies for antimalarials and rapid diagnostic testing on fever management decisions in the retail sector: results from a factorial randomised trial in western Kenya" [in en]. *BMJ Global Health* 1, no. 2 (September). ISSN: 2059-7908, accessed November 21, 2019. https://doi.org/10.1136/bmjgh-2016-000101. https://gh.bmj.com/content/1/2/e000101.

Pauly, Mark. 1980. "Doctors and Their Workshops: Economic Models of Physician Behavior." In *Doctors and Their Workshops: Economic Models of Physician Behavior,* 119–122. University of Chicago Press.

Peabody, John W, Jeff Luck, Peter Glassman, Timothy R Dresselhaus, and Martin Lee. 2000. "Comparison of vignettes, standardized patients, and chart abstraction: a prospective validation study of 3 methods for measuring quality." *Jama* 283 (13): 1715–1722.

Peabody, John W, Riti Shimkhada, Stella Quimbo, Orville Solon, Xylee Javier, and Charles McCulloch. 2014. "The impact of performance incentives on child health outcomes: results from a cluster randomized controlled trial in the Philippines." *Health Policy and Planning* 29 (5): 615–621.

Podgursky, Michael J, and Matthew G Springer. 2007. "Teacher performance pay: A review." *Journal of policy analysis and management* 26 (4): 909–949.

Prendergast, Canice. 1999. "The provision of incentives in firms." *Journal of economic literature* 37 (1): 7–63.

Sulis, Giorgia, Pierrick Adam, Vaidehi Nafade, Genevieve Gore, Benjamin Daniels, Amrita Daftary, Jishnu Das, Sumanth Gandra, and Madhukar Pai. 2020. "Antibiotic prescription practices in primary care in low-and middle-income countries: a systematic review and meta-analysis." *PLoS medicine* 17 (6): e1003139.

Whitehead, Margaret, Göran Dahlgren, and Timothy Evans. 2001. "Equity and health sector reforms: can low-income countries escape the medical poverty trap?" *The Lancet* 358 (9284): 833–836.

WHO. 2021. *WHO Guidelines for malaria.* Licence: CC BY-NC-SA 3.0 IGO.

Witter, Sophie, Atle Fretheim, Flora L Kessy, and Anne Karin Lindahl. 2012. "Paying for performance to improve the delivery of health interventions in low-and middle-income countries." *Cochrane database of systematic reviews,* no. 2.

Yip, Winnie, Timothy Powell-Jackson, Wen Chen, Min Hu, Eduardo Fe, Mu Hu, Weiyan Jian, Ming Lu, Wei Han, and William C Hsiao. 2014. "Capitation combined with pay-for-performance improves antibiotic prescribing practices in rural China." *Health affairs* 33 (3): 502–510.

Tables

Table 1: Incentive amount details, by treatment arm (Back: 3)

Subsidy and incentive amounts

	Control (C)	**Patient discount (T1)**	**Pharmacy incentive (T2)**	**Both (T3)**
Patient discounts? (USD)				
Rapid test	-	$0.90	-	$0.60
ACT (malaria +)	-	$1.10	-	$0.80
ACT (malaria -)	-	$0.00	-	$0.00
Provider incentives (USD)				
Rapid test	-	-	$0.90	$0.20
ACT (malaria +)	-	-	$0.80	$0.10
ACT (malaria -)	-	-	$0.00	$0.00
Transaction completion	-	-	$0.30	$0.30
Total incentive amount (USD)	$0.00	$2.00	$2.00	$2.00

Table 2: Impact on rapid test and antimalarial medication uptake (Back: 5.1)

	Rapid test uptake	ACT uptake	Non-ACT uptake
	(1)	(2)	(3)
Patient discount	.239*	-.126^{+}	-.00479
(γ_{T1})	(0.120)	(0.067)	(0.012)
Pharmacy incentive	.188*	-.0844	-.00391
(γ_{T2})	(0.078)	(0.062)	(0.018)
Patient discount and	.224**	-.153**	.000455
pharmacy incentive (γ_{T3})	(0.062)	(0.041)	(0.013)
Control mean	0.079	0.891	0.037
Wald test p-val ($\gamma_{T1} \neq \gamma_{T2} \neq \gamma_{T3}$)	0.8979	0.5611	0.9122
Wald test p-val ($\gamma_{T1} \neq \gamma_{T2}$)	0.7038	0.5784	0.9613
Wald test p-val ($\gamma_{T1} \neq \gamma_{T3}$)	0.9125	0.7107	0.6766
Wald test p-val ($\gamma_{T2} \neq \gamma_{T3}$)	0.7039	0.2843	0.8108
N	56374	56374	56374

Notes: Standard errors are clustered at the facility level. Controls: months active on platform, urban location, female owner, strata and calendar month FE. Wald test comparisons of difference in marginal effects (γ) between treatment arms. Denominator is all patients that purchased malaria product during study period.Col 1: 45 obs dropped b/c all malaria transactions logged in stratum (1 facility) included an RDT. Col 3: 305 obs dropped b/c no malaria transactions logged in strata (5 facilities) were non-ACTs. $+$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

Table 3: Mechanisms: price pass-through and provider counseling (Back: 5.3)

	Log Prices			Provider counseled
	p_r	$p_{a\|r}$	p_a	patient
	(1)	(2)	(3)	(4)
Patient discount (T1)	-1.25^{**}	-1.35^{**}	0.05	0.03
	(0.18)	(0.10)	(0.14)	(0.07)
Pharmacy incentive (T2)	-0.11	0.16	-0.02	0.16^{*}
	(0.08)	(0.10)	(0.14)	(0.07)
Mixed (T3)	-0.74^{**}	-0.60^{**}	-0.01	0.14^{+}
	(0.10)	(0.09)	(0.14)	(0.07)
Control group mean	\$0.97	\$1.18	\$1.18	0.31
Test p-val ($\gamma_{T1} \neq \gamma_{T2} \neq \gamma_{T3}$)	0.00	0.00	0.84	0.17
Test p-val ($\gamma_{T1} \neq \gamma_{T2}$)	0.00	0.00	0.58	0.07
Test p-val ($\gamma_{T1} \neq \gamma_{T3}$)	0.01	0.00	0.63	0.16
Test p-val ($\gamma_{T2} \neq \gamma_{T3}$)	0.00	0.00	0.96	0.70
N	125	138	138	411
Data source	Admin	Admin	Admin	SP
Unit of analysis	Pharmacy	Pharmacy	Pharmacy	SP

Notes: Columns 1-3 are site-level average outcomes, while column 4 are SP-level observations. SP (Admin) denotes that the data source is the Standardized Patient activity (Point of sales administrative data). Wald test comparisons of differences between arms are included. Column 1 shows mean expenditures on any type of malaria test available in a site. Column 2 shows the price of ACTs, conditional on buying an RDT. Column 3 shows the price ACTs without an RDT purchase. When missing transaction data, prices are filled with study design averages. The analysis in this table is at the price per-unit level. $+\ p < 0.1$, $*\ p < 0.05$, $**\ p < 0.01$.

