Bednets, Information and Malaria in Orissa[*]

Aprajit Mahajan
Alessandro Tarozzi
Joanne Yoong
Brian Blackburn

July 21, 2011

Abstract

We study the identification and estimation of key parameters in a basic model of technology adoption when specifically collected information on subjective beliefs and expectations about the technology's impact is available. We discuss identification with both non-parametrically and parametrically specified utility as well as parametric and semi-parametric specifications for unobserved heterogeneity. We propose parametric and semi-parametric estimation methods to recover underlying preferences and use the model to study the adoption of bednets among poor households in rural Orissa (India). We carry out counterfactual exercises to examine the effects of price and belief changes on net ownership decisions. The results suggest that net purchase decisions are relatively insensitive to changes from current prices and beliefs. The methods proposes here should have applicability to other discrete choice settings with non-linear indices.

JEL: I1,I3
Key words: Malaria, Expectations, Bednets, Identification, Median Restrictions

1 Introduction

Economists and public health researchers have found that relatively inexpensive welfare-improving technologies are often not adopted by poor households.[1] Bednets are an exemplary case. Research has demonstrated that bednets are very effective at protecting from malaria, particularly among pregnant women and children. In addition, the protective power is significantly increased when the nets are appropriately treated with insecticide.[2] However, the purchase and use of bednets remain low in many malaria-prone areas (Webster et al. 2005, Monasch et al. 2004).

The public health literature suggests many factors contributing to low adoption rates, with cost often being the most cited. Poor households may have lower or simply insufficient willingness to pay for bednets (Guyatt et al. 2002ab, Onwujekwe et al. 2000, 2004). Even households that are willing to pay may not be able to do so if they lack the cash at hand and do not have access to credit.

Subjective preferences and expectations about malaria and the effectiveness of bednets have also been proposed as explanations for low net takeup. A basic difficulty of trying to analyze demand for bednets using observed choices (or any purchase decision more generally) is that such choices may be consistent with many combinations of expectations and preferences. In particular, familiarity with malaria and bednets as well as cultural factors, preferences, and perceptions are important considerations (Onwujekwe et al. 2000, Alaii et al. 2003). For example, individuals may weigh against cost other benefits of nets, such as better sleep, and ignore their usefulness in malaria prevention. In many instances, the purchase decision for bednets is primarily made based on the nuisance level of mosquitoes, rather than any desire to prevent malaria (Louis et al. 1992, Van Bortel et al. 1996, Klein et al. 1995). Studies also show that individuals who understand the connection between mosquitoes and malaria may also attribute the disease to other additional factors, making it difficult to convince them to adopt bednets as a control measure (Agyepong 1992, Ahorlu et al. 1997, Agyepong and Manderson 1999, Hill et al. 2003 and Adongo et al. 2005).

Disentangling the different reasons for non-adoption has important policy implications. For instance, policy recommendations would be quite different if cost played a more important role than beliefs. To the extent that standard expected utility theory is relevant, prices and beliefs (along with other components of the choice problem) affect behavior in non-separable and non-linear ways so that separating the effects of the two will require explicitly modeling household expectations about the risk of contracting malaria as part of the decision process. Such a choice model will usually yield estimating equations (or likelihoods) that are non-linear in parameters and whose identification properties are typically unknown.

This paper is a first attempt at providing both a careful identification analysis of such models as well as the estimation of one such model using data from rural Orissa (India). Our data is unusual in that it contains detailed information on household level beliefs about the likelihood of malaria under various scenarios. We posit a simple discrete choice model that incorporates beliefs, costs and preferences in a standard choice framework. We then study identification of key model parameters while allowing for unobserved heterogeneity in preferences by adopting a semi-parametric framework.[3] Finally, we carry

---

[1]See Duflo et al. (2009) for a review of these arguments in a more general context.

[2]See e.g the extensive survey in Lengeler (2004).

[3]The results here are closest to Matzkin (1991) who studies the problem of non-parametric identification of sub-utility functions when the distribution of the error terms is parametrically specified (see also Matzkin (1992)). See also Pierre-

out a series of counterfactual exercises to provide a first order answer to some of the questions raised above. The results suggest that net purchase decisions are relatively insensitive to changes from current prices and beliefs.

The identification results proposed here are also relevant for other discrete choice settings with non-linear index function specifications. In particular, we show identification results for preference parameters (such as risk aversion parameters) that enter the index function non-linearly while at the same time placing relatively few restrictions on the unobserved heterogeneity in household preferences. Concretely, we show identification results for household varying constant relative risk aversion (CRRA) preferences in the presence of cluster fixed effects while imposing only a conditional median restriction on the unobserved heterogeneity in preferences. Non-linear index functions and unobservable heterogeneity are present in a variety of economic models so that the identification results presented here should be applicable to those contexts as well.

The paper is organized as follows: Section 2 provides a brief literature overview. The nature of the data plays an important role in the theoretical formulation of the model and Section 3 provides an overview of the study location and design and a first description of the data. Section 4 develops a basic static model of technology adoption. Section 5 discusses alternative formulations for preferences and Section 6 discusses possible methods for modeling the unobserved components of preferences. Sections 7 and 8 provide identification results and estimation of the identified model is carried out in Section 9. Section 10 carries out a set of counter-factual exercises based on the estimated model and the conclusions follow.

2   Literature Overview

Recent work by economists has used historical evidence to document the significant improvements in literacy, education and income arising from malaria eradication (Bleakley 2010, Cutler et al. 2010). In the absence of a vaccine, prophylaxis and treatment remain the only two avenues for dealing with malaria. Multiple studies have linked ITN use to reductions in malaria infection as well as related morbidity and mortality.[4] In Orissa, the field setting for this project, ITN use was associated with vector reduction and a 50-60% decline in malaria prevalence in Malkangiri district (Sahu et al. 2003) and Sundargarh District (Yadav et al. 2000). The best universally-accepted evidence of ITN efficacy comes from a large medical trial conducted in Western Kenya where sufficient free ITNs were distributed in randomly-chosen villages to allow all persons to sleep under a net, reducing clinical malaria and moderate-severe anemia by 60% in children under five.[5]

However, despite such evidence, ITN coverage remains woefully low. Estimates for all areas at risk in India indicate only 20% coverage (Korenromp 2005). Many public health specialists attribute low adoption rates to costs as well as beliefs and information broadly construed. Adongo et al. (2005) state that no study examining the competing explanations does so carefully enough to disentangle the unique

---

Andre Chiappori and Salanie (2009) which also studies the problem of inferring risk preferences from discrete choices.

    [4]See for instance Beach et al. (1993); Stich et al. (1994); Abdulla et al. (2001); Leenstra et al. (2003); ter Kuile et al. (2003)a,b.

    [5]See Alaii et al. (2003); Leenstra et al. (2003); Hawley et al. (2003)ab, ter Kuile et al. (2003)a,b and other articles in the same issue.

contributions of cost, lack of knowledge about malaria and the number of alternative attributions. This gap in the literature suggests the need for a model of consumer choice that incorporates perceptions about malaria and its costs, but also importantly, beliefs about the relative efficacy of ITNs in preventing malaria.

The standard approach to integrating expectations into structural models of consumer choice under uncertainty relies on making assumptions about individual information as well as the process of expectations formation. For example, Gonul and Srinivasan (1996) estimate a dynamic structural model of diaper purchase that takes into account endogenously determined expectations of future coupons. Erdem et al. (2003), Sun et al. (2003) and Hendel and Nevo (2006) also estimate structural models of purchase decisions that include consumers' price expectations, while Erdem et al. (2005)'s model of learning about computers and consumer purchase choices incorporates expectations of both price and quality. However, inferring expectations from realizations may be problematic, as misspecification of either the information set or the expectations formation process may lead to incorrect estimates (Manski 2004). Eliciting subjective probability distributions, however, allows researchers to replace these assumptions with data. Since the 1990s, an increasing number of surveys have successfully elicited probabilistic expectations from their respondents.[6] In the developing country context, subjective probabilities have been elicited to study HIV risk (Delavande et al. 2010), weather forecasts and livestock and crop planting decisions (Luseno et al. 2003, Giné et al. 2007, Lybbert et al. 2007) and migration decisions (McKenzie et al. 2009).[7]

More economists have begun using subjective expectations to explain behavioral choices as such data becomes available. Nyarko and Schotter (2002) use stated beliefs about opponents' behavior to predict behavior in a series of experimental games. They find that choice models estimated with stated beliefs outperform those estimated with standard models of belief formation. Lochner (2007) links expectations to criminal behavior and finds that youth with a low perceived probability of arrest are significantly more likely to commit crimes. Hurd et al. (2004) shows that individuals with low subjective survival probabilities retire earlier. Delavande (2008) measures probabilistic expectations about contraceptive methods from young women, and combines them with data on actual contraception choice to estimate a structural model of birth control choice. Heterogeneity in beliefs is used to identify respondents' preference parameters for each pregnancy. In a recent paper, De Paula et al. (2010) study how beliefs about HIV status affect risky behaviour. Van der Klaauw W. (2000) and van der Klaauw and Wolpin (2008) use information about expected future choices to improve precision in the context of a structural dynamic model. Mahajan and Tarozzi (2011) uses expectations about state transitions to identify and estimate a dynamic discrete choice model with unobserved types and time-inconsistent agents. Attanasio and Kaufmann (2009) and Kaufmann (2010) explore the role of agents expectations about the returns to schooling in their decision to invest in schooling in a context of credit constraints while Jensen (2010) examines the link experimentally in the Dominican Republic.

Our work is also related to recent studies that analyze behavioral issues surrounding ITN use and

---

[6]See, for example Dominitz (1998) for a discussion of the Survey of Economic Expectations and Hurd and McGarry (1995) and Hurd et al. (2004) for an analysis of the predictive value of subjective survival probabilities collected during the U.S. Health and Retirement Survey.

[7]See Delavande et al. (2010) for a recent survey.

cost variation. Cohen and Dupas (2010) use an experimental intervention in Kenya to demonstrate that ITN usage conditional on adoption are not significantly affected by variations in the cost of the net (including a price of zero). Hoffmann (2008) uses experimental data to document differences in the intra-household allocation of free versus purchased ITNs in Uganda. Tarozzi et al. (2011) also examine these issues using experimental evidence.

3 Location and Data

With a population of 37 million (2001 Census), Orissa accounts for only about four percent of the total Indian population, but it has one quarter of India's annual malaria cases, 44 percent of *P. falciparum* malaria (the most severe form) and 18 percent of malaria deaths (NVBDCP 2008).[8] The Orissa Human Development Report cites malaria as the "number one public health problem" in the state (Government of Orissa 2004). The state Department of Health and Family Welfare reported 366,000 cases of malaria in 2007, 87% of which were *P. falciparum* (NVBDCP 2008).[9] The 1998-99 round of the National Family and Health Survey (NFHS) revealed self-reported malaria incidence rates between 8.5% and 17.2% in our study districts.

The data used in this paper are part of a household survey completed in May-June 2007 in rural Orissa, India. The survey covered 150 villages in the five district of Bargarh, Balangir, Kandhamal, Keonjhar and Sambalpur.[10] Data were collected for a sample of 1947 households, with a total of 10,641 members. Households were sampled from 150 villages selected from a list of 878 villages where BISWA, a local micro-lender, had an established presence at the end of 2006.[11] We treated each of the five districts as a separate stratum and selected 33 villages from Balangir, 48 from Bargarh, 30 from Keonjhar, 9 from Phulbani and 30 from Sambalpur. Within each selected village we selected up to 15 households from lists of borrowers provided by BISWA.[12]

A household-level questionnaire administered to adult respondents included a comprehensive survey of household demographic, socioeconomic and health characteristics. Importantly, the instrument also includes a detailed section on household beliefs and practices related to malaria and bednets. A subset of the household was tested for malaria infection using rapid diagnostic blood tests. Malaria prevalence (as well as hemoglobin levels) were recorded for all children under five (U5) and their mothers, all

---

[8]These figures likely understate the actual burden of malaria in India, because they are largely based on active and passive case detection which monitor certain geographical areas and population groups disproportionately (Dutta 2000). Such monitoring systems often miss individuals who are parasitemic but asymptomatic because of acquired partial immunity (Vinetz and Gilman 2002). For a recent, broad overview of the malaria burden throughout India see also Kumar et al. (2007).

[9]In the same year, 214 malaria deaths were reported. Malaria mortality rates in India are much lower than in many malarious African countries, due to the relatively better vector-control and health care system in India (Alles et al. 1998).

[10]The study blocks were chosen because the corresponding district medical officers did not plan free ITN disbursals in those areas.

[11]The reason for this choice is that the data are part of the baseline survey for a randomized controlled trial carried out in collaboration with BISWA.

[12]The sampling scheme generates different probability of selection for households from different villages. Sampling weights were computed as the inverse of such probabilities and have to be used as inflation factors for the calculation of statistics representative of the study population.

pregnant women and a random sample of working-age adults (15 to 60-year old).[13]

Table 1 includes selected summary statistics. The 1947 households in the sample have 5.5 members on average. The vast majority are Hindu (with only 6 percent of Christians in Sambalpur). Overall, scheduled castes and tribes and "other backward castes" account for more than 90 percent of the study population. Mean per capita expenditure—calculated by summing reports for 18 different item categories—is Rs 655, which is approximately twice as large as the poverty line in rural Orissa in 2004-05.[14] The mean number of bednets per head is close to 0.3, but very few of these are treated with insecticide, so that the overall mean number of ITNs per head is only 0.04. This is despite the high malaria incidence: 55 percent of households report at least one case of malaria in the six months preceding the interview. The results of the biomarkers, which measure prevalence at the time of the test (and only in rare cases can detect past infections more than one-two month old) confirm high prevalence, with 12 percent of blood samples testing positive, almost always with *falciparum* infection. Overall, the fraction of tested individuals with anemia (that is, hemoglobin levels below 10g/dl) is 23 percent for males and 35 percent among women. Note also that the survey was completed in late spring, that is, during the dry season, when malaria rates drop. That malaria prevalence may be higher during the rainy summer is also confirmed by reports about bednet use. While 13 percent of individuals slept under bednets the night immediately before the interview, a much higher fraction (55 percent) report regularly sleeping under a net during times of peak mosquito activity.

Respondents' beliefs about the effectiveness of ITNs in malaria protection are an important element in our purchase decision model. Accurate measurement of beliefs is complicated by the low schooling level of most of our respondents and their unfamiliarity with the concept of probability. To elicit subjective probabilities of events we asked respondents to hold up a number of fingers increasing in the perceived likelihood that the event will happen. Hypothetical examples were first introduced by the interviewer to make sure that the respondent understood the rationale. We estimate the subjective probabilities by dividing the number of fingers held by ten.[15] The survey instrument included questions about the probability for an adult, a child under the age of six (U6) or a pregnant woman (PW) of contracting malaria in the next year. Crucially, separate beliefs were elicited for hypothetical scenarios where the individuals make regular use of an ITN, an untreated net or no net at all.[16]

Overall, only two individuals did not respond to the corresponding questions. Reassuringly, very few respondents report that the probability of getting malaria is higher if one sleeps regularly under a net than if one doesn't (19 for U6, 23 for adults and 21 for PW). The vast majority of respondents (1797 for U6, 1775 for adults and 1771 for PW) report probabilities lower in the former than in the latter case, while equal probabilities are also relatively rare (130 for U6, 148 for adults and 154 for PW), and

---

[13]Individuals were tested for malaria with the Binax Now® malaria RDT. This antigen-based test provides accurate diagnosis for current malaria infection (Moody 2002).

[14]The line was Rs 326, which is close to one US dollar per day per person using PPP conversion rates (Government of India 2007).

[15]Note that we do not attempt to measure ranges of probability, so that our data do not allow to identify the degree of uncertainty around the reports.

[16]For instance, one question asked: "imagine first that your household [or a household like yours] does not own or use a bed net. In your opinion, and on a scale of 0-10, how likely do you think it is that a child under 6 that does not sleep under a bed net will contract malaria in the next 1 year?" Questions for different demographic groups and bednet use were asked using analogous wording.

almost always limited to answers equal to .50 or 1. Figure 1 reports histograms of the elicited beliefs. The means and standard deviations of the beliefs are reported in Table 2.

The three graphs at the top in Figure 1 show the distribution of the beliefs for individuals who sleep regularly under a treated bednet. The graphs in the middle and bottom rows show the distributions for individuals who sleep under an untreated net or who do not use nets at all. A few clear conclusions emerge. First, there is remarkably little difference in the reported beliefs for individuals belonging to different demographic categories. This is somewhat surprising because adults, who are likely to have developed partial immunity, are less likely to develop symptoms. If perceptions about malaria risk depend on the observing symptoms, one could have expected to find higher perceived risk for children than for adults. Second, both bednets and re-treatment with insecticide appear to be widely recognized as very effective at reducing malaria risk. Third, the elicited beliefs are strongly concentrated over the focal figures 0, 5 and 10. About three quarters of respondents believe that if nets are not used one will certainly get malaria, and approximately the same fraction believes that regular use of treated nets will virtually wipe out all risk. According to about half of respondents, there is instead a 50 percent chance of developing malaria if an untreated net is used.[17] On the other hand, there remains a degree of variation in the beliefs which can be exploited in the structural estimation that will follow in Section 9.

4   A Basic Model of Net Adoption

In principle, the decision to purchase an ITN is most appropriately modeled as a dynamic problem with net purchase occurring in the first period and realized malaria status in the second period. This is particularly true when analyzing retreatment decisions. However, given the cross-sectional nature of the data, we consider a static optimization problem focusing on the issue of net ownership. In this framework, therefore, both net purchase and malaria status are revealed in the same period so we abstract away from time discounting issues.

To fix ideas, consider a static individual optimization problem for an agent who needs to decide whether to purchase a net. Let $v(c, b, s, \mathbf{z})$ denote the agent's utility function where c denotes consumption of all other goods, b is a binary variable equal to one if the agent purchases a net, s is a dichotomous random variable equal to m ("malaria") if the agent contracts malaria and h ("healthy") otherwise. Finally, $\mathbf{z}$ is a vector of exogenous variables. The components of $\mathbf{z}$ are allowed to depend upon malaria status s, but that dependence is suppressed in the notation. Throughout the paper we use bold fonts to indicate (column) vectors and matrices, while " $'$ " denote transpose.

Assuming a unique interior solution for consumption, the agent will purchase a net if

$$\mathbb{E}\left[v(c^*(b, s, \mathbf{z}), b, s, \mathbf{z}) \mid b = 1\right] \geq \mathbb{E}\left[v(c^*(b, s, \mathbf{z}), b, s, \mathbf{z}) \mid b = 0\right]$$

where $c^*(b, s, \mathbf{z})$ denotes planned optimal consumption given purchase status b, malaria status s and exogenous variables $\mathbf{z}$. Expectations are taken with respect to malaria status s, and we assume that the

---

[17]The clumping of eliciting probabilities has elicited some interest. Manski and Molinari (2010) argue that this clumping is usefully thought of as rounding error and that in the presence of multiple responses and a uniformity assumption the reported point estimates can be converted into a interval within which the true belief is contained. This is an interesting approach, but combining it with a structural approach is left for future work.

purchase decision is taken based on *subjective* beliefs about the probability of falling sick conditional on the purchase decision b. Comparing indirect utilities in this fashion forms the basis of the estimation procedure and also provides the rationale for the variables we include.

To allow for unobserved heterogeneity across agents, we partition $\mathbf{z}$ as $(\tilde{\mathbf{x}}, \epsilon(0), \epsilon(1))$ where $\tilde{\mathbf{x}}$ are observed exogenous variables and $\epsilon(b)$ is an unobserved component of preferences that depends upon net ownership status. For instance, ϵ may capture the agent's comfort level with a net. To simplify the problem further, we assume that ϵ enters additively in the utility representation as

$$v(c, b, s, \mathbf{z}) = u_{bs}(c, \tilde{\mathbf{x}}) + \epsilon(b)q(\tilde{\mathbf{x}}). \tag{1}$$

Allowing $\epsilon(b)$ to enter non-additively will complicate identification considerably. We do, however, allow for purchase decisions to depend upon interactions between observables and unobservables through the function $q(\tilde{\mathbf{x}})$. Using the specifications above, and suppressing dependence upon $(c, \tilde{\mathbf{x}})$ for simplicity, the agent's net purchase decision is given by

$$b^* = \mathbb{I}\left\{\mathbb{E}\left(u_{1s} - u_{0s} \mid b = 0\right) + \delta(u_{1h} - u_{1m}) + (\epsilon(1) - \epsilon(0))q \geq 0\right\} \tag{2}$$

where $\pi \equiv P(s = m \mid b = 0)$ and $\delta \equiv \pi - P(s = m \mid b = 1)$ so that δ is the perceived reduction in the probability of contracting malaria when using a net. The econometrician observes the vector $(b, c(b, s), \tilde{\mathbf{x}}, \pi, \delta)$ where consumption is elicited from households for each possible net purchase and malaria status combination.[18] To clarify the exposition that follows, we rewrite the optimal purchase decision as an explicit function of preferences (captured by $u(\cdot)$)

$$b^* = \mathbb{I}\left\{g(u(\cdot), \mathbf{x}) + \Delta\epsilon q(\mathbf{x}) \geq 0\right\}, \tag{3}$$

where $\mathbf{x}$ is redefined to equal the vector $(c, \pi, \delta, \tilde{\mathbf{x}})$, $\Delta\epsilon = \epsilon(1) - \epsilon(0)$ and $g(\cdot)$ is defined by a comparison with (2) above. Note that the formulation in (3) is slightly more general than (1), because it allows the function $q(.)$ to depend on all the components of $\mathbf{x}$ and not just $\tilde{\mathbf{x}}$. This will be useful when accounting for measurement error in beliefs, as discussed in Appendix B. The object of interest is the utility function $u(\cdot)$. The extent to which we can learn about preferences depends critically upon the assumptions we are willing to make about $u(\cdot)$ and the unobservable component of choice $\Delta\epsilon q(\mathbf{x})$.

In what follows, we will explore the limits to knowledge of $u(\cdot)$ under two alternative assumptions on the unobservable component $\Delta\epsilon q(\mathbf{x})$. We will first assume a particular parametric form for this term and impose strong independence restrictions between the observable and unobservable determinants of choice. Second, we will adopt a much weaker assumption that allows for dependence between the errors and observed covariates. The first assumption has the advantage that estimation and inference

---

[18]In practice, this is done by eliciting purchase costs and malaria costs from survey respondents and subtracting these from stated consumption. A more correct (and non-parametric) approach would require positing

$$c(b, s) = \sum_{b \in \{0,1\}, s \in \{m,h\}} C_{bs}.$$

Then, the C_{bs} would be unknown parameters of interest and the econometrician would only observe stated consumption in one state of the world. This addition to the model would significantly complicate estimation and so is ignored here. We note though that the identification results stated in the subsequent sections would still hold with this addition to the model.

proceed along well established lines (conditional upon identification) whereas the second assumption is much more robust to the presence of dependencies between observable and unobservable determinants of choice.

Mirroring the discussion above, we also explore identification under alternative assumptions on $u(\cdot)$ for a given set of assumptions on the unobservable components of preferences. We start by specifying a parametric form for preferences and proving a point identification result. We next explore identification when the utility function is not specified parametrically.

5 Utility Specifications

5.1 Non-Parametric Utility Specification

Rewriting (3) to show the dependence of preferences upon covariates explicitly we obtain

$$b^* = \mathbb{I}\left\{\mathbb{E}\left(u_{1s}\left(\mathbf{x}\right) - u_{0s}\left(\mathbf{x}\right) \mid b = 0\right) + \delta\left(u_{1h}\left(\mathbf{x}\right) - u_{1m}\left(\mathbf{x}\right)\right) + \Delta\epsilon q(\mathbf{x}) \geq 0\right\}.$$

Next, we define the quantities

$$
\begin{aligned}
\Delta_1\left(\mathbf{x}\right) &\equiv u_{1m}\left(\mathbf{x}\right) - u_{0m}\left(\mathbf{x}\right) \\
\Delta_2\left(\mathbf{x}\right) &\equiv u_{1h}\left(\mathbf{x}\right) - u_{0h}\left(\mathbf{x}\right) \\
\Delta_3\left(\mathbf{x}\right) &\equiv u_{1h}\left(\mathbf{x}\right) - u_{1m}\left(\mathbf{x}\right)
\end{aligned}
\tag{4}
$$

which are the fundamental objects of interest in the non-parametric setting and can be interpreted intuitively. The quantity $\Delta_1\left(\mathbf{x}\right)$ measures the impact of bednet use on observables-dependent utility among agents with malaria, $\Delta_2\left(\mathbf{x}\right)$ measures the corresponding impact among healthy individuals and $\Delta_3\left(\mathbf{x}\right)$ denotes the difference in observables-dependent utility among ITN users with and without malaria. The non-parametric model can then be written as

$$b^* = \mathbb{I}\left\{\pi\Delta_1(\mathbf{x}) + (1 - \pi)\Delta_2(\mathbf{x}) + \delta\Delta_3(\mathbf{x}) + \Delta\epsilon q(\mathbf{x}) \geq 0\right\}.
\tag{5}$$

Identification in this model therefore refers to identification of the $\Delta(\cdot)$ functions. Note that with a non-parametric specification we cannot directly identify the level utilities $u_{bs}(\cdot)$ since the $\Delta(\cdot)$ functions do not uniquely identify them without further normalizations.

5.2 Parametric Utility Specification

The non-parametric specification outlines the limits of what can be learned without placing assumptions on the utility function. However, in principle, much more can be learned if one is willing to impose more structure on the utility function. In this section, we show that restricting attention to a class of parametric utility functions enables us to study whether the entire utility function itself (and not just the differentials identified above) can be identified given the data.

Understanding household attitudes towards risk and the relationship between these attitudes and the household's net purchase decision allows us to test whether the marginal utility of consumption is state dependent or whether the household's coefficient of risk aversion is state dependent. In a strictly

non-parametric setting such an analysis is not possible since, as we saw above and formally show below, the utility function itself is not directly identified. In such a case it is *a fortiori* not possible to identify a household's attitudes towards risk.[19] For this purpose, we therefore consider a convenient parametric representation of preferences that belong to the constant relative risk aversion (CRRA) family. We begin by positing a variation on this class that accounts for household level covariates and which is flexible in how net ownership affects utility. We specify the utility function in (1) as

$$v(c, b, s; \mathbf{x}_a, \mathbf{x}_b, \mathbf{x}_m) = (\alpha_0 + \alpha_1 b)c(b, s)^{1 - \gamma(\mathbf{x}_a, s; \boldsymbol{\tau})} + b\alpha_2 + b\mathbf{x}_b'\boldsymbol{\alpha}_3$$
$$+ \, \mathbb{I}(s = m)\alpha_4 + \mathbf{x}_m'\boldsymbol{\alpha}_5 \mathbb{I}(s = m) + \alpha_6 b \mathbb{I}\{s = m\} + \epsilon(b)\, q(\mathbf{x}), \tag{6}$$

where $\mathbf{x}_a$, $\mathbf{x}_b$ and $\mathbf{x}_m$ are subsets of the vectors of observable covariates $\mathbf{x}$, $(\alpha_0, ..., \alpha_6, \boldsymbol{\tau})$ are parameters and where we impose that the risk aversion coefficient $\gamma \in (0, 1)$.[20] The restriction to the unit interval is easily relaxed to allow γ to belong to any open interval $(0, c)$ for a known constant c. The choice of c is consistent with findings in experimental studies of the elicitation of risk-aversion parameters (see e.g. Binswanger, 1980) which find that a substantial fraction of poor households in rural India have rates of risk-aversion within this interval.[21]

The risk-aversion parameter γ is allowed to depend upon exogenous agent characteristics $\mathbf{x}_a$ as well as malaria status (which implies state dependent utility). The dependence upon malaria status is intended to capture the notion that households may be more (less) risk averse when they are sick (healthy). Note that in principle, the optimal consumption decision also depends directly upon $\mathbf{x}$, but we ignore this dependence in the subsequent analysis for tractability considerations. Besides appearing in the unobserved component $\epsilon(b)$, net ownership affects utility directly (through α_2 and, for malarious individuals, α_6), through consumption c (via the budget constraint) as well as through the vector $\mathbf{x}_b$ (via the coefficient vector $\boldsymbol{\alpha}_3$). The latter captures the heterogeneity of preferences over net ownership. It seems reasonable that households vary in their attitudes towards net ownership and to the extent that this variation is captured by the observables $\mathbf{x}_b$, we can account for it directly. Analogously, the coefficient α_4 capture the direct effect of malaria upon utility while the parameter vector $\boldsymbol{\alpha}_5$ captures the heterogeneity across households in their utility losses from malaria. Households are likely to have diverse coping mechanisms for dealing with illness in general and the variables in $\mathbf{x}_m$ are intended to capture the resultant utility differentials across households. The interaction terms between net ownership and malaria status (α_6) captures a possible notion of regret where households may derive lower utility from net ownership if they do succumb to malaria while owning a net. Alternatively, the utility of a bednet could be higher for malarious individuals because of the perceived reduction in the risk of spreading the disease.

Given this specification of preferences, the latent index b^* can be written as

$$b^* = \mathbb{I}\{g(\mathbf{x}; \boldsymbol{\alpha}, \boldsymbol{\tau}) + \Delta\epsilon \, q(\mathbf{x}) \geq 0\}, \tag{7}$$

---

[19]If the utility function were directly identified and were sufficiently differentiable one could directly construct estimates of risk aversion based on calculating the second derivative of the utility function with respect to consumption.

[20]Note that the formulation in (6) does not include the main effects of $\mathbf{x}$ on $v(.)$. We choose to do so because these main effects are not identified, as the identification proof will rely on differences between utility by malaria status s, see the proof of Lemma 2 below.

[21]Note that since γ is between 0 and 1 by assumption we do not need to divide the utility function by $(1 - \gamma)$ as is done in unrestricted specifications of the CRRA function.

where $\Delta\epsilon \equiv \epsilon(1) - \epsilon(0)$, $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \alpha_2, \boldsymbol{\alpha}_3', \alpha_4, \boldsymbol{\alpha}_5', \alpha_6)$ and

$$g(\mathbf{x}; \boldsymbol{\alpha}, \boldsymbol{\tau}) \equiv \alpha_2 + \mathbf{x}_b' \boldsymbol{\alpha}_3 - \alpha_4 \delta - \mathbf{x}_m' \boldsymbol{\alpha}_5 + \alpha_6(\pi - \delta)$$

$$+ (\alpha_0 + \alpha_1) h_1(\pi, \delta, c, \mathbf{x}_a, \boldsymbol{\tau}) - \alpha_0 h_0(\pi, c, \mathbf{x}_a, \boldsymbol{\tau})$$

$$h_1(\pi, \delta, c, \mathbf{x}_a; \boldsymbol{\tau}) \equiv (1 - \pi + \delta) c(1, h)^{1 - \gamma(\mathbf{x}_a, h; \boldsymbol{\tau})} + (\pi - \delta) c(1, m)^{1 - \gamma(\mathbf{x}_a, m; \boldsymbol{\tau})}$$

$$h_0(\pi, c, \mathbf{x}_a; \boldsymbol{\tau}) \equiv (1 - \pi) c(0, h)^{1 - \gamma(\mathbf{x}_a, h; \boldsymbol{\tau})} + \pi c(0, m)^{1 - \gamma(\mathbf{x}_a, m; \boldsymbol{\tau})}.$$

The parameter vector of interest is $(\boldsymbol{\alpha}, \boldsymbol{\tau})$. To ensure that the risk aversion parameter is bounded on the unit interval, we use the following parameterization

$$\gamma(\mathbf{x}_a, s; \boldsymbol{\tau}) = \frac{e^{\mathbf{x}_a' \boldsymbol{\tau}_1 + \tau_0 \mathbb{I}\{s = m\}}}{1 + e^{\mathbf{x}_a' \boldsymbol{\tau}_1 + \tau_0 \mathbb{I}\{s = m\}}}, \tag{8}$$

where $\boldsymbol{\tau} \equiv (\boldsymbol{\tau}_1', \tau_0)'$.

6   Error Specifications

We specify two alternative models for the error term in (3). The first specification supposes that $q(\mathbf{x}) = 1$ and imposes a parametric assumption on $\Delta\epsilon$ that includes an additive village fixed effect. This allows us to develop a parametric non-linear maximum likelihood procedure that, conditional upon identification, has standard statistical properties (that is, parametric convergence rates and limiting normal distributions) that are easy to analyze. This model, however, has the weakness that household level unobserved heterogeneity (as captured in the last term of (3)) must be independent of observable heterogeneity.

The second model overcomes this weakness by making no assumptions on the form of $q(\mathbf{x})$ or any parametric assumption on $\Delta\epsilon$ beyond assuming that it is median independent of the covariates $\mathbf{x}$. This model allows for interactions between the observed and unobserved components of preferences as well as for the presence of village level fixed effects. The disadvantage of this specification is that the statistical properties of the estimator are not straightforward. Conditional upon identification, we conjecture (based on the results of Kim and Pollard 1990) that the estimator is cube root consistent with a non-normal limiting distribution. We use subsampling to conduct inference. Taken together, the results from these two strategies will provide a robust picture of net adoption determinants in the data.

6.1   Parametric Error Specification

In this subsection we assume that unobserved household level heterogeneity is independent of the observed determinants of choice with a known distribution. We work with the standard logistic model since it is analytically tractable when dealing with fixed effects (in fact, it is perhaps the only tractable error distribution in dealing with fixed effects in binary choice, see Chamberlain 1993). We assume that the researcher observes a random sample on $(b_{ch}, \mathbf{x}_{ch})$, where recall that the covariates $\mathbf{x}$ also include beliefs as well as total expenditure, and where the index $h \in 1, ..., H_c$ denotes a household within cluster (village) c, and H_c is the number of households sampled within the cluster.

ASSUMPTION 1. *The unobserved term $\Delta\epsilon_{ch} q(\mathbf{x}_{ch}) = f_c + \eta_{ch}$ where f_c is unobserved village level heterogeneity and conditional on $(\mathbf{x}_{ch}, f_c)$ the household specific error terms $(\eta_{1c}, ..., \eta_{Hc})$ are i.i.d. with standard logistic distribution.*

Under Assumption 1, the model therefore assumes that unobserved heterogeneity consists of two additively separable components one of which (η) is independent of $\{\mathbf{x}, f\}$ with parametrically specified distribution. The other component (f) is allowed to be arbitrarily correlated with observables but is restricted to be constant across households within a cluster. Indexing by households and clusters, equation (3) is then given by

$$b_{ch}^* = \mathbb{I}\{g(u(\cdot), \mathbf{x}_{ch}) + f_c + \eta_{ch} \geq 0\} \tag{9}$$

and given the i.i.d. assumption, the probability of purchase is given by

$$\mathbb{P}\left(b_{ch} = 1 \left| \{\mathbf{x}_{ch}, f_c\}_{h=1}^{H_c}; u(.)\right.\right) = \frac{\exp(g(u(.), \mathbf{x}_{ch}) + f_c)}{1 + \exp(g(u(.), \mathbf{x}_{ch}) + f_c)}, \tag{10}$$

where we make explicit that the likelihood is a function of the choice-dependent utility. The non-standard aspects of the model are contained in the index function $g(\cdot)$. In the non-parametric case, the index function contains the $\Delta(\cdot)$ functions described in (4)—which are the objects of interest—whereas in the parametric case the function contains the finite dimensional parameters of interest $(\boldsymbol{\alpha}, \boldsymbol{\tau})$. The non-linear and non-standard nature of $g(\cdot)$ means that we must first develop identification results for the model given by (10) before we can proceed to estimation. As a first step, we write down the likelihood function. Using the standard conditioning argument to eliminate the fixed effect (see Chamberlain 1984), the likelihood of a particular sequence of net ownership decisions conditional on the total number of adoption decisions within cluster c is given by

$$\mathbb{P}\left(\mathbf{b}_c = \mathbf{b} \left| \sum_{h=1}^{H_c} b_{ch} = \bar{b}_c, \{\mathbf{x}_{ch}, f_c\}_{h=1}^{H_c}; u(\cdot)\right.\right) = \frac{\exp\left(\sum_{h=1}^{H_c} g(u(\cdot), \mathbf{x}_{ch}) b_{ch}\right)}{\sum_{a \in A(H_c, \bar{b}_c)} \exp\left(\sum_{h=1}^{H_c} g(u(\cdot), \mathbf{x}_{ch}) a_h\right)}$$

where $A(H_c, \bar{b}_c)$ is a set consisting of all sequences of binary variables $\{a_h\}_{h=1}^{H_c}$ such that $\sum_{h=1}^{H_c} a_h = \bar{b}_c$. The likelihood function is then

$$\Pi_{c=1}^C \frac{\exp\left(\sum_{h=1}^{H_c} g(u(\cdot), \mathbf{x}_{ch}) b_{ch}\right)}{\sum_{a \in A\left(H_c, \sum_{h=1}^{H_c} b_{ch}\right)} \exp\left(\sum_{h=1}^{H_c} g(u(\cdot), \mathbf{x}_{ch}) a_h\right)}, \tag{11}$$

where C denotes the number of clusters in the sample. Note that in order to carry out counter-factual exercises one needs to know (or otherwise estimate) f_c but these are not identified in the conditional logit model. Below, we suggest a bounding exercise to handle this problem.