Table 4: Marginal effect of contract characteristics on patient demand for testing

	$D(\delta)$
$p_r - p_a$ (α)	-0.23**
	(0.08)
$p_{a\|r}$ ($\alpha_{a\|r}$)	-0.22*
	(0.11)
Provider incentives ($g(\pi, t)$)	0.40**
	(0.12)
Profit margin control (π_a)	0.00
	(0.12)
Implied patient's $P(m\|i)$	0.95
	p = 0.02
Control group mean	0.081
N	121

Note: This table presents the results from equation 27, using price variables described in appendix 8.1. Variables $p_r - p_a$, $p_{a|r}$, and $g(\pi, t)$ are in USD. Estimated through 2SLS using treatment assignment indicators, marginal costs of treatment, and marginal costs of testing as instruments. First stage F-statistics are 39, 65, and 67 respectively. The last row shows the implied patient's belief about their probability of having malaria using the following formula: $\alpha_{a|r}/\alpha$. The p value for the patient belief corresponds to a test of whether the implied belief is equal or lower to the true positivity rate of 0.24 using the Bias-Corrected and Accelerated (BCa) bootstrap method. An indicator for whether the price for RDTs was filled-in with the average of each treated group as well as the average profit margin of non-loyalty ACTs were included as control variables.

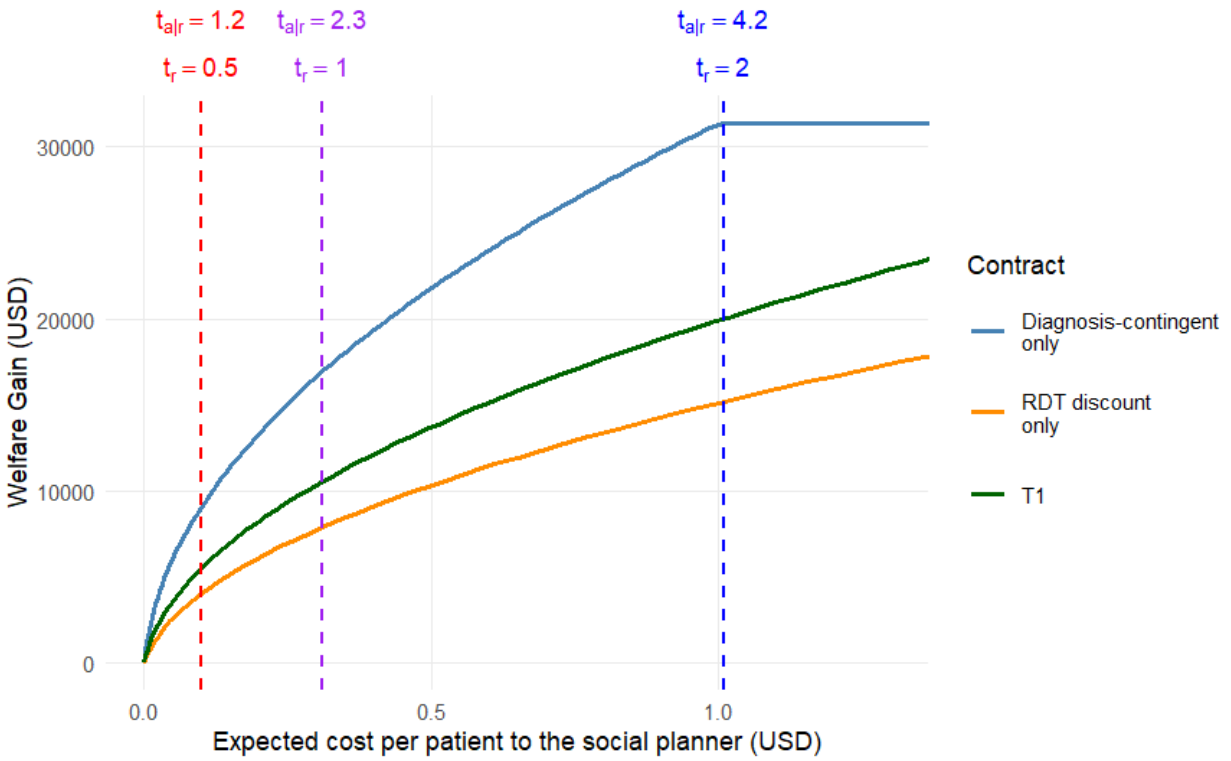
Table 5: Welfare estimates and counterfactual contracts

	Δ_D (1)	*Cost* (2)	Weight (3)	C.E. (4)	Welfare (5)
Panel A: Direct welfare estimates from RCT study arms					
Patient discount (T1)	0.24	0.37	0.64	0.65	1.41*
					(0.21)
Pharmacy incentive (T2)	0.19	0.37	0.64	0.51	1.32*
					(0.13)
Mixed patient and pharmacy (T3)	0.22	0.40	0.64	0.56	1.36**
					(0.10)
Panel B: Counterfactual patient contracts					
(B1) RDT discounts only	0.23	0.31	0.64	0.74	1.47**
					(0.18)
(B2) Diagnosis-contingent ACT discounts only	0.25	0.09	0.64	2.79	2.79*
					(0.90)
(T1) Experimental contract	0.30	0.44	0.64	0.68	1.44**
					(0.10)
Test: B1 = T1					p = 0.83
Test: B2 = T1					p = 0.11
Test: B1 = B2					p = 0.20
Panel C: Counterfactual provider contracts					
(C1) RDT incentives only	0.41	0.49	0.64	0.83	1.53**
					(0.16)
(C2) Diagnosis-contingent ACT incentives only	0.11	0.05	0.64	2.14	2.37**
					(0.42)
(T2) Experimental contract	0.25	0.46	0.64	0.55	1.35**
					(0.10)
Test: C1 = T2					p < 0.01
Test: C2 = T2					p < 0.01
Test: C1 = C2					p < 0.01
Test: B2 = C2					p = 0.63

Notes: Panel A displays the estimated welfare effects from study arms, using inputs from table C12 and following equation 19. Panels B and C display the welfare effects of different counterfactual contract designs. Counterfactual effects are estimated using the fitted results from the linear probability model estimated in Table 4 using the average input values in the control group as baseline. Column 4 is calculated by dividing (1) by (2). Column 5 is calculated by multiplying (4) by (3) and adding 1 (to display welfare gains for patients and pharmacies per dollar of cost to the social planner). Standard errors calculated through the delta method holding columns (2) and (4) as constant. Stars denote significance of the test for whether welfare increased as a result of the policy ($Welfare = 1$). $+\ p < 0.1$, $*\ p < 0.05$, $**\ p < 0.01$.

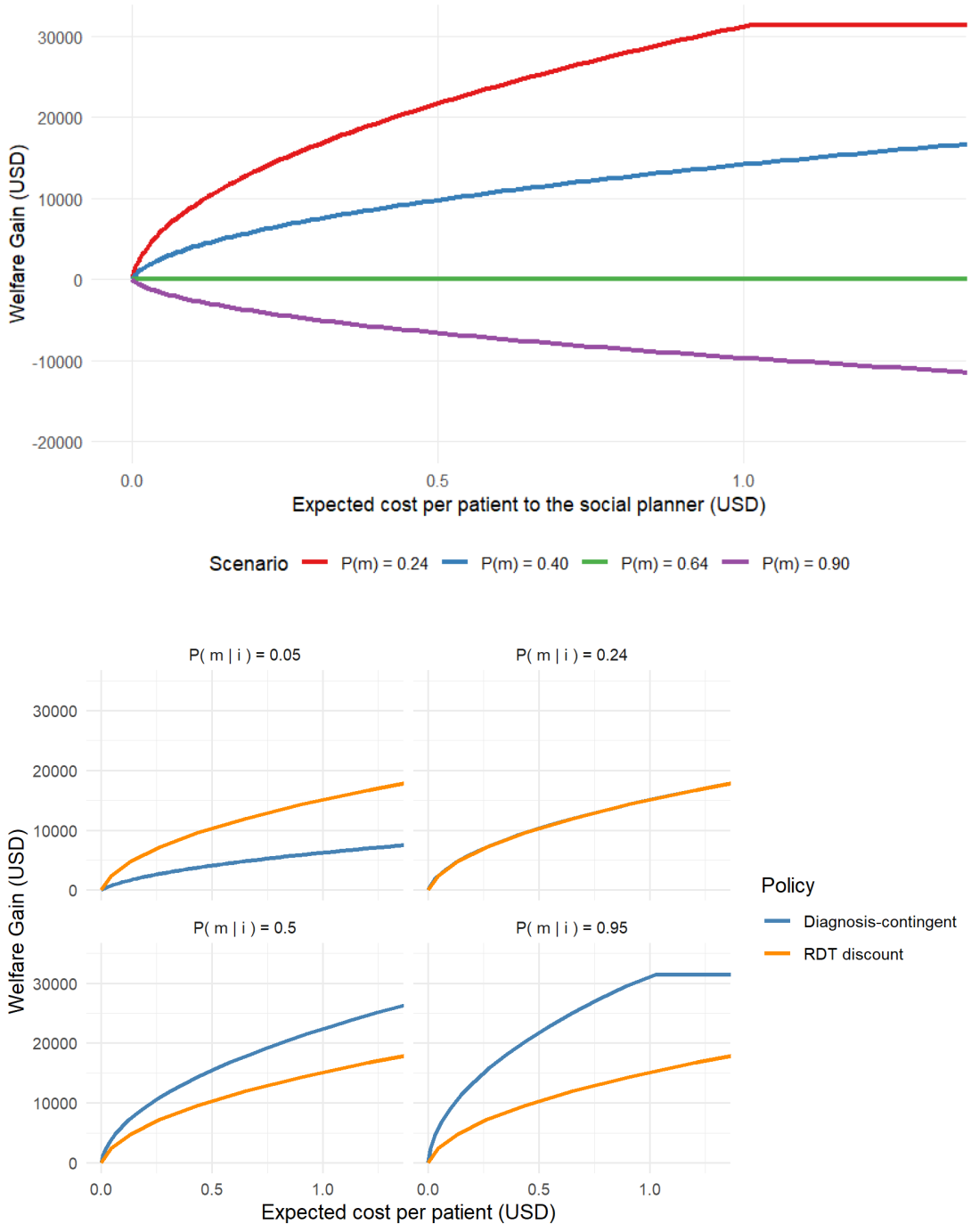
Figures

Figure 1: Total Welfare as a function of subsidy size (Back: 7.2)



Notes: This plot displays the counterfactual expected total welfare in the study sample as a function of the expected cost per patient to the social planner of three alternative contracts directed at the patient: diagnosis-contingent ACT incentives only, RDT discounts only, and a mixture of both contracts (using the ratios from T1). The dotted lines mark three reference contracts: one that gives a free ACT contingent on a positive diagnosis (red), one that offers a free RDT (purple), and one that maximizes social welfare through a diagnostic-contingent ACT incentive only. The labels above the dotted lines denote the subsidy sizes ($t_{a|r}$ and t_r) that correspond to each of the non-mixture contract types at that point in USD.

Figure 2: Total Welfare effects as a function of positivity rates and patient beliefs (Back: 7.2)



Notes: This figure is analogous to figure 1 but tests for different assumptions on $P(m)$ and $P(m|i)$. The top sub-figure shows welfare effects implied by models assuming different true positivity rates $P(m)$ for diagnosis-contingent contracts only (RDT discounts follow a similar pattern). The bottom sub-figure shows the relative cost-effectiveness of diagnosis contingent only and RDT discount only contracts as a function of patient beliefs $P(m|i)$. For additional details see notes in figure 1.