6.2 Conditional Median Error Specification

The second estimation strategy begins by making far weaker assumptions on the error term $\Delta\epsilon$ and the function $q(\mathbf{x})$. Similar errors specifications have been used frequently following the introduction of the Maximum Score estimator in Manski (1975). This is important for at least four distinct reasons. First,

given the observational nature of the data, there is no compelling reason to believe that the observed and the unobserved determinants of net ownership decisions do not interact with each other. One way of capturing such interactions is via the existence of a non-constant $q(\cdot)$ function. Second, self-reported beliefs may be measured with error for various reasons and incorporating this into the model is important since they play an important role in identification. Third, the parameters in (7) might reasonably be understood as random coefficients. Finally, there may be no compelling reason to believe that the errors follow the parametric form specified above and at the very least we would like to examine the sensitivity of our conclusions to the error specification. In all of these cases, implementing an estimation procedure based on the conditional logit model in (11) would lead to inconsistent inference.

We therefore begin by describing the following assumption:

ASSUMPTION 2. *The error term $\Delta\epsilon_{ch}$ satisfies*

$$median\left(\Delta\epsilon_{ch} \mid \{\mathbf{x}_{ch}\}_{h=1}^{H}\right) = 0. \tag{12}$$

This assumption allows for a non-constant function $q(\mathbf{x})$ in (3) as well as allowing for other forms of dependencies between the unobserved heterogeneity and the observed covariates. In Appendix B we outline formal models for measurement error and random coefficients that can be placed within this framework. The weakness of the zero conditional median assumption, however, is that it is very difficult to recover all the parameters of interest in the index function. In particular, we will have to make scale normalizations that will imply that only part of the vector $\boldsymbol{\alpha}$ is identified. In addition, given the non-linear nature of the index function we will have to make stronger assumptions on the support of the consumption variable for identification.

We also consider the following weakening of the conditional logit specification.

ASSUMPTION 3. *The unobserved term in (3) can be decomposed as $\Delta\epsilon_{ch}q(\mathbf{x}_{ch}) = f_c + \eta_{ch}$, where f_c and η_{ch} represent unobserved heterogeneity at the village and household level respectively. Conditional on $(\{\mathbf{x}_{ch}\}_{h=1}^{C}, f_c)$, the random variable η_{ch} has the same distribution for all $h = 1, ..., H$ and has support over the entire real line.*

This assumption weakens the joint logit parametric assumption made previously but like the conditional logit allows for endogeneity. However, note that we do not place any restrictions on the joint distribution (beyond imposing that they have the same marginals) so that we also allow for general correlation in the error terms across households. This is unlike the conditional logit which imposes strong independence assumptions. Insofar as these correlations are an important issue with clustered data, it is important to account for them. The maximum score fixed effect model provides one method for doing so in addition to making minimal assumptions on the nature of the unobserved heterogeneity. Relative to Assumption 2 this assumption allows for the presence of a fixed effect but it does not allow for the direct interaction between household level observables and unobservables. All the caveats with respect to support conditions and normalizations hold here *a fortiori*. We discuss identification of the parametric utility model under Assumption 3 but do not do so for the non-parametric case since the proof is very similar to the case where Assumption 2 holds.

7 Identification with Parametric Errors

7.1 Identification with Non-Parametric Utility

If there were no observable covariates (i.e. no $(\mathbf{x}_a, \mathbf{x}_b, c)$ or more weakly if c were not in the model) in the utility functions, then the model would reduce to one that is linear in parameters and standard linear conditional logit identification conditions would suffice. However, once observable covariates enter the indirect utilities, stronger assumptions are required. To simplify exposition, we first focus on the case where there are only two households per cluster. We begin by assuming that beliefs—which are observable covariates in our data set—remain stochastic even after conditioning on other covariates. In what follows, for given vectors $\mathbf{A}$ and $\mathbf{B}$, $\mathbf{B} \subset \mathbf{A}$, we use "$\mathbf{A} \backslash \mathbf{B}$" to denote the elements of $\mathbf{A}$ that do not belong to $\mathbf{B}$. All proofs are included in Appendix A.

ASSUMPTION 4. *The distribution of δ_{ch} conditional on $(\mathbf{x}_{ch} \backslash \delta_{ch}, \mathbf{x}_{ch'})$ contains at least two points of support for almost every conditioning value.*

ASSUMPTION 5. *The distribution of π_{ch} conditional on $(\mathbf{x}_{ch} \backslash \pi_{ch}, \mathbf{x}_{ch'})$ contains at least two points of support for almost every conditioning value.*

In our context, assumptions 4 and 5 are intuitively sensible, because the perceived protective power of bednets likely depends on unobservable characteristics such as access to information about malaria or mosquito density. If one could perfectly predict beliefs using observables, there would be no identifying variation for the parameters of interest. In our data, OLS regressions of beliefs upon observables (not reported here) are encouraging in the sense that observables do not predict beliefs well with the R-squared of the regressions being very small. The following lemma shows that such variation in beliefs is required for identification regardless of the utility's functional form.

Lemma 1. *Consider the model given by (5) with the errors satisfying Assumption (1). Then, under Assumptions 4 and 5 the utility differential functions $\Delta_j(\cdot)$ for $j = 1, 2, 3$ are identified.*

The proof uses variation in beliefs across households to identify the model by utilizing differencing arguments. Note that although the differentials themselves are identified, the individual components of the differentials are not.

7.2 Identification With Parametric Utility

We now describe identification results assuming a parametric utility function as in (7). For simplicity we consider the case where there are 2 households in the cluster, but the argument holds for larger cluster sizes as well. Suppose that the model is not identified. Then, there exist parameters $(\tilde{\boldsymbol{\alpha}}, \tilde{\mathbf{t}})$ such that the likelihood is identical under the true values $(\boldsymbol{\alpha}, \mathbf{t})$ and under $(\tilde{\boldsymbol{\alpha}}, \tilde{\mathbf{t}})$ so that for any $\mathbf{b}$

$$\mathbb{P}\left(\mathbf{b}_c = \mathbf{b} \left| \sum_{h=1}^{2} b_{ch} = \bar{b}, \{\mathbf{x}_{ch}\}_{h=1}^{2} ; \boldsymbol{\alpha}, \mathbf{t} \right.\right) = \mathbb{P}\left(\mathbf{b}_c = \mathbf{b} \left| \sum_{h=1}^{2} b_{ch} = \bar{b}, \{\mathbf{x}_{ch}\}_{h=1}^{2} ; \tilde{\boldsymbol{\alpha}}, \tilde{\mathbf{t}} \right.\right).$$

We first note that the parameter α_2 is not identified since it is an intercept term that does not vary across households within a cluster. Next, identification of τ requires further restrictions on the $\boldsymbol{\alpha}$ parameters. This is formalized in the next assumption.

ASSUMPTION 6. *The parameters α_0 and α_1 satisfy $\alpha_0 + \alpha_1 \neq 0$.*

This assumption is not demanding, because its failure would imply that the marginal utility of consumption for individuals who own a net is zero. The necessity of this assumption follows from examining the form of the $g(.)$ function and noting that if $\alpha_0 + \alpha_1$ were equal to zero, then the function $h_1(.)$ which is used for identification would not enter the likelihood. More generally, if both α_0 and $\alpha_0 + \alpha_1$ were equal to zero, then $\boldsymbol{\tau}$ would not enter the likelihood at all and hence would not be identified. However, since the parameters can be interpreted as marginal utilities of consumption, it is reasonable to assume that they do not equal zero. In fact, in some of the estimation exercises we will set this sum equal to one so that this assumption will be automatically satisfied.

ASSUMPTION 7. *The consumption variable $c_{ch}(1, m)$ conditional on $(\mathbf{x}_{c1}, \mathbf{x}_{c2}) \backslash c_{ch}(1, m)$ is continuously distributed over some range.*

The continuity assumption is required to enable differentiation of the $h_1(\cdot)$ function with respect to consumption, an important step in showing identification.[22] Assumption 7 is relatively undemanding, because it implies, for instance, that consumption cannot be perfectly predicted using covariates for all households in the cluster and that the residual variation after conditioning is continuously distributed. It is the residual continuous variation in consumption that is central for identification. In our data, OLS regressions of consumption on household observables indicate that while some of the demographic variables (household size and education) are important predictors of consumption, the R-squared from these regressions is quite low, approximately .10. The next assumption is similar to a standard second-moment condition except that it involves the true value of the risk aversion parameters. Note, however, that since these parameters are identified prior to the use of this assumption in the proof, in principle this assumption is verifiable.

ASSUMPTION 8. *At the true value of $\boldsymbol{\tau}$, the random vector*

$$(\mathbf{x}_{b,c1} - \mathbf{x}_{b,c2}, \delta_{c1} - \delta_{c2}, \mathbf{x}_{m,c1} - \mathbf{x}_{m,c2}, \pi_{c1} - \delta_{c1} - \pi_{c2} + \delta_{c2}, h_1(\mathbf{x}_{c1}; \boldsymbol{\tau}) - h_1(\mathbf{x}_{c2}; \boldsymbol{\tau}), h_0(\mathbf{x}_{c1}; \boldsymbol{\tau}) - h_0(\mathbf{x}_{c2}; \boldsymbol{\tau}))$$

and the vector $\mathbf{x}_{a,ch}$ have a non-singular second moment matrix.

Lemma 2. *Suppose that the model is given by (7) with the errors satisfying Assumption (1) and that Assumptions 4-8 hold. Then, the parameters $\alpha_0, \alpha_1, \boldsymbol{\alpha}_3', \alpha_4, \boldsymbol{\alpha}_5', \alpha_6, \boldsymbol{\tau})$ in model (7) are identified.*

The proof relies on using the support conditions to respectively take differences of the likelihood function and then take second order derivatives. The second moment assumption is then finally used to ensure that $\boldsymbol{\tau}$ is identified. Conditional upon identification, the large sample properties of the estimator follow from standard results in the literature (for instance Theorem 3.3. in Newey and McFadden 1994) and we can conclude that the estimator converges at the parametric rate and has a limiting normal distribution with the limiting variance matrix equal to the negative of the inverse of the Fisher Information matrix.

---

[22]Showing identification with discrete support for consumption requires more work and is not pursued here.

7.2.1 Monte Carlo Results

To assess the practical validity of the identification and large sample results for the non-linear index conditional logit model, we present the results from a series of Monte Carlo experiments. The experiments are designed to illustrate that the parameters are well estimated in finite samples and that the parametric rate of convergence holds.

The model is given by

$$b_{ch}^* = \mathbb{I}\left\{g\left(\mathbf{x}_{ch}; \boldsymbol{\alpha}, \boldsymbol{\tau}\right) + f_c + \epsilon_{ch} \geq 0\right\}, \tag{13}$$

where $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \alpha_3, \alpha_4, \alpha_5, \alpha_6)$ and equation (7) takes the following form (suppressing the data subscripts)

$$g(\mathbf{x}; \boldsymbol{\alpha}, \boldsymbol{\tau}) \equiv 3x_b - \delta - x_m - (\pi - \delta)$$
$$+ h_1(\pi, \delta, c, x_a, \boldsymbol{\tau}) - h_0(\pi, c, x_a, \boldsymbol{\tau})$$
$$h_1(\pi, \delta, c, x_a; \boldsymbol{\tau}) \equiv (1 - \pi + \delta)c(1, h)^{1 - \gamma(\mathbf{x}_a, h; \boldsymbol{\tau})} + (\pi - \delta)c(1, m)^{1 - \gamma(\mathbf{x}_a, m; \boldsymbol{\tau})}$$
$$h_0(\pi, c, x_a; \boldsymbol{\tau}) \equiv (1 - \pi)c(0, h)^{1 - \gamma(\mathbf{x}_a, h; \boldsymbol{\tau})} + \pi c(0, m)^{1 - \gamma(\mathbf{x}_a, m; \boldsymbol{\tau})}.$$

Beliefs (π, δ) are each i.i.d. $\mathcal{B}eta(1, 1)$ random variables, the (x_b, x_m) are independent random variables with a discrete uniform distribution on $[0, 5]$ and a Bernoulli distribution with $p = .6$ respectively. State contingent consumption is given by

$$c(b, s) = c - .05c_b b - .25c_s s + f_c$$

where (c, c_b, c_s, f_c) are independent and log-normally distributed. The function

$$\gamma(x_{ch,a}, s_{ch}; \boldsymbol{\tau}) = \frac{e^{-0.5s_{ch}}}{1 + e^{-0.5s_{ch}}}$$

where $(\tau_0, \tau_1) = (-0.5, 0)$. The errors ϵ_{ch} have a logit distribution. Note that the consumption variable is correlated with the village fixed effect f_c so that a fixed-effect formulation is necessary. The parameter vector of interest is then $(\tau_0, \tau_1, \alpha_0, \alpha_1, \alpha_3, \alpha_4, \alpha_5, \alpha_6) = (-.5, 0, 1, 0, 3, 1, 1, -1)$. In order to carry out the estimation, we use a simplification noted in Arellano and Honore (2001) that constructs the objective function by carrying out pairwise comparisons of observations (within cluster) and is much faster to compute than the MLE. The resulting estimator is an M-estimator and its asymptotic properties therefore also follow from standard results (see e.g. Newey and McFadden 1994). The simulations are carried out for four sample sizes (150, 200, 300 and 600) and the results are reported in Table 4. The estimator is well behaved for sample sizes relevant for our empirical application with means and medians both close to the true values. Finally, the doubling of sample size from 150 to 300, or from 300 to 600, lead to a ratio of standard errors approximately equal to $\sqrt{n}$, which provides encouraging evidence that the parametric rate of convergence holds.

8 Identification with Median Restrictions

8.1 Identification with Non-Parametric Utility

Identification in the non-parametric utility case with median restrictions is quite weak in the sense that we are only able to recover one potential utility differential of interest and that too is sensitive to

normalization. However, for completeness we record the argument. Recall from (5) that the model is given by

$$b = \mathbb{I}\{\Delta_1 + \pi(\Delta_1 - \Delta_2) + \delta\Delta_3 + \Delta\epsilon \geq 0\},$$

where we have suppressed for simplicity the household-specific subscripts and the dependence of the differentials from the covariates $\mathbf{x}$. As is usual in models with median restrictions, location and scale normalizations are required since they determine the relationship between the identified parameters.[23] One possible location normalization is to set $\Delta_1 = 0$, which is equivalent to assuming that in the state of the world where the agent experiences malaria, no utility is derived from net ownership. One possible scale normalization is to set $\Delta_3 = 1$. This normalizes the marginal utility from avoiding malaria (when the household owns a net) to being positive and equal to 1. These are by no means the only normalizations one could choose of course but have the advantage of being simple to impose. With this normalizations, the only remaining unidentified parameter is the differential $(\Delta_1 - \Delta_2)$. Recalling the definitions of Δ_1 and Δ_2 from equation (4), this measures the difference between the change in utility due to owning a net when one has malaria and the change when one does not have malaria.

ASSUMPTION 9. *For every* $\mathbf{x}\backslash(\pi, \delta)$ *there exists an* $\xi > 0$ *and an open set* $\mathcal{N}$ *such that* $\mathbb{P}_{\mathbf{x}\backslash(\pi,\delta)}(\pi \in \mathcal{N}) > 0$ *and for all* $\pi \in \mathcal{N}$*, the distribution of* $\pi(\Delta_1 - \Delta_2) + \delta$ *(conditional upon* $\mathbf{x}\backslash(\pi, \delta)$*) has positive probability density over* $[-\xi, \xi]$ *(i.e. an interval containing zero).*

Lemma 3. *Consider the model given by (5) and suppose that Assumption 9 holds. Then the parameter* $\Delta_1 - \Delta_2$ *is identified.*

The result is straightforward, since effectively, conditional upon $\mathbf{x}\backslash(\pi, \delta)$ the model is a linear index model and standard identification conditions apply. However, these conditions (captured in Assumption 9) place strong restrictions on the support of δ and its ability to vary sufficiently freely conditional upon the remaining covariates in $\mathbf{x}$. The fact that both π and δ are by construction included in the unit interval makes the requirement somewhat less demanding, but the condition still requires, for example, that the parameters (Δ_1, Δ_2) are not such that $\pi(\Delta_1 - \Delta_2) + \delta$ never changes sign.[24] Given the definition of $(\Delta_1 - \Delta_2)$, there is no obvious reason why such assumption should hold and it it seems therefore worthwhile to relax the non-parametric utility specification. A parametric specification, although clearly relying on functional form assumptions, will allow us to exploit additional variation in other covariates to achieve identification and so require less stringent conditions.

8.2 Identification with Parametric Utility

We begin first by stating the following assumption:

ASSUMPTION 10. *At the true parameter value* $(\boldsymbol{\alpha}, \boldsymbol{\tau})$*, the index function* $g(\mathbf{x}; \boldsymbol{\alpha}, \boldsymbol{\tau})$ *conditional on* $s(\boldsymbol{\alpha}, \boldsymbol{\tau}, \mathbf{x}\backslash c(1, m))$ *has strictly positive density over an interval containing zero, where*

$$s\left(\boldsymbol{\alpha}, \boldsymbol{\tau}, \mathbf{x}\backslash c(0, m)\right)$$
$$= \left[\frac{1}{\pi}\left(\alpha_1 h_1\left(\mathbf{x}_a, c(1, h), c(1, m); \boldsymbol{\tau}\right) - (1 - \pi)c(0, h)^{1 - \gamma(\mathbf{x}_a, h; \boldsymbol{\tau})} + \boldsymbol{\alpha}_3 \mathbf{x}_b\right)\right]^{\frac{1}{1 - \gamma(\mathbf{x}_a, m; \boldsymbol{\tau})}} \tag{14}$$

---

[23]See Horowitz (1998) for a clear explanation on the need for such normalizations.

[24]Note, however, that we do not require that this support be all of the real line.

Like Assumption 9, this is a very strong assumption, because it imposes boundary restrictions which are not obviously satisfied. Some version of this kind of assumption is required for point identification in the maximum score (see for instance Horowitz 1998 for a textbook discussion on identification issues for the maximum score in the linear index case). As the proof will show, this is a key assumption in showing identification. Next, we need a stronger version of Assumption 7 for the current model.

ASSUMPTION 11. *There exists a variable in $\mathbf{x}\backslash c(1, m)$ that is continuously distributed over some interval conditional on all the other variables in $\mathbf{x}\backslash c(1, m)$.*

Similarly, we need to replace Assumption 8 with the following:

ASSUMPTION 12. *At the true value of $(\boldsymbol{\alpha}, \boldsymbol{\tau})$, the random vector $(\mathbf{x}_b, h_1(\mathbf{x}; \boldsymbol{\tau}))$ has a non-singular second moment matrix. The vector $\mathbf{x}_a$ has a non-singular second moment matrix.*

Assumption 12 performs a role analogous to Assumption 8 in Lemma 2. In particular, it is used to argue that if $\boldsymbol{\tau}$ is identified, then the coefficients in $\boldsymbol{\alpha}$ are also identified. Finally, we require a scale normalization, which is necessary given the weak assumptions on the error terms.

ASSUMPTION 13. *The parameter vector $\alpha_0 = 1$.*

Setting $\alpha_0 = 1$ normalizes the utility level in the "no net ownership" state of the world relative to which utility comparisons will be made (alternative normalizations are discussed in the appendix). We can now state the point-identification result for the maximum score estimator

Lemma 4. *Consider the model given by (3) and Assumption (2). Then under Assumptions 10-13 the parameters (α_1, α_3) and $\boldsymbol{\tau}$ are identified.*

The model in this section enables more flexibility than the parametric error model described in Section 7 in at least two distinct ways. First, unobserved heterogeneity is allowed to interact with observables in determining net purchase. This is important because there seems to be no *a priori* reason to rule these out and indeed such interactions are likely important given that we only observe a subset of household characteristics. Second, the model makes no parametric assumption on the nature of the unobserved heterogeneity which is also important given that we know very little about these unobserved household characteristics. However, the added flexibility comes at a cost. First, not all the parameters of interest are identified. Second, the estimates of the identified parameters converge at a rate slower than the usual parametric rate and the limiting distribution is non-standard, making inference non-trivial. Finally, the assumptions for point identification are quite strong and may not hold in our particular data set.

8.2.1 Maximum Score Fixed Effect Identification

We next discuss identification of the parametric model under Assumption 3. This assumption allows for household observables to be correlated with unobserved village level variables and also allows error terms across households to be correlated with each other. We state results assuming two households per cluster but the reasoning extends straightforwardly to clusters with more than two households. In what follows it is convenient to define $\mathbf{w}_c = (\mathbf{x}_{c1}, \mathbf{x}_{c2})$ and by $c_{ch}(b, m)$ we denote the consumption level

for household h, given bednet purchase decision b and malaria status m. We begin with an appropriate strengthening of Assumption 10.

ASSUMPTION 14. *At the true parameter value* $(\boldsymbol{\alpha}, \boldsymbol{\tau})$*, the function* $g(\mathbf{x}_{c1}; \boldsymbol{\alpha}, \boldsymbol{\tau}) - g(\mathbf{x_{c2}}; \boldsymbol{\alpha}, \boldsymbol{\tau})$ *conditional on* $\tilde{s}(\boldsymbol{\alpha}, \boldsymbol{\tau}, \mathbf{w}_c \backslash c_{c2}(0, m))$ *has strictly positive density over an interval containing zero, where*

$$\tilde{s}(\boldsymbol{\alpha}, \boldsymbol{\tau}, \mathbf{w_c} \backslash c_{c2}(0, m)) \equiv \left[-\frac{1}{\pi_2} \left(\Delta\bar{\mathbf{x}}' \bar{\boldsymbol{\alpha}} + \alpha_1 \Delta h_1(\boldsymbol{\tau}) - h_{1,0}(\boldsymbol{\tau}) + (1 - \pi_2) c_2(0, h)^{1-\gamma(\mathbf{x}_{a,c2}, m; \boldsymbol{\tau})} \right) \right]^{\frac{1}{1-\gamma(\mathbf{x}_{a,c2}, m; \boldsymbol{\tau})}}$$

The quantities $(\Delta\bar{\mathbf{x}}, \bar{\boldsymbol{\alpha}}, \Delta h_1(\cdot), h_{1,0}(\cdot), \gamma_{sh}(\cdot))$ are defined in Appendix A. This assumption, like the analogous ones described in the previous subsection, imposes a strong support condition but is key in showing point identification. Next, we need a stronger version of Assumption 11.

ASSUMPTION 15. *There exist two variables in* $\mathbf{w}_c \backslash c_{c2}(0, m)$ *which are continuously distributed over some open interval conditional on all the other variables.*

In the model, the variables $c_{c2}(0, h)$ and $c_{c1}(0, m)$ are presumed to satisfy this requirement. Similarly, we need to replace Assumption 12 with the following:

ASSUMPTION 16. *At the true vector* $(\boldsymbol{\alpha}, \boldsymbol{\tau})$ *the random vector* $(\Delta\bar{\mathbf{x}}, \Delta h_1(\boldsymbol{\tau}))$ *has a non-singular second moment matrix. The vector* $\mathbf{x}_a$ *has a non-singular second moment matrix.*

Assumption 16 performs a role analogous to Assumption 8 in Lemma 2. In particular, it is used to argue that $\boldsymbol{\tau}$ is identified and, as a consequence, that certain coefficients in $\boldsymbol{\alpha}$ are also identified. Finally, we require a scale normalization, which is necessary given the weak assumptions on the error terms. We can now state the point-identification result for the maximum score fixed effects estimator with non-linear index function.

Lemma 5. *Consider the model given by Assumption (3) and equation (7). Then under Assumptions 13-16 the parameters* $(\alpha_1, \boldsymbol{\alpha}_3', \alpha_4, \boldsymbol{\alpha}_5', \alpha_6, \boldsymbol{\tau})$ *are identified.*

Note that, relative to the result of Lemma 2, here we lose the identification of α_0, that is, the constant that multiplies $c(.)^{1-\gamma(.)}$ in the utility function for agents who did not purchase a bednet.

8.2.2 Monte Carlo Results

We carry out a small set of Monte Carlo simulations to assess the practical performance of the maximum score estimator in small samples. We use genetic algorithms to maximize the objective function. The model is the same as was estimated in Section 7.2.1 with the exception that α_0 is normalized to be 1. The results in Table 5 illustrate that the parameters are estimable in small samples, although the performance seems worse than in the parametric case. Finally, looking at standard deviations across sample sizes indicates that the standard parametric rate of convergence does not hold.

9 Estimation

9.1 Parametric Errors and Preferences

We first discuss estimation of the model given by the parametric utility framework in (7). In this specification, the risk aversion coefficient is a function of $\mathbf{x}_a$ via the index function (8). We assume

that $\mathbf{x}_a$ includes household size, age and education of the household head and the number of children under the age of five (U5). In addition, we assume that $\mathbf{x}_b$ includes a binary variable equal to one if the respondent thinks that more than 50% of households in the village sleep regularly under a bednet. This interaction between net ownership and household perceptions of community level net ownership captures (in an admittedly ad hoc manner) the potential interdependence between household utility and the perceived behavior of other households.[25] The presence of this variable is motivated both by field observations as well as results from reduced form regressions that suggested that net ownership is affected by the perceptions of other households' behavior. Since we use a fixed effect assumption on the error term, we directly control for actual net usage levels in the village. As a result, any measured effect of the perceptions variable will be net of actual usage levels. Finally, we assume that $\mathbf{x}_m$ includes household size and the number of U5, so that we allow the disutility from malaria to vary by household size as well as by the fraction of young children in the household. This allows us to capture further potential heterogeneity in the impact of malaria across households (beyond that captured in the risk aversion parameter). Finally, the parameter α_6 allows for the interaction between malaria status and net ownership to affect utility.

The model is estimated using the (conditional) maximum likelihood (11) using gradient-based methods and the results are displayed in Table 6. We first estimate the parameters $\boldsymbol{\tau}$ that characterize the coefficient of relative risk aversion. All the variables included in $\mathbf{x}_a$ are positively associated with the level of risk aversion. In particular, households with more young children are more risk averse than households with fewer children. This suggests that the reduced form positive correlation between net ownership and young children arises partly because households with young children are more risk averse. In addition, older and more educated household heads are more risk averse which again provides an economic rationale for the positive correlation we observe in reduced form between these variables and net ownership. The point estimate for malaria status is positive, implying that utility is state dependent and households with malarious individuals exhibit more risk aversion towards consumption lotteries in the malaria state of the world (so that marginal utility of consumption is higher). However, the estimate is not significant at conventional levels. The coefficient α_4 (which multiplies $1(s = m)$, is negative, although to evaluate the impact of malaria on utility it would be necessary to take into account that $1(s = m)$ enters the index in a complex fashion. In addition, there is evidence that malaria has heterogeneous impact across households with the elements of α_5 negative and jointly statistically significant. The results suggest that malaria decreases utility more in larger households as well as households with more young children. This is consistent with the findings about households risk aversion. We turn next to the coefficient α_3 that captures the heterogeneity across households in the utility derived from net ownership. We see that perceptions of village level ownership are important in that the direct utility derived from net ownership is significantly higher for households believing that more than half of the village owns a net. Since we control for village level fixed effects in the analysis, it is not actual ownership that is directly driving these results but rather the perception that most other households own a net. This result suggests that perceptions of group-level behavior directly affect utility and thus

---

[25]We do not, however, model the implied simultaneity across agent choices. Doing so would greatly complicate the model and require an assumption of rational expectations to find solutions.

influence individual decision-making.[26] As discussed earlier, a fully general model that incorporates the strategic inter-dependence of all agent purchase decisions would be intractable to estimate and so we leave that inter-dependence un-modeled. Also, since we directly account for malaria related beliefs in the analysis, we already allow for any interdependence across household decisions arising from common information.

9.2 Parametric Preferences and Median Restrictions on Errors

Section 6.2 outlined four distinct reasons why a median restriction on the error term could be a useful an improvement on the previous model. In addition, it would be of interest to estimate the maximum score model as a robustness check on the stability of our counterfactual exercises. To this end, we re-estimate the model with parametric utility under the set of assumptions required for Lemma 5. We need to make standard location and scale normalizations and we achieve them by setting $\alpha_1 = 0$ and $\alpha_0 = 1$. The coefficients across the two specifications are then not directly comparable because of these normalizations. However, it is possible to compare the signs of the coefficients as well as carry out counterfactual exercises using the maximum score coefficients subject to the normalizations. Because the objective function for the maximum score is not everywhere differentiable, standard gradient-based optimization algorithms cannot be used and estimation was carried out using genetic algorithms.[27] Abrevaya and Huang (2005) show that the bootstrap is inconsistent for the maximum score, so we used subsampling procedures to generate confidence intervals. The results for the point estimates are displayed in Table 8. With the exception of $\hat{\alpha}_{51}$ (which turns from negative and significant at the 10% level to positive and not significant) the signs of the estimated parameters that enter the index linearly ($\hat{\boldsymbol{\alpha}}$) are the same for both estimation methods. The parameters that enter the index non-linearly ($\hat{\boldsymbol{\tau}}$) do not agree in sign with the parametric case in three instances but, in each of these cases, they are very imprecisely estimated and we cannot reject the null that they are equal to zero at conventional levels.

10 Counterfactual Exercises

We next propose several manipulations of the model's exogenous variables and study their effect on net uptake. There is, however, an important methodological impediment to this analysis. In conventional binary choice models, one would carry out counterfactuals by examining changes in choice probabilities (see, for instance, Delavande 2008). However, in a fixed effects model, the choice probabilities $\mathbb{P}\left(b_{ch} = 1 | \mathbf{x}_{ch}, f_c\right)$ are not identified. Further assumptions, typically on the conditional distribution $f_c | \mathbf{x}_{ch}$, are required to identify them. Such assumptions are not attractive since they inevitably place considerable structure on the fixed effect, an object on which one would like to place as little structure as possible. As an alternative to such ad-hoc assumptions, we propose a method for evaluating

---

[26]Unlike standard models of social interactions, utility here does not depend upon a linear functional of the (perceived) group level distribution of usage. The variable used is equal to one whenever the household thinks that more that half of the village owns nets. This is a non-linear (though still smooth) functional of the distribution of household beliefs about village net usage. The information was recorded in this fashion because fieldwork suggested that households were more comfortable with this wording rather than the standard question eliciting probabilities directly. The usual fixed-point solution concepts in standard models are not directly applicable here since they rely heavily on the linearity assumption.

[27]We also experimented with simulated annealing algorithms.

policy changes based on a bounding analysis. Specifically, let $\{g\left(\mathbf{x}_{ch}; \hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\tau}}\right)\}_{h,c}$ denote the estimate of the observable component of the index functions (9) for all households in the sample (evaluated at the household specific value $\mathbf{x}_{ch}$. Consider a manipulation of $\mathbf{x}_{ch}$ to $\mathbf{x}'_{ch}$ and denote the associated index functions by $\{g\left(\mathbf{x}'_{ch}; \hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\tau}}\right)\}_{h,c}$. Next, identify all households who do not purchase a net and for whom

$$g\left(\mathbf{x}_{ch}; \hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\tau}}\right) < 0 < g\left(\mathbf{x}'_{ch}; \hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\tau}}\right). \tag{15}$$

In words, these are households that, in the absence of the unobserved component in (9), would change their bednet ownership decision from non-purchase to purchase when the covariates change from $\mathbf{x}_{ch}$ to $\mathbf{x}'_{ch}$. These households provide therefore an upper bound on the possible effect of the policy intervention net uptake, because, in reality, the unobserved component of utility will likely lead only some of them to modify their purchase decision. If a household were to purchase a net under the new policy regime, then it must be the case that the unobserved component of the index, that is, $f_c + \eta_{ch}$, satisfies the following condition:

$$f_c + \eta_{ch} \in \left(-g(\mathbf{x}'_{ch}; \hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\tau}}), -g(\mathbf{x}_{ch}; \hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\tau}})\right). \tag{16}$$

Counting all households satisfying the inequality (15) as "switchers" makes the implicit assumption that for all of them (16) holds which in general need not be true. This exercise then provides us with an upper bound on the possible effects of the policy intervention.

10.1 Evaluating the effects of Price Changes

A major issue in the public policy debate on net provision is net prices. Several global institutions have called for largely subsidized or free net provision to poor households (see the introduction). In our sample, the mean price paid for a net on the open market was about 80 Rupees.[28] In the first set of counterfactual exercises we study the effect of a 50% reduction in net prices on uptake. We model the price reduction as the increase in non-bednet expenditure $c(b, s)$ allowed by the price reduction, leaving total outlay unchanged. We present the results in Table 7 and show that net uptake is minimally effected by a price change. A 50% decline in the price paid for a net leads to a maximum of four households in the sample switching their purchase decision from 0 to 1. The results suggest that, in this sample and conditional on the model being correctly specified, price changes will have little effect on uptake.

10.2 Evaluating the effects of Changes in Beliefs

The other major component of malaria eradication programs has been raising awareness of malaria and of the benefits of using bednets. We next evaluate the effect of interventions that exogenously alter beliefs both about the efficacy of bednets and of community bednet use. We first evaluate the effect of an intervention that exogenously increases households' perceptions on the efficacy of bednets by 50%.[29] This means that on average a household believes that the regular use of a net will reduce the likelihood

---

[28]Net costs are a relatively low fraction of expenditures (total monthly household expenditures are about 3600 Rupees and the corresponding per capita figure is about 700 Rupees).

[29]Specifically, this is done by evaluating the impact of increasing the perceived protective power from δ_{ch} to $\min\{1.5 \times \delta_{ch}, 1\}$.

of malaria over a malaria season by 80%. In this case, our estimates indicate that at most 23 households (only about 1%) would switch their purchase decisions based on these revised beliefs.

Our empirical analysis showed that household beliefs regarding village level net ownership influenced household utility from net purchase. We therefore evaluate the effect of a change in such perceptions such that 90% of respondents believe that more than half the village owns nets. This change leads five percent of households to alter their purchase decisions.

In addition to considering these counterfactuals, we also considered combinations of these counter-factuals. In particular, we considered a simultaneous increase in beliefs about the protective efficacy of nets as well as an increase in beliefs about their community level usage. These combinations provide the strongest results with about 10% of all households (at the upper bound) changing their decisions as a result of these changes in their environment. These results point to the potential importance of considering multiple interventions simultaneously to improved ITN coverage and usage.

In sum, the results from the counterfactual exercises suggest that price reduction and increased perceived protection will have limited impact on net uptake rates although there is some suggestive evidence that changes along multiple margins may have better results. These conclusions are not unexpected given the general failure of such programs to improve net take up in practice. There is some evidence of social interaction effects so that there may be a strong feedback effect between individual and village level net ownership. These results, however, depend upon model specification as well as the strong assumption that the interventions will not directly alter the structure of preferences (that is, the parameters $(\boldsymbol{\alpha}, \boldsymbol{\tau})$) themselves. Also, our model assumes that beliefs about bednet efficacy and community bednet use are exogenously given, but belief formation is more likely the result of cognitive processes which should ideally be modeled explicitly.

The results for the counterfactual exercises using only median restriction assumptions on the error term are presented in Table 9. The results from the counterfactual exercises are also broadly in line with those from the parametric case, indicating that uptake responses to changes in prices and beliefs are modest at best. Using experimental data from Kenya, Cohen and Dupas (2010) document large drops in uptake when cost-sharing is introduced. Our analysis, however, is not suited to estimate the price elasticity at a price of zero since the model assumes strictly positive prices.