Appendix A. Site-level test positivity rates detail

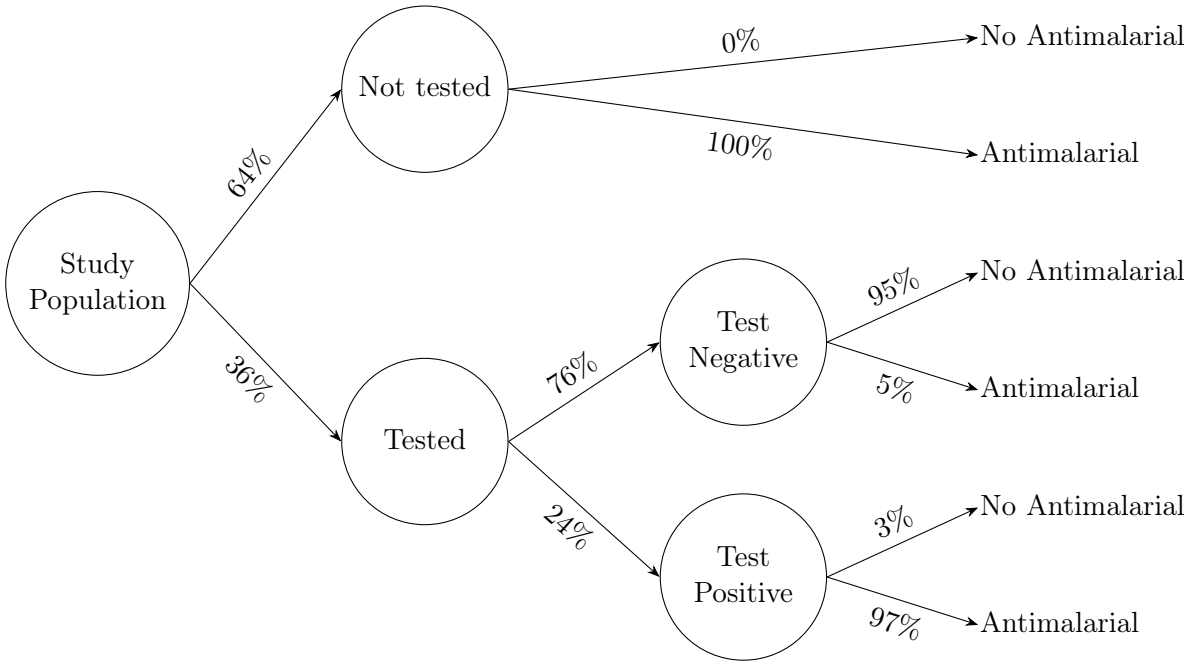
In the intervention arms (T1, T2, T3), test positivity for the tested sample - Pr(malaria positive | tested) - is observed directly from transaction records for patients that tested for malaria using the incentivized rapid tests. We calculate facility-level average test positivity rates for these arms. In the control group, we do not observe test positivity for individual patients in the transaction records. Instead, we estimate Pr(malaria positive | tested) using administrative aggregate testing records provided by a subset of control group sites that keep records on malaria positivity rates. Eight control group sites kept aggregate testing and positivity rate information of all tests conducted between January - February 2022. We calculate facility-level positivity rates using this information. We obtain an overall Pr(malaria positive | tested) that is the weighted average of these treatment-arm facility-level probabilities, weighted by the number of facilities that contributed to these estimates (Control group: 8, T1: 30, T2: 31, T3: 32).

We also estimate the overall, unconditional malaria positivity rate - Pr(malaria positive) by leveraging an independent lab testing activity done in the control group sites. Trained technicians enrolled eligible adult pharmacy clients (eligibility = over 18, not pregnant, seeking care for suspected malaria, not taking antimalarials) and offered them a free RDT to test for malaria after they had concluded their routine pharmacy visit. Technicians collected basic demographic data, data on purchases made at that pharmacy visit, and RDT test result. We calculate test positivity rate for clients that purchased a malaria product, in order to draw a comparable sample to that in the main administrative data. Thirteen control group sites contributed data from 35 randomly administered tests for a final Pr(malaria positive) = 0.4. This measure is not significantly different to the positivity rate conditional on testing obtained through administrative records. Given that this measure is only available for 13 sites and is very noisy, in contrast to the conditional on testing measure (available for 101 sites), we assume that $P(m|tested) = P(m)$ in our analysis and use the administrative data as our preferred measure. Sensitivity tests for the analysis are included assuming $P(m) = 0.4$ instead.

Appendix B. Supplementary Tables and Figures

8.1 Appendix Tables and Figures

Figure B1: Patient participation in the program, positivity rate, and treatment decisions in treatment arms



Notes: This diagram summarizes conditional probabilities of purchasing tests and ACTs among the population exposed to treatment in the study using available sources of data. The first level summarizes the share of transactions that included a test (including RDTs and alternative methods outside of the program) across malaria transactions in treatment arms. The second level presents the positivity rate conditional on testing using administrative data from RDT results. The third level summarizes the conditional probability of a transaction including an antimalarial. This figure shows that absent testing, the vast majority of patients chose an ACT as tratment. Conditional on testing, almost all patients follow their test results when making a treatment decision.