11   Conclusions

This paper develops a simple static model of net purchase decisions using specifically collected data on subjective beliefs about the protective power of bednets against malaria. We study conditions for iden-tification of key preference parameters under alternative parametric and non-parametric assumptions about the utility function and about the structure of the error. We find that non-parametric utility specifications are generally not point identified. However, providing a parametric structure for utility allows to recover the preference structure under an appropriate set of conditions. These identification results are novel and can be extended to more general discrete choice settings with non-linear index function specifications. Such results are potentially important because they show identification for pref-erence parameters (such as covariate-dependent risk aversion parameters) that enter the index function non-linearly while at the same time placing relatively few restrictions on the unobserved heterogeneity

in household preferences. Concretely, we show identification results for household varying constant relative risk aversion (CRRA) preferences in the presence of cluster fixed effects while imposing only a conditional median restriction.

We next estimate the model under alternative sets of assumptions using data from rural Orissa and carry out a set of counterfactual exercises to evaluate the effects of possible changes in some features of the underlying economic environment. We find that allowing for the estimation of covariate-dependent heterogeneity in risk aversion is important, and that risk aversion changes with education levels and demographic characteristics of the household. Our results also indicate that perceptions about village-level net ownership is a relatively important predictor of adoption, while counterfactual price reductions and increased perceived protection of bednets are estimated to have very limited impact on net uptake rates. This latter finding is consistent with the often documented failure of public health programs that attempt to increase bednet use through subsidies or social marketing.

A natural next step would be to address some of the current limitations of the model. First, the model assumes that beliefs are exogenously given, but belief formation is more likely the result of cognitive processes which should ideally be modeled explicitly and we hope to do this using follow up data from the same households. A second related limitation is that we take perceptions about village-level bednet use as exogenous, while such perceptions may be strategically interdependent.[30] Third, our estimation strategy relies on specifically collected data on the perceived protective power of bednets. Such information, albeit relevant, may not sufficiently represent the complexity of the purchase decision faced by a poor multi-person household. Finally, it would be of interest to estimate the model using experimental data to validate the choice of functional form as well as other model specification issues.

---

[30]Although note that since in our model it is a non-linear functional of the distribution of (perceived) behavior that enters utility, standard social interaction models are not directly applicable.

References

Abdulla, S., J. A. Schellenberg, R. Nathan, O. Mukasa, T. Marchant, T. Smith, M. Tanner, and C. Lengeler (2001). Impact on malaria morbidity of a programme supplying insecticide treated nets in children aged under 2 years in Tanzania: community cross sectional study. *British Medical Journal 322*(7281), 270–3.

Abrevaya, J. and J. Huang (2005). On the bootstrap of the maximum score estimator. *Econometrica 73*(4).

Adongo, P., B. Kirkwood, and C. Kendall (2005). How local community knowledge about malaria affects insecticide-treated net use in northern Ghana. *Tropical Medicine & International Health 10*(4), 366–378.

Agyepong, I. (1992). Malaria: ethnomedical perceptions and practice in an Adangbe farming community and implications for control. *Social Science and Medicine 35*(2), 131–7.

Agyepong, I. and L. Manderson (1999). Mosquito avoidance and bed net use in the Greater Accra region, Ghana. *Journal of Biosocial Science 31*(01), 79–92.

Ahorlu, C., S. Dunyo, E. Afari, K. Koram, and F. Nkrumah (1997). Malaria-related beliefs and behaviour in southern Ghana: implications for treatment, prevention and control. *Tropical Medicine & International Health 2*(5), 488–499.

Alaii, J. A., W. A. Hawley, M. S. Kolczak, F. O. ter Kuile, J. E. Gimnig, J. M. Vulule, A. Odhacha, A. J. Oloo, B. L. Nahlen, and P. A. Phillips-Howard (2003). Factors affecting use of permethrin-treated bed nets during a randomized controlled trial in western Kenya. *The American Journal of Tropical Medicine and Hygiene 68*(4 Suppl), 137–141.

Alles, H., K. Mendis, and R. Carter (1998). Malaria mortality rates in South Asia and in Africa: Implications for malaria control. *Parasitology Today 14*(9), 369–375.

Arellano, M. and B. Honore (2001). Panel data models: Some recent developments. In J. Heckman and E. Leamer (Eds.), *Handbook of Econometrics*, Volume 5, Chapter 53, pp. 3229–3296. North-Holland.

Attanasio, O. and K. Kaufmann (2009). Educational Choices, Subjective Expectations, and Credit Constraints.

Beach, R. F., T. K. Ruebush, J. D. Sexton, P. L. Bright, A. W. Hightower, J. G. Breman, D. L. Mount, and A. J. Oloo (1993). Effectiveness of permethrin-impregnated bed nets and curtains for malaria control in a holoendemic area of western Kenya. *The American Journal of Tropical Medicine and Hygiene 49*(3), 290–300.

Binswanger, H. P. (1980). Attitudes toward risk: Experimental measurement in rural India. *American Journal of Agricultural Economics 62*(3), 395–407.

Bleakley, H. (2010). Malaria eradication in the Americas: A retrospective analysis of childhood exposure. *American Economic Journal: Applied Economics 2*(2), 1–45.

Chamberlain, G. (1984). Panel Data. In Z. Griliches and M. Intriligator (Eds.), *Handbook of Econometrics*, Volume II. North-Holland.

Chamberlain, G. (1993). Feedback in panel data models. Unpublished Manuscript, Department of Economics, Harvard University.

Chen, X. (2004). Large sample sieve estimation of semi-nonparametric models. In *Handbook of Econometrics*, Volume 6 of *Handbook of Econometrics*.

Cohen, J. and P. Dupas (2010). Free distribution or cost-sharing? Evidence from a randomized malaria prevention experiment. *Quarterly Journal of Economics 125*(1), 1–45.

Cutler, D., W. Fung, M. Kremer, M. Singhal, and T. Vogl (2010). Early-life malaria exposure and adult outcomes: Evidence from malaria eradication in India. *American Economic Journal: Applied Economics 2*(2), 72–94.

De Paula, A., G. Shapira, and P. Todd (2010). How beliefs about hiv status affect risky behaviors: Evidence from malawi, fifth version. *PIER Working Paper Archive*.

Delavande, A. (2008). Pill, patch or shot? Subjective expectations and birth control choice. *International Economic Review 49*(3), 999–1042.

Delavande, A., X. Giné, and D. McKenzie (2010). Measuring subjective expectations in developing countries: A critical review and new evidence. *Journal of Development Economics 94*(2), 151–163.

Dominitz, J. (1998). Earnings expectations, revisions, and realizations. *Review of Economics and Statistics 80*(3), 374–388.

Duflo, E., M. Kremer, and J. Robinson (2009). Nudging farmers to use fertilizer: Theory and experimental evidence from Kenya. NBER Working Paper No. 15131.

Dutta, H. S. (2000). KAP study on malaria in project areas of malaria centres of OVHA, 1999-2000. Technical report, Orissa Voluntary Health Association, Bhubaneswar, Orissa, India. Available at `http://www.orissavha.org/studies/pdf/malaria_02.pdf`.

Erdem, T., S. Imai, and M. Keane (2003). Brand and Quantity Choice Dynamics Under Price Uncertainty. *Quantitative Marketing and Economics 1*(1), 5–64.

Erdem, T. E., M. P. Keane, T. S. Öncü, and J. Strebel (2005). Learning about computers: An analysis of information search and technology choice. *Quantitative Marketing and Economics 3*, 207–247.

Giné, X., R. M. Townsend, and J. Vickery (2007). Rational expectations? Evidence from planting decisions in semi-arid India. Working Paper, The World Bank (DECRG), University of Chicago and Federal Reserve Bank of New York FED.

Gonul, F. and K. Srinivasan (1996). Estimating the Impact of Consumer Expectations of Coupons on Purchase Behavior: A Dynamic Structural Model. *Marketing Science 15*(3), 262–279.

Government of India (2007). Poverty estimates for 2004-05. Available at `http://planningcommission.nic.in/news/prmar07.pdf` (accessed May 2008).

Government of Orissa (2004). *Human Development Report 2004: Orissa.* Government of Orissa.

Guyatt, H. L., M. H. Gotink, S. A. Ochola, and R. W. Snow (2002). Free bednets to pregnant women through antenatal clinics in kenya: a cheap, simple and equitable approach to delivery. *Tropical Medicine & International Health 7*(5), 409–20.

Guyatt, H. L., S. A. Ochola, and R. W. Snow (2002). Too poor to pay: charging for insecticide-treated bednets in highland Kenya. *Tropical medicine & international health 7*(10), 846–50.

Hawley, W. A., P. A. Phillips-Howard, F. O. ter Kuile, D. J. Terlouw, J. M. Vulule, M. Ombok, B. L. Nahlen, J. E. Gimnig, S. K. Kariuki, M. S. Kolczak, and A. W. Hightower (2003). Community-wide effects of permethrin-treated bed nets on child mortality and malaria morbidity in western Kenya. *The American Journal of Tropical Medicine and Hygiene 68*(4 Suppl), 121–7.

Hawley, W. A., F. O. ter Kuile, R. S. Steketee, B. L. Nahlen, D. J. Terlouw, J. E. Gimnig, Y. P. Shi, J. M. Vulule, J. A. Alaii, A. W. Hightower, M. S. Kolczak, S. K. Kariuki, and P. A. Phillips-Howard (2003). Implications of the western Kenya permethrin-treated bed net study for policy, program implementation, and future research. *The American Journal of Tropical Medicine and Hygiene 68*(4 Suppl), 168–73.

Hendel, I. and A. Nevo (2006). Measuring the Implications of Sales and Consumer Stockpiling Behavior. *Econometrica 74*(6), 1637–1673.

Hill, Z., C. Kendall, P. Arthur, B. Kirkwood, and E. Adjei (2003). Recognizing childhood illnesses and their traditional explanations: exploring options for care-seeking interventions in the context of the IMCI strategy in rural Ghana. *Tropical Medicine & International Health 8*(7), 668–676.

Hoffmann, V. (2008). Psychology, gender, and the intrahousehold allocation of free and purchased mosquito nets. Cornell University mimeo.

Horowitz, J. (1998). *Semiparametric Methods in Econometrics.* Springer-Verlag.

Hurd, M. and K. McGarry (1995). Evaluation of the subjective probabilities of survival in the Health and Retirement Study. *Journal of Human Resources 30*(Supplement), 268–292.

Hurd, M. D., J. P. Smith, and J. M. Zissimopoulos (2004). The effects of subjective survival on retirement and social security claiming. *Journal of Applied Econometrics 19*, 761–775.

Imbens, G. and D. Hyslop (2000). Bias from Classical and other forms of Measurement Error. Technical Report 257, National Bureau of Economic Research.

Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *Quarterly Journal of Economics 125*(2), 515–548.

Kaufmann, K. (2010). Understanding the income gradient in college attendance in mexico: The role of heterogeneity in expected returns. *Working Papers.*

Kim, J. and D. Pollard (1990). Cube Root Asymptotics. *Annals of Statistics 18*(1), 191–219.

Klein, R., S. Weller, R. Zeissig, F. Richards, and T. Ruebush (1995). Knowledge, beliefs, and practices in relation to malaria transmission and vector control in Guatemala. *The American Journal of Tropical Medicine and Hygiene 52*(5), 383.

Korenromp, E. (2005). Malaria incidence estimates at country level for the year 2004. Proposed estimates and draft report. Technical report, Roll Back Malaria, World Health Organization, Geneva, Switzerland.

Kumar, A., N. Valecha, T. Jain, and A. P. Dash (2007). Burden of malaria in India: Retrospective and prospective view. *The American Journal of Tropical Medicine and Hygiene 77*(Suppl 6), 69–78.

Leenstra, T., P. A. Phillips-Howard, S. K. Kariuki, W. A. Hawley, J. A. Alaii, D. H. Rosen, A. J. Oloo, B. L. Nahlen, P. A. Kager, and F. O. ter Kuile (2003). Permethrin-treated bed nets in the prevention of malaria and anemia in adolescent schoolgirls in western Kenya. *The American Journal of Tropical Medicine and Hygiene 68*(4 Suppl), 86–93.

Lengeler, C. (2004). Insecticide-treated bed nets and curtains for preventing malaria. *Cochrane database of systematic reviews (Online).* Issue 2. Art. No.: CD000363. DOI: 10.1002/14651858.CD000363.pub2.

Lochner, L. (2007). Individual perceptions of the criminal justice system. *American Economic Review 97*(1), 444–460.

Louis, J., G. Le Goff, A. Trebucq, R. Migliani, F. Louis, V. Robert, and P. Carnevale (1992). Feasibility of mosquito control using insecticide-impregnated bednets in rural areas of Cameroon. *Ann Soc Belg Med Trop 72*(3), 189–95.

Luseno, W. K., J. G. McPeak, C. B. Barrett, P. D. Little, and G. Gebru (2003). Assessing the value of climate forecast information for pastoralists: Evidence from Southern Ethiopia and Northern Kenya. *World Development 31*(9), 1477–1494.

Lybbert, T. J., C. B. Barrett, J. G. McPeak, and W. K. Luseno (2007). Bayesian herders: Updating of rainfall beliefs in response to external forecasts. *World Development 35*(3), 480–497.

Mahajan, A. and A. Tarozzi (2011). Time inconsistency, expectations and technology adoption: The case of insecticide treated nets. Working Paper, Duke University and Stanford University.

Manski, C. F. (1975). Maximum score estimation of the stochastic utility model of choice. *Journal of Econometrics 3*, 205–228.

Manski, C. F. (1987). Semiparametric analysis of random effects linear models from binary panel data. *Econometrica 55*(2), 357–362.

Manski, C. F. (2004). Measuring expectations. *Econometrica 72*(5), 1329–1375.

Manski, C. F. and F. Molinari (2010, April). Rounding Probabilistic Expectations in Surveys. *Journal of business & economic statistics : a publication of the American Statistical Association 28*(2), 219–231.

Matzkin, R. (1991). Semiparametric Estimation of Monotone and Concave Utility Functions for Polychotomous Choice Models. *Econometrica: Journal of the Econometric Society*, 1315–1327.

Matzkin, R. (1992). Nonparametric and Distribution-free Estimation of the Binary Threshold Crossing and the Binary choice models. *Econometrica: Journal of the Econometric Society 60*(2), 239–270.

McKenzie, D., J. Gibson, and S. Stillman (2009). The impacts of international migration on remaining household members: Omnibus results from a migration lottery program. Manuscript.

Monasch, R., A. Reinisch, R. W. Steketee, E. L. Korenromp, D. Alnwick, and Y. Bergevin (2004). Child coverage with mosquito nets and malaria treatment from population-based surveys in African countries: a baseline for monitoring progress in Roll Back Malaria. *The American journal of tropical medicine and hygiene 71*(2 Suppl), 232–8.

Moody, A. (2002). Rapid diagnostic tests for malaria parasites. *Clinical Microbiology Reviews 15*(1), 66–78.

National Vector Borne Disease Control Programme (2008). Malaria situation. Available at http://www.nvbdcp.gov.in/malaria-new.html, Accessed June 20, 2008. National Vector Borne Disease Control Programme, Directorate General of Health Services, Ministry of Health and Family Welfare. Delhi, India.

Newey, W. K. and D. McFadden (1994). Large sample estimation and hypothesis testing. In R. Engle and D. McFadden (Eds.), *Handbook of Econometrics*, Volume IV, Chapter 36, pp. 2111–2245. Elsevier Science.

Nyarko, Y. and A. Schotter (2002). An experimental study of belief learning using elicited beliefs. *Econometrica 70*(3), 971–1005.

Onwujekwe, O., K. Hanson, and J. Fox-Rushby (2004). Inequalities in purchase of mosquito nets and willingness to pay for insecticide-treated nets in Nigeria: challenges for malaria control interventions. *Malaria Journal 3*(6).

Onwujekwe, O., E. Shu, R. Chima, A. Onyido, and P. Okonkwo (2000). Willingness to pay for the retreatment of mosquito nets with insecticide in four communities of south-eastern Nigeria. *Tropical Medicine & International Health 5*(5), 370–376.

Pierre-Andre Chiappori, Amit Gandhi, B. S. and F. Salanie (2009). Identifying Preferences under Risk from Discrete Choices. *American Economic Review forthcoming*. Papers and Proceedings.

Sahu, S. S., P. Jambulingam, T. Vijayakumar, S. Subramanian, and M. Kalyanasundaram (2003). Impact of alphacypermethrin treated bed nets on malaria in villages of Malkangiri district, Orissa, India. *Acta Tropica 89*(1), 55–66.

Stich, A. H., C. A. Maxwell, A. A. Haji, D. M. Haji, A. Y. Machano, J. K. Mussa, A. Matteelli, H. Haji, and C. F. Curtis (1994). Insecticide-impregnated bed nets reduce malaria transmission in rural Zanzibar. *Transactions of the Royal Society of Tropical Medicine and Hygiene 88*(2), 150–4.

Sun, B., S. Neslin, and K. Srinivasan (2003). Measuring the Impact of Promotions on Brand Switching When Consumers Are Forward Looking. *Journal of Marketing Research 40*(4), 389–405.

Tarozzi, A., A. Mahajan, B. Blackburn, D. Kopf, L. Krishnan, and J. Yoong (2011). Micro-loans, bednets and malaria: Evidence from a randomized controlled trial in Orissa (India). Working Paper.

ter Kuile, F. O., D. J. Terlouw, S. K. Kariuki, P. A. Phillips-Howard, L. B. Mirel, W. A. Hawley, J. F. Friedman, Y. P. Shi, M. S. Kolczak, A. A. Lal, J. M. Vulule, and B. L. Nahlen (2003). Impact of permethrin-treated bed nets on malaria, anemia, and growth in infants in an area of intense perennial malaria transmission in western Kenya. *The American Journal of Tropical Medicine and Hygiene 68*(4 Suppl), 68–77.

ter Kuile, F. O., D. J. Terlouw, P. A. Phillips-Howard, W. A. Hawley, J. F. Friedman, S. K. Kariuki, Y. P. Shi, M. S. Kolczak, A. A. Lal, J. M. Vulule, and B. L. Nahlen (2003). Reduction of malaria during pregnancy by permethrin-treated bed nets in an area of intense perennial malaria transmission in western kenya. *The American Journal of Tropical Medicine and Hygiene 68* (4 Suppl), 50–60.

Van Bortel, W., M. Barutwanayo, C. Delacollette, and M. Coosemans (1996). Motivation to acquire and use impregnated mosquito nets in a stable malaria zone in Burundi. *Tropical Medicine & International Health 1* (1), 71–80.

van der Klaauw, W. and K. Wolpin (2008). Social Security and the Retirement and Savings Behavior of Low-Income Households. *Journal of Econometrics 145* (1-2), 21–42.

van der Klaauw W. (2000). On the Use of Expectations Data in Estimating Structural Dynamic Models: An Analysis of Career Choices. Working Paper, University of North Carolina at Chapel Hill.

Vinetz, J. M. and R. H. Gilman (2002). Asymptomatic plasmodium parasitemia and the ecology of malaria transmission. *The American Journal of Tropical Medicine and Hygiene 66* (6), 639–640.

Webster, J., J. Lines, J. Bruce, J. R. Armstrong Schellenberg, and K. Hanson (2005). Which delivery systems reach the poor? A review of equity of coverage of ever-treated nets, never-treated nets, and immunisation to reduce child mortality in Africa. *The Lancet Infectious Diseases 5* (11), 709–17.

Yadav, R. S., R. R. Sampath, and V. P. Sharma (2000). Deltamethrin treated bednets for control of malaria transmitted by *Anopheles culicifacies* (Diptera: Culicidae) in India. *Journal of Medical Entomology 38* (5), 613–622.

Appendix A

Proof of Lemma 1

Proof. For simplicity we assume that the number of households per cluster is equal to 2. Identification with a larger number of households per cluster merely require stronger conditioning statements. As in the standard conditional logit framework, identification only exploits clusters where there is variation in bednet ownership.

Consider the conditional probability of net purchase in this case

$$l\left(b_1, b_2, \mathbf{x}_1, \mathbf{x}_2; u(\cdot)\right) = \frac{\exp\left(g\left(u(\cdot), \mathbf{x}_1\right) b_1 + g\left(u(\cdot), \mathbf{x}_2\right) b_2\right)}{\exp\left(g\left(u(\cdot), \mathbf{x}_1\right)\right) + \exp\left(g\left(u(\cdot), \mathbf{x}_2\right)\right)} \tag{17}$$

Recall that in the non-parametric case, g only depends upon utility though the functions $\Delta(\cdot) \equiv \{\Delta_j(\cdot)\}_{j=1}^3$ which are objects of interest in the likelihood. If the model is not identified, then there exists a function $\tilde{\Delta}$ such that $l(\cdot, \Delta) = l(\cdot, \tilde{\Delta})$ almost everywhere (a.e.). Evaluating the likelihood at the values $(1, 0)$ and $(0, 1)$ for (b_1, b_2) and dividing the two expressions we see that we must have

$$g\left(\Delta, \mathbf{x}_1\right) - g\left(\Delta, \mathbf{x}_2\right) = g\left(\tilde{\Delta}, \mathbf{x}_1\right) - g\left(\tilde{\Delta}, \mathbf{x}_2\right) \tag{18}$$

so that if the model was linear and the Δ functions were constants (e.g. $g\left(\Delta, \mathbf{x}\right) = x_a \Delta_1 + x_b \Delta_2$) identification would follow from the standard conditions for linear index conditional logit models.

By assumption 4, the distribution of δ_1 conditional on all the other random variables in (17) contains at least two points of support. Denote these points by δ_1' and δ_1''. Now recall the definition of $g(\cdot)$ for the non-parametric utility case (see equation 5). Evaluating (17) at δ_1' and δ_1'' while keeping $(\mathbf{x}_1 \backslash \delta_1, \mathbf{x}_2)$ constant, and finally taking differences, we deduce that if the model is not identified then

$$\left(\delta_1' - \delta_1''\right) \Delta_3(\mathbf{x}_1) = \left(\delta_1' - \delta_1''\right) \tilde{\Delta}_3(\mathbf{x}_1)$$

This in turn implies that $\Delta_3(\mathbf{x}_1) = \tilde{\Delta}_3(\mathbf{x}_1)$. The above argument can be applied for every value in the support of $\mathbf{x}_1$ and we can then conclude that the function $\Delta_3(\cdot)$ is identified over the relevant support of $\mathbf{x}_1$. Exactly analogous arguments yield identification of $\Delta_1(\cdot)$ and $\Delta_2(\cdot)$. Note that it is the presence of the fixed effect that requires us to condition δ_1 on $\mathbf{x}_2$. Without a fixed effect we would not need to condition on the second household's covariates. Note also that the utility functions themselves are not identified since we have a system of three linear equations in four unknowns. The preceding argument demonstrates identification of an infinite dimensional parameter within a maximum likelihood framework. We note that *estimation* in such a setting is much more complicated and is deferred to future work (see Chen 2004 for a discussion of sieve based procedures in such models). □

Proof of Lemma 2

Proof. From the previous arguments we know that the $\Delta_3(\cdot)$ is non-parametrically identified. We first use this result to prove that the parameter vector $\boldsymbol{\tau}$ is identified[31] and then show that the parameters $\boldsymbol{\alpha} \backslash \alpha_2$ are identified as well. Substituting the parametric utility specification (7) into $\Delta_3(\cdot)$

$$\Delta_3(\mathbf{x}) = (\alpha_0 + \alpha_1) \left(c\left(1, h\right)^{1 - \gamma(\mathbf{x}_a, h; \boldsymbol{\tau})} - c\left(1, m\right)^{1 - \gamma(\mathbf{x}_a, m; \boldsymbol{\tau})}\right) - \alpha_4 - \mathbf{x}_m' \alpha_5 - \alpha_6.$$

We next compute the ration of the second derivative of this function to its first derivative. By the continuity Assumption (7) we can differentiate the function respect to $c\left(1, m\right)$ to obtain

$$\frac{\partial \Delta_3(\mathbf{x})}{\partial c(1, m)} = (\alpha_0 + \alpha_1) \left(1 - \gamma\left(\mathbf{x}_a, m; \boldsymbol{\tau}\right)\right) c\left(1, m\right)^{-\gamma(\mathbf{x}_a, m; \boldsymbol{\tau})}$$

---

[31]We thank a referee for suggesting a clearer proof.

The second derivative yields

$$\frac{\partial^2 \Delta_3(\mathbf{x})}{\partial c(1,m)^2} = (\alpha_0 + \alpha_1)\left(1 - \gamma\left(\mathbf{x}_a, m; \boldsymbol{\tau}\right)\right)\left(-\gamma(\mathbf{x}_a, m; \tau)\right)c\left(1, m\right)^{-\gamma(\mathbf{x}_a, m; \boldsymbol{\tau})-1}$$

so that we can identify $\gamma(\cdot)$ as

$$-c(1,m)\frac{\frac{\partial^2 \Delta_3(\mathbf{x})}{\partial c(1,m)^2}}{\frac{\partial \Delta_3(\mathbf{x})}{\partial c(1,m)}} = \gamma(\mathbf{x}_a, m, \tau)$$

Given the logit parametrization of γ in (8) and Assumption 8, if $\mathbf{x}_a$ has a non-singular second moment matrix, then $\boldsymbol{\tau}$ is identified.

We now show that certain elements of $\boldsymbol{\alpha}$ are identified. The parameter α_2 is not identified since it is an intercept term in the equation. Relaxing the fixed effect assumption will permit this parameter to be identified as well. The remaining parameters are identified. Using equation (7), recall that the function $g(\cdot)$ can be rewritten as

$$g(\mathbf{x}; \boldsymbol{\alpha}, \boldsymbol{\tau}) = \alpha_2 + \boldsymbol{\alpha}_3' \mathbf{x}_b - \alpha_4 \delta - \mathbf{x}_m' \alpha_5 - \alpha_6(\pi - \delta) + (\alpha_0 + \alpha_1)h_1(\pi, \delta, c, \mathbf{x}_a, \boldsymbol{\tau}) - \alpha_0 h_0(\pi, c, \mathbf{x}_a, \boldsymbol{\tau}).$$

Since $\boldsymbol{\tau}$ is identified the objects $h_0(\mathbf{x}, \boldsymbol{\tau}), h_1(\mathbf{x}, \boldsymbol{\tau})$ are also identified. Therefore g is now linear in the remaining unknown parameters. If these parameters are not identified, then there exists a $\tilde{\boldsymbol{\alpha}}$ such that

$$g(\mathbf{x}_1; \boldsymbol{\alpha}, \boldsymbol{\tau}) - g\left(\mathbf{x}_2; \boldsymbol{\alpha}, \boldsymbol{\tau}\right) = g\left(\mathbf{x}_1; \tilde{\boldsymbol{\alpha}}, \boldsymbol{\tau}\right) - g\left(\mathbf{x}_2; \tilde{\boldsymbol{\alpha}}, \boldsymbol{\tau}\right) \tag{19}$$

or rewriting,

$$
\begin{aligned}
& \boldsymbol{\alpha}_3'(\mathbf{x}_{1b} - \mathbf{x}_{2b}) - \alpha_4(\delta_1 - \delta_2) - (\mathbf{x}_{1m} - \mathbf{x}_{2m})'\boldsymbol{\alpha}_5 - \alpha_6(\pi_1 - \delta_1 - \pi_2 + \delta_2) \\
& + (\alpha_0 + \alpha_1)\left(h_1(\mathbf{x}_1, \boldsymbol{\tau}) - h_1(\mathbf{x}_2, \boldsymbol{\tau})\right) - \alpha_0(h_0(\mathbf{x}_1\boldsymbol{\tau}) - h_0(\mathbf{x}_2, \boldsymbol{\tau})) \\
= \ & \tilde{\boldsymbol{\alpha}}_3'(\mathbf{x}_{1b} - \mathbf{x}_{2b}) - \tilde{\alpha}_4(\delta_1 - \delta_2) - (\mathbf{x}_{1m} - \mathbf{x}_{2m})'\tilde{\boldsymbol{\alpha}}_5 - \tilde{\alpha}_6(\pi_1 - \delta_1 - \pi_2 + \delta_2) \\
& + (\tilde{\alpha}_0 + \tilde{\alpha}_1)\left(h_1(\mathbf{x}_1, \boldsymbol{\tau}) - h_1(\mathbf{x}_2, \boldsymbol{\tau})\right) - \tilde{\alpha}_0(h_0(\mathbf{x}_1\boldsymbol{\tau}) - h_0(\mathbf{x}_2, \boldsymbol{\tau}))
\end{aligned}
$$

But by Assumption (8) this can only happen if $(\alpha_0, \alpha_1 + \alpha_0, \boldsymbol{\alpha}_3', \alpha_4, \boldsymbol{\alpha}_5, \alpha_6)$ is equal to $(\tilde{\alpha}_0, \tilde{\alpha}_1 + \tilde{\alpha}_0, \tilde{\boldsymbol{\alpha}}_3', \tilde{\alpha}_4, \tilde{\boldsymbol{\alpha}}_5, \tilde{\alpha}_6)$ We therefore conclude that the vector $(\alpha_0, \alpha_1, \boldsymbol{\alpha}_3', \alpha_4, \boldsymbol{\alpha}_5, \alpha_6)$ is identified. Note instead that the parameter α_2 is the same for all individuals, so that it cancels out in equation (19) and is therefore not identified. $\qquad \square$

Remark 1. *The identification argument above is non-parametric in the form of $\gamma(\cdot)$ in the sense that even without the logit parametrization we have shown that $\tilde{\gamma}(\mathbf{x}_a, m) = \gamma(\mathbf{x}_a, m)$.*

Proof of Lemma 3

Proof. The conditional median zero assumption implies that the conditional median of b^* is equal to an indicator function

$$median(b^*|\mathbf{x}) = \mathbb{I}\{g(\Delta_1 - \Delta_2, \mathbf{x}) \geq 0\}$$

Following the analogy principle then one could estimate the parameters by minimizing the sample version of

$$\mathbb{E}|b^* - \mathbb{I}\{g(\Delta_1 - \Delta_2, \mathbf{x}) \geq 0\}|.$$

Conditions for identification then reduce to showing that if

$$\mathbb{I}\{g(\Delta_1 - \Delta_2, \mathbf{x}) \geq 0\} = \mathbb{I}\{g(\tilde{\Delta}_1 - \tilde{\Delta}_2, \mathbf{x}) \geq 0\} \tag{20}$$

with probability one conditional on $\mathbf{x}$, then $\Delta_1 - \Delta_2 = \tilde{\Delta}_1 - \tilde{\Delta}_2$. We will show the negation of this statement, namely that if two candidate parameter values are not equal then with positive probability the two indicator functions associated with them are also not equal.

Fix an alternative parameter value $\Delta \equiv \Delta_1 - \Delta_2 \neq \tilde{\Delta} \equiv \tilde{\Delta}_1 - \tilde{\Delta}_2$ and define the sets over which the associated indicator functions disagree

$$S_1\left(\tilde{\Delta}\right) = \left\{\mathbf{x} : g\left(\Delta, \mathbf{x}\right) \geq 0 > g\left(\tilde{\Delta}, \mathbf{x}\right)\right\}$$

$$S_2\left(\tilde{\Delta}\right) = \left\{\mathbf{x} : g\left(\tilde{\Delta}, \mathbf{x}\right) \geq 0 > g\left(\Delta, \mathbf{x}\right)\right\}$$

We show that under the assumptions in the lemma, $\mathbb{P}\left(S_1 \cup S_2\right) > 0$. First, redefine the sets as

$$S_1\left(a, t\right) = \left\{\mathbf{x} : g\left(\Delta, \mathbf{x}\right) - g\left(\tilde{\Delta}, \mathbf{x}\right) > g\left(\Delta, \mathbf{x}\right) \geq 0\right\}$$

$$S_2\left(a, t\right) = \left\{\mathbf{x} : g\left(\Delta, \mathbf{x}\right) - g\left(\tilde{\Delta}, \mathbf{x}\right) \leq g\left(\Delta, \mathbf{x}\right) < 0\right\}.$$

By assumption, the distribution of $g\left(\Delta, \mathbf{x}\right) = \pi\left(\Delta_1 - \Delta_2\right) + \delta$ conditional upon $(\mathbf{x}\backslash\delta)$ has density in a neighborhood of zero. This ensures that one of the sets above will always have positive probability. $\square$

Proof of Lemma 4

Proof. By Assumption 2 we know that

$$median\left(b|\mathbf{x}\right) = \mathbb{I}\left\{g(\mathbf{x}; \boldsymbol{\alpha}, \boldsymbol{\tau}) \geq 0\right\}$$

so that identification is equivalent to showing that

$$(\boldsymbol{\alpha}, \boldsymbol{\tau}) \neq (\mathbf{a}, \mathbf{t}) \Rightarrow \mathbb{I}\left\{g(\mathbf{x}; \boldsymbol{\alpha}, \boldsymbol{\tau}) \geq 0\right\} \neq \mathbb{I}\left\{g(\mathbf{x}; \mathbf{a}, \mathbf{t}) \geq 0\right\},$$

where the last inequality is interpreted as holding with positive probability. Next, define the sets over which the associated indicator functions disagree

$$S_1(\mathbf{a}, \mathbf{t}) = \left\{\mathbf{x} : g\left(\mathbf{x}; \boldsymbol{\alpha}, \boldsymbol{\tau}\right) \geq 0 > g(\mathbf{x}; \mathbf{a}, \mathbf{t})\right\}$$

$$S_2\left(\mathbf{a}, \mathbf{t}\right) = \left\{\mathbf{x} : g\left(\mathbf{x}; \mathbf{a}, \mathbf{t}\right) \geq 0 > g\left(\mathbf{x}; \boldsymbol{\alpha}, \boldsymbol{\tau}\right)\right\}$$

We will show that under the assumptions in the lemma, $\mathbb{P}\left(S_1 \cup S_2\right) > 0$. First, redefine the sets as

$$S_1\left(\mathbf{a}, \mathbf{t}\right) = \left\{\mathbf{x} : g\left(\mathbf{x}; \boldsymbol{\alpha}, \boldsymbol{\tau}\right) - g\left(\mathbf{x}; \mathbf{a}, \mathbf{t}\right) > g\left(\mathbf{x}; \boldsymbol{\alpha}, \boldsymbol{\tau}\right) \geq 0\right\}$$

$$S_2\left(\mathbf{a}, \mathbf{t}\right) = \left\{\mathbf{x} : g\left(\mathbf{x}; \boldsymbol{\alpha}, \boldsymbol{\tau}\right) - g\left(\mathbf{x}; \mathbf{a}, \mathbf{t}\right) \leq g(\mathbf{x}; \boldsymbol{\alpha}, \boldsymbol{\tau}) < 0\right\}$$

In order to facilitate analysis of these sets we rewrite the median as

$$\mathbb{I}\left\{g(\mathbf{x}; \boldsymbol{\alpha}, \boldsymbol{\tau}) \geq 0\right\} = \mathbb{I}\left\{c(0, m) - s\left(\boldsymbol{\alpha}, \boldsymbol{\tau}, \mathbf{x}\backslash c(0, m)\right) \leq 0\right\}$$

where

$$s\left(\boldsymbol{\alpha}, \boldsymbol{\tau}, \mathbf{x}\backslash c(0, m)\right) = \left[\frac{1}{\pi}\left(\alpha_1 h_1\left(\mathbf{x}_a, c(1, h), c(1, m); \boldsymbol{\tau}\right) - (1 - \pi)c(0, h)^{1-\gamma(\mathbf{x}_a, h; \boldsymbol{\tau})} + \boldsymbol{\alpha}_3 \mathbf{x}_b\right)\right]^{\frac{1}{1-\gamma(\mathbf{x}_a, m; \boldsymbol{\tau})}}.$$

Then we can rewrite the sets as

$$\begin{aligned} S_1(\mathbf{a}, \mathbf{t}) &= \left\{\mathbf{x} : s\left(\boldsymbol{\alpha}, \boldsymbol{\tau}, \mathbf{x}\backslash c(0, m)\right) - s\left(\mathbf{a}, \mathbf{t}, \mathbf{x}\backslash c(0, m)\right) < c(0, m) - s\left(\boldsymbol{\alpha}, \boldsymbol{\tau}, \mathbf{x}\backslash c(0, m)\right) \leq 0\right\} \\ S_2\left(\mathbf{a}, \mathbf{t}\right) &= \left\{\mathbf{x} : s\left(\mathbf{a}, \mathbf{t}, \mathbf{x}\backslash c(0, m)\right) - s\left(\boldsymbol{\alpha}, \boldsymbol{\tau}, \mathbf{x}\backslash c(0, m)\right) \geq c(0, m) - s\left(\boldsymbol{\alpha}, \boldsymbol{\tau}, \mathbf{x}\backslash c(0, m)\right) > 0\right\}. \end{aligned}$$