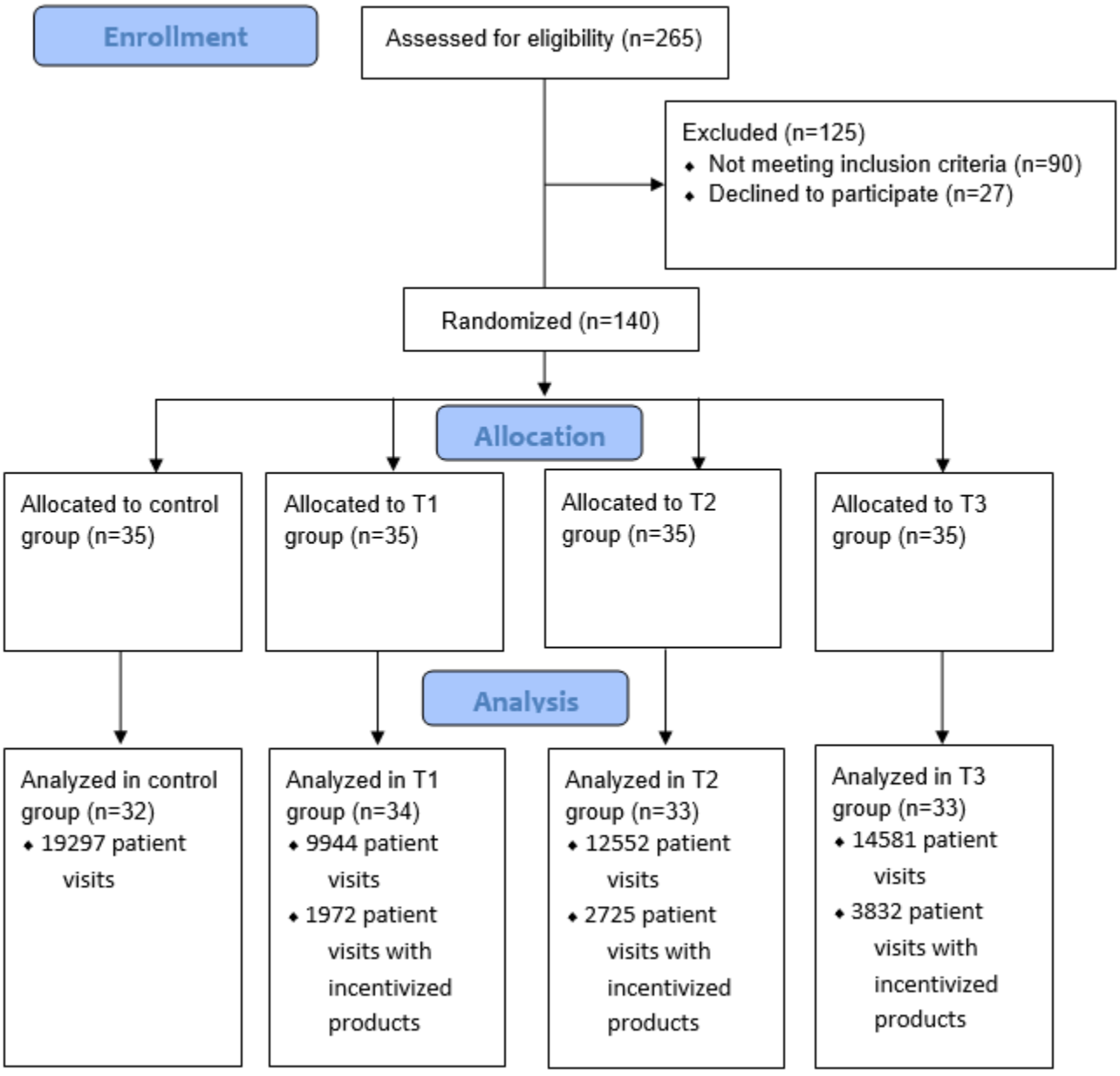


Figure B2: Study flow diagram (Back: 3)

Table B6: Study timeline (Back: 3.2)

Jun-Dec '21	**Experiment launch**: baseline pharmacy survey with 233 pharmacy owners and staff from all 140 sites; staggered onboarding of 140 pharmacies to intervention and study
Aug '21-Feb '22	**Monitoring**: implementation team monitors intervention implementation through regular outreach calls and random site visits; ongoing administrative data collection through digital platform
Oct '21-Jan '22	**Patient exit survey**: survey of random sample of 1654 adult clients who seek care for malaria-like symptoms
Dec '21-Feb '22	**Standardized patient visits**: 412 mystery shopper visits by enumerators presenting as suspected malaria patients, to obtain data on patient-pharmacist interaction, implementation fidelity, and quality of care
Jan-Feb '22	**Control group testing**: testing of random subset of 230 pharmacy clients at control group sites to obtain test positivity rate
Mar '22	**Pharmacy endline survey**: survey of all pharmacy staff and owners at conclusion of the data collection period

Table B7: Baseline balance table (Back: 4)

	Patient discount	Pharmacy incentive	Mixed patient and pharmacy	Control mean
In malaria-endemic county	-0.11	-0.09	0.03	0.89
	(0.09)	(0.09)	(0.09)	
Urban	0.09	0.20+	0.14	0.20
	(0.11)	(0.11)	(0.11)	
Participated in earlier study phase	0.03	0.00	-0.09	0.17
	(0.09)	(0.09)	(0.09)	
Below median malaria sales	0.09	-0.00	0.09	0.37
	(0.12)	(0.12)	(0.12)	
Number of staff	-0.03	-0.04	-0.12	1.54
	(0.12)	(0.12)	(0.12)	
Female owner	-0.20*	-0.21*	-0.17+	0.36
	(0.10)	(0.10)	(0.10)	
Average staff age	0.34	0.08	0.20	29.37
	(1.29)	(1.31)	(1.29)	
Average owner age	-1.51	-0.36	-1.42	37.43
	(1.79)	(1.81)	(1.79)	
Percent female staff	-0.09	0.05	-0.01	0.44
	(0.10)	(0.10)	(0.10)	
Pharmacy only, no clinic capability	0.10	0.06	0.12	0.71
	(0.09)	(0.09)	(0.09)	
Average monthly RDT sales	2.68	1.10	-0.12	4.30
	(1.79)	(1.81)	(1.78)	
Average monthly ACT sales	-1.97	-3.55	-7.67	52.84
	(9.98)	(10.10)	(9.96)	
Average monthly malaria sales	-6.29	-9.81	-12.78	64.80
	(11.37)	(11.50)	(11.34)	
Total months active on POS platform	-1.45	-4.15*	-0.74	10.14
	(1.98)	(2.00)	(1.98)	

Notes: Linear probability and linear regression with strata fixed effects. Multinomial logit test for joint orthogonality produces p-value from χ^2 test of 0.46. Significance stars from comparisons with control group: $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$.

Table B8: Baseline balance between facilities in sample and refusals (Back: 4)

Variable	(1) In sample	(2) Declined	(3) P-value difference
Number of months active on digital sales management tool	12.04	16.81	0.01
	(9.43)	(8.52)	
Average monthly malaria sales, 2019-2020	63.39	66.47	0.83
	(63.56)	(75.45)	
Average monthly quality treatment sales, 2019-2020	54.41	61.25	0.59
	(54.26)	(72.61)	
Average monthly rapid test sales, 2019-2020	6.48	5.39	0.62
	(9.93)	(11.26)	
Site was in earlier pilot study phase	0.16	0.23	0.32
	(0.37)	(0.43)	
Site is in an urban area	0.31	0.34	0.69
	(0.46)	(0.48)	
Site is in a malaria endemic county	0.84	0.86	0.84
	(0.37)	(0.36)	
Site is a pharmacy	0.56	0.56	1.00
	(0.50)	(0.51)	
Observations	140	35	

Notes: In sample facilities include those that have been onboarded and those that are pending.
P-value difference between group means from chi-squared tests or t-tests.

Table B9: Effects on log pharmacy profit margins and log patient expenditures

	Log Profits	Log Expenditures
Patient discount (T1)	−0.04	−0.39*
	(0.13)	(0.16)
Pharmacy incentive (T2)	0.22+	−0.29*
	(0.12)	(0.13)
Mixed (T3)	−0.02	−0.37*
	(0.12)	(0.15)
Control group mean	\$0.74	\$2.67
Test p-val ($\gamma_{T1} \neq \gamma_{T2} \neq \gamma_{T3}$)	0.08	0.74
Test p-val ($\gamma_{T1} \neq \gamma_{T2}$)	0.06	0.54
Test p-val ($\gamma_{T1} \neq \gamma_{T3}$)	0.85	0.95
Test p-val ($\gamma_{T2} \neq \gamma_{T3}$)	0.05	0.53
N	131	134

Notes: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$. This table shows the treatment effects on log profits and log expenditures for each study site using an OLS regression. In column 1, the outcome variable is the average markup of products sold within a transaction, aggregated at the site level. In column 2, the outcome variable is the sum of expenditures in a transaction, aggregated at the site level.

Table B10: Inputs for social welfare effect calculations

	T1	T2	T3
(A) Average incentive cost inputs			
$D(\delta)$	0.318	0.267	0.303
$\boldsymbol{P(m)}$	0.24	0.24	0.24
RDT incentive cost	\$0.90	\$1.20	\$1.10
Diagnosis contingent ACT incentive cost	\$1.10	\$0.80	\$0.90
(B) Social welfare weight inputs			
$\boldsymbol{P(m')}$	0.76	0.76	0.76
c_a	\$1.59	\$1.59	\$1.59
c_r	\$0.57	\$0.57	\$0.57

Notes: This table presents the inputs used to estimate the welfare effects in table 5, Panel A. Panel A presents the inputs used to estimate the average per patient cost of the incentives using administrative data on transactions and program implementation costs. Panel B presents the inputs for the social welfare weight lower-bound, using administrative data on marginal costs of treatment and testing, and estimated positivity rates from the random testing activity.

Table B11: Model fit

	Δ_D (1)	$Cost$ (2)	Weight (3)	C.E. (4)	Welfare (5)
Experimental patient contract					
Estimated directly	0.24	0.37	0.64	0.65	1.41*
					(0.21)
Model estimate	0.30	0.44	0.64	0.68	1.44**
					(0.10)
Experimental provider contract					
Estimated directly	0.19	0.37	0.64	0.51	1.32*
					(0.13)
Model estimate	0.25	0.46	0.64	0.55	1.35**
					(0.10)

Note: This table compares the welfare estimates using parameters estimated in table 4 against the estimates computed directly using equation 19 and inputs from table B10. Column 4 is calculated by dividing (1) by (2). Column 5 is calculated by multiplying (4) by (3) and adding 1. Standard errors calculated through the delta method holding columns (2) and (4) as constant. Stars denote significance of the test $Welfare = 1$ (whether welfare increased as a result of the policy). + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

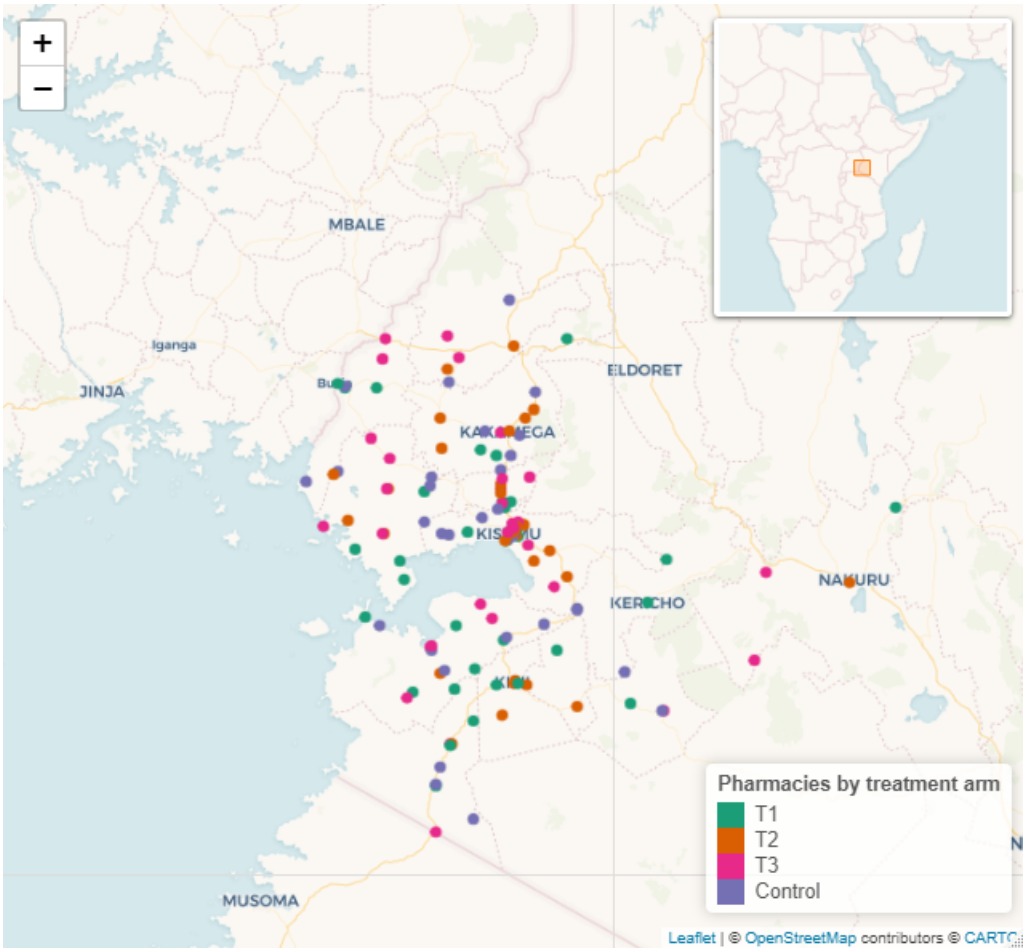


Figure B3: Map of study sites (Back: 3.1)