By assumption, the distribution of $c(0,m)$ conditional on $\mathbf{x}\backslash c(0,m)$ has support in a neighborhood of zero so that one of these sets will always have positive probability as long as

$$s\left(\boldsymbol{\alpha}, \boldsymbol{\tau}, \mathbf{x}\backslash c(0,m)\right) \neq s(\mathbf{a}, \mathbf{t}, \mathbf{x}\backslash c(0,m)). \tag{21}$$

We next show that under the stated assumptions (21) holds with positive probability. First, assume the statement is not true so that there does exist an (a, t) such that two functions coincide with probability one. Then, we must have

$$\left(\pi^{-1}\right)^{\frac{1}{1-\gamma(\mathbf{x}_a, m; \boldsymbol{\tau})} - \frac{1}{1-\gamma(\mathbf{x}_a, m; \mathbf{t})}} = \frac{\left(\alpha_1 h_1\left(\boldsymbol{\tau}, \mathbf{x}_a, c(1,h), c(1,m)\right) - (1-\pi)\, c(0,h)^{1-\gamma(\mathbf{x}_a, h; \boldsymbol{\tau})} + \boldsymbol{\alpha}_3 \mathbf{x}_b\right)}{\left(a_1 h_1\left(\mathbf{t}, \mathbf{x}_a, c(1,h), c(1,m)\right) - (1-\pi)\, c(0,h)^{1-\gamma(\mathbf{x}_a, h; \mathbf{t})} + \mathbf{a}_3 \mathbf{x}_b\right)}$$

and defining

$$\tilde{s}\left(\boldsymbol{\alpha}, \boldsymbol{\tau}, \mathbf{x}\backslash c(0,m)\right) \equiv \left(\alpha_1 h_1\left(\boldsymbol{\tau}, \mathbf{x}_a, c(1,h), c(1,m)\right) - (1-\pi)\, c(0,h)^{1-\gamma(\mathbf{x}_a, h; \boldsymbol{\tau})} + \boldsymbol{\alpha}_3 \mathbf{x}_b\right)$$

and

$$e\left(\pi, \delta, \mathbf{x_a}, \mathbf{t}, \boldsymbol{\tau}\right) \equiv \left(\pi^{-1}\right)^{\frac{1}{1-\gamma(\mathbf{x}_a, m; \boldsymbol{\tau})} - \frac{1}{1-\gamma(\mathbf{x}_a, m; \mathbf{t})}}$$

we obtain

$$e\left(\pi, \delta, \mathbf{x_a}, \mathbf{t}, \boldsymbol{\tau}\right) \tilde{s}\left(\mathbf{a}, \mathbf{t}, \mathbf{x}\backslash c(0,m)\right) = \tilde{s}\left(\boldsymbol{\alpha}, \boldsymbol{\tau}, \mathbf{x}\backslash c(0,m)\right)$$

By assumption, there exists a variable in $\tilde{s}\left(\cdot\right)$ that does not exist in $h\left(\cdot\right)$ that is continuously distributed over an interval conditional upon $\mathbf{x}\backslash c(1,m)$, namely the variable $c(0,h)$. This justifies taking derivatives with respect to this variable on both sides of the equality above and we obtain (suppressing dependence upon other covariates)

$$e\left(1 - \gamma(h; \boldsymbol{\tau})\right) c(0,h)^{-\gamma(h; \boldsymbol{\tau})} = \left(1 - \gamma(h; \mathbf{t})\right) c(0,h)^{-\gamma(h; \mathbf{t})}$$

and collecting terms we obtain

$$c(0,h)^{\gamma(h; \mathbf{t}) - \gamma(h; \boldsymbol{\tau})} = \frac{\left(1 - \gamma(h; \mathbf{t})\right)}{\left(1 - \gamma(h; \boldsymbol{\tau})\right)} e,$$

where the right hand side is well defined since the denominator is never equal to zero and as long as $\pi \neq 0$ with positive probability (note we can always restrict attention to the set over which π is positive and condition all subsequent arguments on π belonging to this set). Taking derivatives once again we obtain

$$c(0,h)^{\gamma(h; \mathbf{t}) - \gamma(h; \boldsymbol{\tau}) - 1}\left(\gamma(h; \mathbf{t}) - \gamma(h; \boldsymbol{\tau})\right) = 0 \quad a.e.$$

By assumption, consumption is almost everywhere positive so we must have $\gamma(h; \mathbf{t}) = \gamma(h; \boldsymbol{\tau})$. Since by assumption the second moment matrix of $\mathbf{x}_a$ is assumed non-singular, we must therefore have $\mathbf{t} = \boldsymbol{\tau}$ so that $\boldsymbol{\tau}$ is identified. Once $\boldsymbol{\tau}$ is identified then as long as the vector $(h_1, \mathbf{x}_b)$ has non-singular second moment matrix, $\boldsymbol{\alpha}$ will also be identified.

$\square$

Proof of Lemma 5

Proof. Under Assumption 3, it is straightforward to show that the following condition hold for households 1 and 2 in a cluster (see e.g. Manski 1987),

$$median\left(b_1 - b_2 | \mathbf{w}, b_1 \neq b_2\right) = sign\left[g\left(\mathbf{x}_1; \boldsymbol{\alpha}, \boldsymbol{\tau}\right) - g\left(\mathbf{x}_2; \boldsymbol{\alpha}, \boldsymbol{\tau}\right)\right].$$

Identification requires then that for any candidate pair $(\mathbf{a}, \mathbf{t})$

$$(\mathbf{a}, \mathbf{t}) \neq (\boldsymbol{\alpha}, \boldsymbol{\tau}) \Rightarrow sign\left[g\left(\mathbf{x}_1; \boldsymbol{\alpha}, \boldsymbol{\tau}\right) - g\left(\mathbf{x}_2; \boldsymbol{\alpha}, \boldsymbol{\tau}\right)\right] \neq sign\left[g\left(\mathbf{x}_1; \boldsymbol{\alpha}, \boldsymbol{\tau}\right) - g\left(\mathbf{x}_2; \boldsymbol{\alpha}, \boldsymbol{\tau}\right)\right]$$

In order to study this further, consider the sets (on the support of $\mathbf{w} = (\mathbf{x_1}, \mathbf{x_2})$) where these functions disagree

$$
\begin{aligned}
S_1(\mathbf{a}, \mathbf{t}) &= \left\{\mathbf{w} : g\left(\mathbf{x}_1; \boldsymbol{\alpha}, \boldsymbol{\tau}\right) - g\left(\mathbf{x}_2; \boldsymbol{\alpha}, \boldsymbol{\tau}\right) \geq 0, g\left(\mathbf{x}_1; \mathbf{a}, \mathbf{t}\right) - g\left(\mathbf{x}_2; \mathbf{a}, \mathbf{t}\right) < 0\right\} \\
S_2(\mathbf{a}, \mathbf{t}) &= \left\{\mathbf{w} : g\left(\mathbf{x}_1; \boldsymbol{\alpha}, \boldsymbol{\tau}\right) - g\left(\mathbf{x}_2; \boldsymbol{\alpha}, \boldsymbol{\tau}\right) < 0, g\left(\mathbf{x}_1; \mathbf{a}, \mathbf{t}\right) - g\left(\mathbf{x}_2; \mathbf{a}, \mathbf{t}\right) \geq 0\right\}.
\end{aligned}
$$

We show that at least one of these sets always occurs with positive probability for any choice of $(\mathbf{a}, \mathbf{t})$ so that the model is point identified. First, the index function for household s when the normalization imposed in Lemma 5 hold can be written as

$$
\begin{aligned}
g\left(\mathbf{x}_s; \boldsymbol{\alpha}, \boldsymbol{\tau}\right) &\equiv \mathbf{x}'_{s,b}\boldsymbol{\alpha}_3 - \alpha_4\delta_s - \mathbf{x}'_{s,m}\boldsymbol{\alpha}_5 + \alpha_6(\pi_s - \delta_s) + (1 + \alpha_1)h_{s,1}(\boldsymbol{\tau}) - h_{s,0}(\boldsymbol{\tau}) \\
&= \bar{\mathbf{x}}'_s\bar{\boldsymbol{\alpha}} + (1 + \alpha_1)h_{1,s}(\boldsymbol{\tau}) - h_{0,s}(\boldsymbol{\tau}),
\end{aligned}
$$

where the subscript s reflects the dependence of the $h(\cdot)$ functions upon the data, and we have for convenience rewritten the linear part of the index as

$$\bar{\mathbf{x}}'_s\bar{\boldsymbol{\alpha}} \equiv \mathbf{x}'_{s,b}\boldsymbol{\alpha}_3 - \alpha_4\delta_s - \mathbf{x}'_{s,m}\boldsymbol{\alpha}_5 + \alpha_6(\pi_s - \delta_s).$$

Using this redefinition of the index we obtain

$$
\begin{aligned}
&g\left(\mathbf{x}_1; \mathbf{a}, \mathbf{t}\right) - g\left(\mathbf{x}_2; \mathbf{a}, \mathbf{t}\right) \\
&= \left(\bar{\mathbf{x}}'_1 - \bar{\mathbf{x}}'_2\right)\bar{\mathbf{a}} + (1 + a_1)\left(h_{1,1}(\mathbf{t}) - h_{2,1}(\mathbf{t})\right) - \left(h_{1,0}(\mathbf{t}) - h_{2,0}(\mathbf{t})\right) \\
&= \Delta\bar{\mathbf{x}}'\bar{\mathbf{a}} + (1 + a_1)\Delta h_1(\mathbf{t}) - \left(h_{1,0}(\mathbf{t}) - h_{2,0}(\mathbf{t})\right) \\
&= \Delta\bar{\mathbf{x}}'\bar{\mathbf{a}} + (1 + a_1)\Delta h_1(\mathbf{t}) - \left[h_{1,0}(\mathbf{t}) - (1 - \pi_2)c_2(0, h)^{1-\gamma(h_2;\mathbf{t})} - \pi_2 c_2(0, m)^{1-\gamma(m_2;\mathbf{t})}\right],
\end{aligned}
$$

Consider now the inequality

$$g\left(\mathbf{x}_1; \mathbf{a}, \mathbf{t}\right) - g\left(\mathbf{x}_2; \mathbf{a}, \mathbf{t}\right) < 0 \tag{22}$$

which can be rewritten as

$$c_2(0, m) < \left[-\frac{1}{\pi_2}\left(\Delta\bar{\mathbf{x}}'\bar{\mathbf{a}} + (1 + a_1)\Delta h_1(\mathbf{t}) - h_{1,0}(\mathbf{t}) + (1 - \pi_2)c_2(0, h)^{1-\gamma(h_2;\mathbf{t})}\right)\right]^{\frac{1}{1-\gamma(m_2;\mathbf{t})}},$$

and recall that we define

$$s\left(\mathbf{a}, \mathbf{t}, \mathbf{x}\backslash c_2(0, m)\right) \equiv \left[-\frac{1}{\pi_2}\left(\Delta\bar{\mathbf{x}}'\bar{\mathbf{a}} + a_1\Delta h_1(\mathbf{t}) - h_{1,0}(\mathbf{t}) + (1 - \pi_2)c_2(0, h)^{1-\gamma(h_2;\mathbf{t})}\right)\right]^{\frac{1}{1-\gamma(m_2;\mathbf{t})}},$$

so that, suppressing the dependence of $s(.)$ on $(\mathbf{x}\backslash c_2(0, m))$ for brevity, we can rewrite the inequality as

$$\pi_2 c_2(0, m)^{1-\gamma(m_2;\boldsymbol{\tau})} < \pi_2 s(\mathbf{a}, \mathbf{t})^{1-\gamma(m_2;\boldsymbol{\tau})},$$

and subtracting $\pi_2 s(\boldsymbol{\alpha}, \boldsymbol{\tau})^{1-\gamma(m_2;\boldsymbol{\tau})}$ on both sides we obtain

$$\pi_2 c_2(0, m)^{1-\gamma(m_2;\boldsymbol{\tau})} - \pi_2 s(\boldsymbol{\alpha}, \boldsymbol{\tau})^{1-\gamma(m_2;\boldsymbol{\tau})} < \pi_2 s(\mathbf{a}, \mathbf{t})^{1-\gamma(m_2;\boldsymbol{\tau})} - \pi_2 s(\boldsymbol{\alpha}, \boldsymbol{\tau})^{1-\gamma(m_2;\boldsymbol{\tau})}$$

which is equivalent to

$$g\left(\mathbf{x}_1; \boldsymbol{\alpha}, \boldsymbol{\tau}\right) - g\left(\mathbf{x}_2; \boldsymbol{\alpha}, \boldsymbol{\tau}\right) < \pi_2 s(\mathbf{a}, \mathbf{t})^{1-\gamma(m_2;\boldsymbol{\tau})} - \pi_2 s(\boldsymbol{\alpha}, \boldsymbol{\tau})^{1-\gamma(m_2;\boldsymbol{\tau})}.$$

Consider the case where the term on the right hand side of the inequality is positive (if it is negative, one works instead with the reverse inequality from that in (22) and focusses on the set S_2). Then we can write

$$S_1(\mathbf{a}, \mathbf{t}) = \left\{ \mathbf{x} : 0 \leq g(\mathbf{x}_1; \boldsymbol{\alpha}, \boldsymbol{\tau}) - g(\mathbf{x}_2; \boldsymbol{\alpha}, \boldsymbol{\tau}) < \pi_2 s(\mathbf{a}, \mathbf{t})^{1-\gamma(m_2; \boldsymbol{\tau})} - \pi_2 s(\boldsymbol{\alpha}, \boldsymbol{\tau})^{1-\gamma(m_2; \boldsymbol{\tau})} \right\}.$$

By assumption, the distribution of $g(\mathbf{x}_1; \boldsymbol{\alpha}, \boldsymbol{\tau}) - g(\mathbf{x}_2; \boldsymbol{\alpha}, \boldsymbol{\tau})$ conditional on $\mathbf{x} \backslash c_2(0, m)$ has support in a neighborhood of zero so to complete the proof we have to show that

$$s(\mathbf{a}, \mathbf{t}) \neq s(\boldsymbol{\alpha}, \boldsymbol{\tau}) \tag{23}$$

holds with positive probability (note we are assuming $\pi_2 > 0$ and the argument can be viewed as being restricted to the set where this is true). The proof is by contradiction. Suppose that (23) does not hold so that $s(\mathbf{a}, \mathbf{t}) = s(\boldsymbol{\alpha}, \boldsymbol{\tau})$. This equality can be rewritten as

$$\kappa \left(\Delta \bar{\mathbf{x}}' \bar{\boldsymbol{\alpha}} + (1 + \alpha_1) \Delta h_1(\boldsymbol{\tau}) - h_{1,0}(\boldsymbol{\tau}) + (1 - \pi_2) c_2(0, h)^{1-\gamma(h_2; \boldsymbol{\tau})} \right)^{\frac{1}{1-\gamma(m_2; \boldsymbol{\tau})}}$$

$$= \left(\Delta \bar{\mathbf{x}}' \bar{\mathbf{a}} + (1 + a_1) \Delta h_1(\mathbf{t}) - h_{1,0}(\mathbf{t}) + (1 - \pi_2) c_2(0, h)^{1-\gamma(h_2; \mathbf{t})} \right)^{\frac{1}{1-\gamma(m_2; \mathbf{t})}},$$

where

$$\kappa = \left(-\frac{1}{\pi_2} \right)^{\frac{1}{1-\gamma(m_2; \boldsymbol{\tau})} - \frac{1}{1-\gamma(m_2; \mathbf{t})}}.$$

Next, by assumption the distribution of $c_2(0, h)$ conditional on $\mathbf{w} \backslash \{c_2(0, m), c_2(0, h)\}$ is continuously distributed over some range so that we can take derivatives with respect to $c_2(0, h)$ in that range and rearrange to obtain

$$\left(\frac{1 - \gamma(h_2; \boldsymbol{\tau})}{1 - \gamma(h_2; \mathbf{t})} \right) \left(\frac{1 - \gamma(m_2; \mathbf{t})}{1 - \gamma(m_2; \boldsymbol{\tau})} \right) \left(\Delta \bar{\mathbf{x}}' \bar{\mathbf{a}} + (1 + a_1) \Delta h_1(\mathbf{t}) - h_{1,0}(\mathbf{t}) + (1 - \pi_2) c_2(0, h)^{1-\gamma(h_2; \mathbf{t})} \right)$$

$$= \left(\Delta \bar{\mathbf{x}}' \bar{\boldsymbol{\alpha}} + (1 + a_1) \Delta h_1(\boldsymbol{\tau}) - h_{1,0}(\boldsymbol{\tau}) + (1 - \pi_2) c_2(0, h)^{1-\gamma(h_2; \boldsymbol{\tau})} \right) c_2(0, h)^{\gamma(h_2; \boldsymbol{\tau}) - \gamma(h_2; \mathbf{t})}.$$

Finally, by assumption the distribution of $c_1(0, m)$ conditional on $\mathbf{w} \backslash \{c_2(0, m), c_1(0, m)\}$ is continuously distributed over some range so that taking derivatives with respect to $c_1(0, m)$ and rearranging:

$$\left(\frac{1 - \gamma(h_2; \boldsymbol{\tau})}{1 - \gamma(h_2; \mathbf{t})} \right) \left(\frac{1 - \gamma(m_2; \mathbf{t})}{1 - \gamma(m_2; \boldsymbol{\tau})} \right) \left(\frac{1 - \gamma(m_1; \mathbf{t})}{1 - \gamma(m_1; \boldsymbol{\tau})} \right) = c_2(0, h)^{(\gamma(h_2; \boldsymbol{\tau}) - \gamma(h_2; \mathbf{t}))} c_1(0, m)^{(\gamma(m_1; \mathbf{t}) - \gamma(m_1; \boldsymbol{\tau}))},$$

and taking derivatives with respect to $c_1(0, m)$ again we obtain the result that if the model is not identified then we must have

$$c_2(0, h)^{(\gamma(h_2; \boldsymbol{\tau}) - \gamma(h_2; \mathbf{t}))} c_1(0, m)^{(\gamma(m_1; \mathbf{t}) - \gamma(m_1; \boldsymbol{\tau}) - 1)} (\gamma(m_1; \mathbf{t}) - \gamma(m_1; \boldsymbol{\tau})) = 0$$

almost everywhere. Since by assumption consumption is strictly positive, this must imply $\gamma(m_1; \mathbf{t}) - \gamma(m_1; \boldsymbol{\tau}) = 0$ almost everywhere. However, since by assumption the vector $\mathbf{x}_a$ has a non-singular second moment matrix, we must have $\mathbf{t} = \boldsymbol{\tau}$. If that is the case, then we must have

$$\left(\Delta \bar{\mathbf{x}}' \bar{\boldsymbol{\alpha}} + (1 + \alpha_1) \Delta h_1(\boldsymbol{\tau}) - h_{1,0}(\boldsymbol{\tau}) + (1 - \pi_2) c_2(0, h)^{1-\gamma(h_2; \boldsymbol{\tau})} \right)$$

$$= \left(\Delta \bar{\mathbf{x}}' \bar{\mathbf{a}} + (1 + a_1) \Delta h_1(\boldsymbol{\tau}) - h_{1,0}(\boldsymbol{\tau}) + (1 - \pi_2) c_2(0, h)^{1-\gamma(h_2; \boldsymbol{\tau})} \right)$$

which simplifies to

$$\left(\Delta \bar{\mathbf{x}}' \bar{\boldsymbol{\alpha}} + (1 + \alpha_1) \Delta h_1(\boldsymbol{\tau}) \right) = \left(\Delta \bar{\mathbf{x}}' \bar{\mathbf{a}} + (1 + a_1) \Delta h_1(\boldsymbol{\tau}) . \right)$$

By assumption the second moment matrix of $(\Delta \bar{\mathbf{x}}, \Delta h_1(\boldsymbol{\tau}))$ is non-singular and therefore we must have $\boldsymbol{\alpha} = \mathbf{a}$ and so $(\boldsymbol{\alpha}, \mathbf{t}) = (\mathbf{a}, \mathbf{t})$. This is a contradiction since we started by assuming $(\boldsymbol{\alpha}, \mathbf{t}) \neq (\mathbf{a}, \mathbf{t})$. Therefore the parameters $(\alpha_1, \boldsymbol{\alpha}_3', \alpha_4, \boldsymbol{\alpha}_5', \alpha_6, \boldsymbol{\tau}')$ are identified.