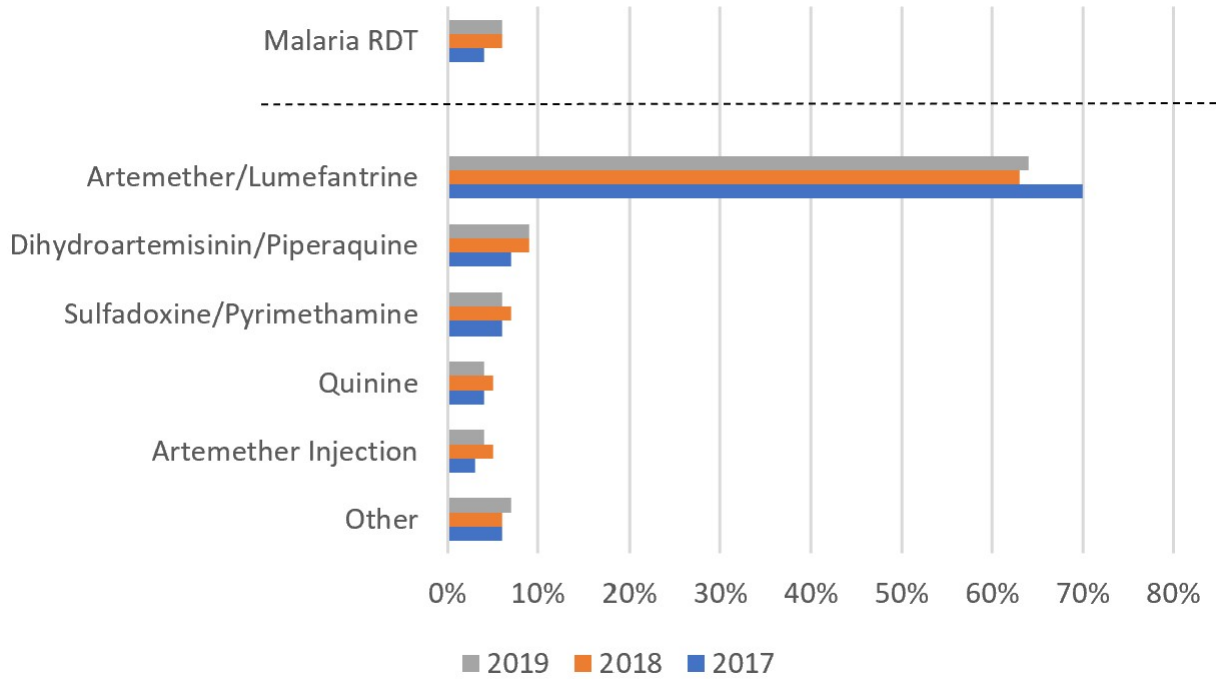


Figure B4: Malaria product sales volumes by type, 2017-2019 (Back: 5.1)

Appendix C. Proxy for provider's incentives

In this appendix we discuss the proxy used for the financial incentives of the provider to recommend $f(p, c, t)$. Given that we do not use a measure of θ for estimation, we need to proxy for $f(p, c, t)$. We first introduce some notation. Let $\Delta P(k) = P(k|\theta = 1) - P(k|\theta = 0)$ for product k. Also, note that in the data $t_a = 0$ in all arms. Then we can express $f(p, c, t)$ as:

$$f(p, c, t) = (\pi_r + t_r)\Delta P(r) + (\pi_{a|r} + t_{a|r})\Delta P(a|r) + (\pi_a)\Delta P(a)$$

Note that $P(a|r) = \boldsymbol{P(m)}P(r)$, since only malaria positive patients who test purchase incentivized treatment. Therefore, $\Delta P(a|r) = \boldsymbol{P(m)}\Delta P(r)$. Hence:

$$\begin{aligned} f(p, c, t) &= \Delta P(r)[(\pi_r + t_r) + \boldsymbol{P(m)}(\pi_{a|r} + t_{a|r})] + (\pi_a)\Delta P(a) \\ &= \Delta P(r)g(p, c, t) + (\pi_a)\Delta P(a) \end{aligned}$$

We argue that $g(p, c, t) = (\pi_r + t_r) + \boldsymbol{P(m)}(\pi_{a|r} + t_{a|r})$ can be used as a proxy for $f(p, c)$ for the purposes of our counterfactuals when using treatment indicators from the RCT as instruments in equation 27. First, note that this component of $f(p, c, t)$ is a measure of only the contract components. Second, note that $\Delta P(r)$ is a constant, so the coefficient for the financial incentives when using $g(p, c, t)$ will be scaled by this constant. Third, the experiment generates exogenous variation in the price of RDTs and diagnosis contingent incentives. If

we assume that pharmacy prices for ACTs do not respond to the interventions, then the treatment assignment is orthogonal to $(\pi_a + t_a)\Delta P(a)$. This is because $\Delta P(a) = -\Delta P(r)$ by definition, where $\Delta P(r)$ is a scaling factor as discussed above. This assumption is reasonable and consistent with the reduced form results on prices since, as seen in table 3, there is no pass-through of pharmacy incentives on prices for RDTs and ACTs and there is no effect on the price of non-loyalty ACTs. However, marginal costs of non-loyalty ACTs (used as an instrument to recover α) can be correlated with π_a. If excluded from estimation, this term could be part of the error term and bias the results, so we include π_a as a control variable in the model estimates.

Appendix D. Inputs for the empirical model

Prices in the empirical model come from administrative transaction data. For all products in our sample, we observe whether they are for a test or for treatment but information on dosage is not always available. Therefore, for all price measures derived from the POS system we use the average transaction expenditures for each category.

In order to make these variables analogous to the prices from loyalty program products, we normalize them in the following way. For tests, we divide the mean expenditure variable by the average number of doses sold in a transaction for testing across all sites. For treatment, let p_a^* denote the dose-level price of AL ACTs outside of the loyalty program, and let P_{treat} be the vector of mean transaction expenditures on treatment outside of the loyalty program at the site level. Then:

$$p_a = P_{treat}\left(\frac{\mathbb{E}[p_a^*]}{\mathbb{E}[P_{treat}]}\right) \tag{28}$$

Hence, this normalization accounts for the fact that treatments in non-loyalty sales might include multiple doses and might be a combination of non-ACT and ACT treatments available in the study site.

With these normalizations in hand, table C12 shows the first stage regression of the instruments used to estimate the empirical model.

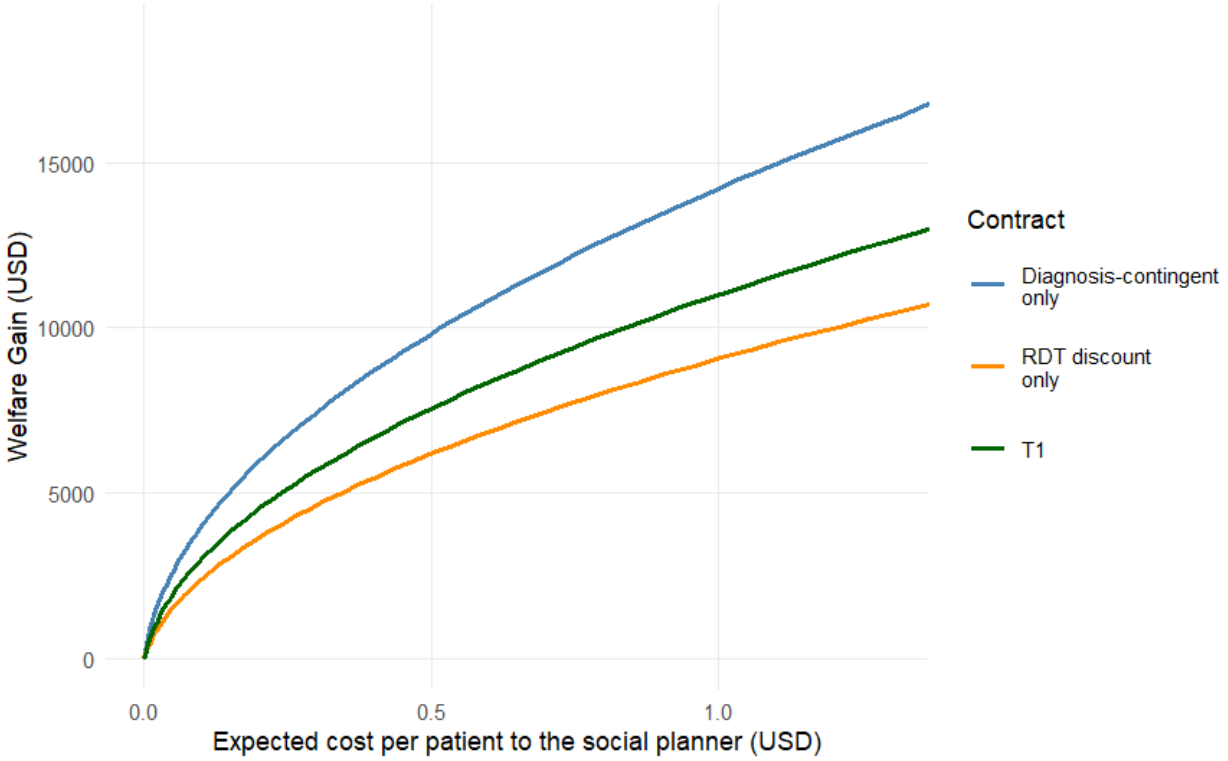
Table C12: Empirical model first stage

	$p_r - p_a$	$p_{a\|r}$	$g(p, c, t)$
Patient discount (T1)	-28.73^{*}	-88.75^{**}	-15.42^{*}
	(13.81)	(11.81)	(7.19)
Pharmacy incentive (T2)	14.15	9.40	66.70^{**}
	(14.62)	(14.10)	(7.73)
Mixed incentive (T3)	-11.20	-59.68^{**}	2.57
	(13.33)	(12.38)	(7.80)
Marginal cost ACTs (c_a)	-0.53^{**}	0.06^{**}	0.03
	(0.08)	(0.02)	(0.02)
Marginal cost RDTs (c_r)	0.82^{**}	-0.03	0.03
	(0.22)	(0.18)	(0.12)
Num. obs.	121	121	121
F statistic	39	65	67

Notes: This table shows the first stage regressions of the instruments used to estimate the empirical model. The f-statistics correspond to the first stage wald-tests of each endogenous variable. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Appendix E. Robustness checks

Figure C1: Total Welfare as a function of subsidy size when $P(m) = 0.40$ (Back: 7.2)



Notes: This figure is analogous to figure 1 but assuming a higher positivity rate of $P(m) = 0.4$. Higher positivity rates make diagnosis contingent contracts have a smaller effect on welfare. For additional details see figure 1.

Table C13: Marginal effect of contract characteristics on patient demand for testing for various positivity rates

	$D(\delta)$	$D(\delta)$	$D(\delta)$
$p_r - p_a$ (α)	-0.23**	0.23**	0.23**
	(0.08)	(0.08)	(0.08)
$p_{a\|r}$ $(\alpha_{a\|r})$	-0.21*	-0.21*	-0.21*
	(0.11)	(0.10)	(0.10)
Provider incentives $(g(\pi, t))$	0.34**	0.28**	0.23**
	(0.10)	(0.08)	(0.07)
Implied patient's $P(m\|i)$	0.93	0.90	0.89
Control group mean	0.081	0.081	0.081
Positivity rate $P(m)$	0.40	0.64	0.90
N	121	121	121

Note: This table presents analogous analysis to 4 but assuming a true probability of having malaria of 0.40, 0.64, and 0.90.