$\square$

Appendix B

Measurement Errors in Beliefs

The model for measurement error for which the maximum score assumption is robust is non-standard since it is not the traditional classical errors-in-variable model. In fact, identification results for maximum score type models with classical measurement error are not known. The type of measurement error to which the maximum score is robust is one in which the measurement error is best thought of as being a forecast error given current information (see Imbens and Hyslop 2000) . In this case, assumptions on the symmetry of the error distribution given the realized value of the mismeasured variable are easier to maintain since it is reasonable to assume that the forecast error is unrelated to the reported (mismeasured) value.

Consider the model given by (5) and suppose that the true probabilities used for decision making by the household are given by (π^*, δ^*). However, at the time of the survey respondents rather than reporting these quantities report their best guess for these quantities given their current information. Such errors are plausible if the respondent has forgotten the true values (or the information set used to generate them) and reports values that are best guesses based on the current information possessed. In both instances, since reports are "optimal" in some sense (e.g. they are based on minimizing some loss function). The difference between the reported and true values is then best thought of as a type of forecast error. Depending upon the precise loss function used, there will be restrictions on the dependence between the forecast errors and the optimal responses. For instance, if responses are best guesses based on minimizing an L_1 distance, then the resulting forecast error will be median independent of the reported belief. We assume a somewhat stronger condition: namely that the error has a distribution that is symmetric around zero conditional on the reported belief (as well as other covariates). Formally, the measurement error $\epsilon_\pi \equiv \pi - \pi^*$ is assumed to have a distribution symmetric around zero conditional on $\mathbf{x}$ (note that $\mathbf{x}$ includes π). We impose a similar assumption on ϵ_δ which is defined analogously to ϵ_π. If we further assume that, conditional on $\mathbf{x}$ each element in the vector of errors $(\epsilon_\pi, \epsilon_\delta, \Delta\epsilon)$ is symmetrically distributed around zero and all three random variables are independent of each other, then we obtain the maximum score model with non-parametric utility. Formally, the purchase decision is given by

$$
\begin{aligned}
b &= \mathbb{I}\{\pi^*\Delta_1 + (1 - \pi^*)\Delta_2 + \delta^*\Delta_3 + \Delta\epsilon \geq 0\} \\
&= \mathbb{I}\{\pi\Delta_1 + (1 - \pi)\Delta_2 + \delta\Delta_3 + \tilde{\epsilon} \geq 0\}
\end{aligned}
$$

where

$$
\tilde{\epsilon} = (\Delta_1 - \Delta_2)\epsilon_\pi + \Delta_3\epsilon_\delta + \Delta\epsilon \tag{24}
$$

By the assumption of conditional symmetry for each of the component random variables we have $median(\tilde{\epsilon}|\mathbf{x} = 0$, so that Lemma 1 can still be used to show that the parameters Δ_1, Δ_2 and Δ_3 are identified.

Random Coefficients

Consider the parametric model (6). However, suppose that the coefficients $\boldsymbol{\alpha}$ are no longer fixed (unknown) constants but suppose that it is the random quantities $\boldsymbol{\alpha}^*$ that belong in the equation

instead of $\boldsymbol{\alpha}$. Further, assume that $\boldsymbol{\alpha}^* = \boldsymbol{\alpha} + \epsilon$ where the vector of errors ϵ are independent of $\boldsymbol{\alpha}$. In addition, assume that the elements of $(\epsilon, \Delta\epsilon)$ are independently and symmetrically distributed around zero conditional on $\mathbf{x}$. Then, the composite error term will satisfy the conditional median restriction (2). Formally,

$$\tilde{b} = \mathbb{I}\left\{\boldsymbol{\alpha}_3^* \mathbf{x}_b + (\alpha_0^* + \alpha_1^*) h_1 - \alpha_0^* h_0 + \Delta\epsilon \geq 0\right\}$$

and given $\alpha_j^* = \alpha_j + \epsilon_j$ we can rewrite this as

$$\tilde{b} = \mathbb{I}\left\{\boldsymbol{\alpha}_3 \mathbf{x}_b + (\alpha_0 + \alpha_1) h_1 - \alpha_0 h_0 + \tilde{\epsilon} \geq 0\right\}$$

where

$$\tilde{\epsilon} = \epsilon_0 (h_1 - h_0) + \epsilon_1 h_1 + \epsilon_3 \mathbf{x}_b$$

and by the assumption of conditional symmetry we can conclude that $median(\tilde{\epsilon}|\mathbf{x}) = 0$.
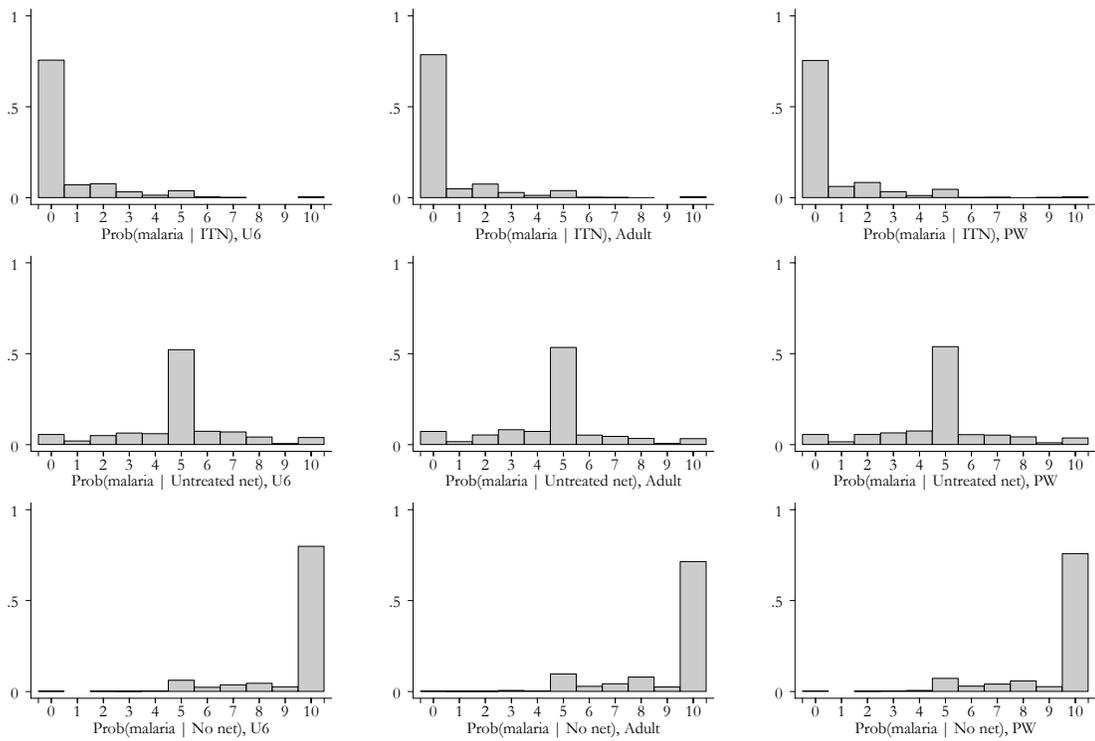
Figure 1: Histograms of subjective beliefs about the protective power of bednets.

Table 1: Summary Statistics

No. of villages	150	
No. of households	1947	
Scheduled Caste (SC)	0.19	
Scheduled Tribe (ST)	0.37	
Other Backward Castes (OBC)	0.37	
Christian[1]	0.02	
	Mean	**s.d.**
Household size	5.46	2.15
# children under 5 in household	0.5	0.7
Highest schooling level in household (years)	8.39	3.72
log(monthly expenditure per head): Itemized[2]	6.32	0.56
Monthly expenditure per head (Rs): Itemized[2]	655	443
No. bednets per head	0.29	0.29
No. ITNs per head	0.04	0.15
Slept under a net last night	0.13	0.29
Slept under an ITN last night	0.03	0.14
Sleeps under a net when mosquitoes peak	0.56	0.45
At least one member had malaria last six months[3]	0.55	0.50
Biomarkers		**obs.**
Malaria	0.12	2621
Moderate anemia (< 10 g/dl) - Males	0.23	965
Moderate anemia (< 10 g/dl) - Females	0.35	1718

Notes: [1]Almost all non-Christian households are Hindu. [2]Estimated by adding reports on expenditure in the month (or year, for some items) before the interview. [3]Calculated from self-reports. All means are calculated using sampling weights.

Table 2: Beliefs about Malaria Risk

	Mean	s.d.
No net (U6)	9.3	1.59
No net (adult)	9.0	1.83
No net (PW)	9.2	1.73
Untreated net (U6)	4.9	2.06
Untreated net (adult)	4.6	2.07
Untreated net (PW)	4.8	2.04
ITN (U6)	0.7	1.47
ITN (adult)	0.6	1.47
ITN (PW)	0.7	1.49

Notes: Figures refer to subjective probabilities that an adult, a child under the age of six or a pregnant woman will contract malaria in the next year, conditional on making regular use of an ITN, of an untreated net or of no net at all.

Table 3: Summary Statistics: Malaria Economic Burden

	All	No Net in hh.	Nets in hh > 0
Any malaria episode reported in last six months	0.52	0.50	0.53
Any positive malaria blood test	0.15	0.16	0.15
	Mean	Median	s.d.
Expected cost of a malaria episode for a working man (Rs)	2791	2300	2175
Expected cost of a malaria episode for a working woman (Rs)	1897	1550	2009
Expected cost of a malaria episode for a non-working member (Rs)	1829	1430	1801

Table 4: Monte Carlo Simulations: Parametric Utility and Parametric Errors

	Mean	Median	SD	IQR
$N = 150$				
τ_0	0.0462	0.0022	0.0736	0.0331
τ_1	-0.5737	-0.5683	0.2166	0.2861
α_0	0.9719	0.9599	0.2206	0.2829
α_1	-0.0014	0.0005	0.0107	0.0086
α_3	3.0157	3.0129	0.1386	0.1853
α_4	1.0065	1.0057	0.1082	0.1490
α_5	1.0866	1.1165	0.8568	1.3036
α_6	-1.0292	-1.0131	0.2026	0.2768
$N = 200$				
τ_0	0.0034	-0.0009	0.0664	0.0251
τ_1	-0.5614	-0.5489	0.2315	0.2439
α_0	0.9819	0.9794	0.2007	0.2511
α_1	-0.0009	0.0005	0.0122	0.0061
α_3	3.0180	3.0207	0.1205	0.1542
α_4	1.0002	0.9948	0.0988	0.1177
α_5	0.9868	0.9749	0.6852	0.9021
α_6	-1.0088	-0.9988	0.1802	0.2461
$N = 300$				
τ_0	0.0009	0.0002	0.0183	0.0207
τ_1	-0.5241	-0.5222	0.1615	0.1941
α_0	0.9943	0.9919	0.1572	0.2049
α_1	-0.0008	-0.0001	0.0059	0.0049
α_3	3.0075	3.0136	0.0910	0.1225
α_4	0.9995	0.9962	0.0788	0.1090
α_5	1.0426	1.0269	0.5915	0.8569
α_6	-1.0027	-1.0041	0.1498	0.1933
$N = 600$				
τ_0	-0.0002	-0.0002	0.0105	0.0129
τ_1	-0.5186	-0.5191	0.1005	0.1173
α_0	0.9891	0.9907	0.1115	0.1431
α_1	-0.0002	0.0003	0.0036	0.0036
α_3	3.0044	3.0023	0.0678	0.0925
α_4	0.9968	0.9989	0.0549	0.0773
α_5	1.0591	1.0443	0.4496	0.6268
α_6	-1.0113	-1.0041	0.0941	0.1289

Notes: Each model was simulated 150 times. The true parameter vector is given by $(0, -.5, 1, 0, 3, 1, 1, -1)$.

Table 5: Monte Carlo Simulations: Parametric Utility with Semi-Parametric Error Specification

	Mean	Median	SD	IQR
$N = 150$				
τ_0	0.0758	0.0434	0.1319	0.0904
τ_1	-0.2666	-0.2032	1.2645	1.6102
α_1	-0.0490	-0.0310	0.0541	0.0664
α_3	2.6338	2.5680	1.3764	1.8634
α_4	1.5467	1.4942	0.5638	0.7459
α_5	1.5630	1.5687	1.2835	1.7698
α_6	-1.7357	-1.5662	0.8569	1.0985
$N = 200$				
τ_0	0.0637	0.0437	0.0686	0.0754
τ_1	-0.4868	-0.4174	1.1876	1.4414
α_1	-0.0454	-0.0317	0.0434	0.0587
α_3	2.8497	2.8045	1.3320	1.7056
α_4	1.5828	1.5520	0.5400	0.7047
α_5	1.8392	1.7481	1.1787	1.6210
α_6	-1.8209	-1.6700	0.7855	0.9344
$N = 300$				
τ_0	0.0590	0.0422	0.0586	0.0678
τ_1	-0.4828	-0.4614	1.0835	1.3860
α_1	-0.0423	-0.0332	0.0383	0.0525
α_3	3.1150	3.0830	1.3318	1.6539
α_4	1.5673	1.5441	0.4655	0.6444
α_5	1.7838	1.7337	1.1888	1.4665
α_6	-1.8205	-1.6560	0.7787	0.9292
$N = 600$				
τ_0	0.0492	0.0378	0.0442	0.0538
τ_1	-0.4785	-0.4669	0.7939	1.0353
α_1	-0.0363	-0.0293	0.0313	0.0421
α_3	3.5461	3.5462	1.1171	1.5641
α_4	1.6296	1.6147	0.3847	0.5450
α_5	1.9149	1.8305	1.0816	1.4428
α_6	-1.8381	-1.7826	0.6835	0.7732

Notes: Each model was simulated 150 times. The true parameter vector is given by $(0, -.5, 0, 3, 1, 1, -1)$.

Table 6: Parametric Utility with Parametric Errors

Variable	Point Estimate	Std.Dev	T-Stat
CRRA function parameters (τ)			
HH Size	0.1187	0.5518	0.21
H. Head Education	1.8980	0.4729	4.01
U5	0.9186	0.4892	1.87
H. Head Age	1.7950	0.6278	2.86
Malaria Status	0.4569	0.4582	1.00
Parameters entering Linearly (α)			
Purchase×Consumption (α_1)	-0.7194	0.6691	-1.07
Purchase×Community Effects (α_3)	0.7095	0.1354	5.24
Malaria Status (α_4)	-0.3468	0.4384	-0.79
Malaria×HH Size (α_{51})	-0.0588	0.0312	-1.88
Malaria×U5 (α_{52})	-0.1954	0.1068	-1.82
Purchase×Malaria (α_6)	0.4658	0.3701	1.25

Notes: Estimated on a cluster of 129 villages with 1865 total households. Maximization was carried out using a gradient based algorithm and standard errors were computed by bootstrapping clusters with 250 replications.

Table 7: Counterfactual Exercises: Parametric Errors

Exogenous Shift	Point Estimate	2.5 Percentile	97.5 Percentile
Price of Nets Falls by 50%			
Average change in index	.002	.001	.0035
% change in index (/100)	.24	.01	.69
Upper bound on households "switching"	4	1	9
Beliefs in net efficacy increase by 50%			
Average change in index	.03	.0314	.0351
% change in index	.10	.002	.33
Upper bound on households "switching"	23	15	32
Beliefs of Community ownership increase by 50%			
Average change in index	.38	.35	.41
% change in index	.89	.62	.91
Upper bound on households "switching"	90	83	114
Beliefs of Community ownership and Net efficacy increase			
Average change in index	.42	.39	.45
% change in index	1.11	.89	1.15
Upper bound on households "switching"	174	152	201

Notes: Standard errors were computed by bootstrapping clusters using 250 replications. The Average change in index is the sample average of the change in the estimated index $g(\hat{\alpha}, \hat{\tau})$ associated with the indicated exogenous shift.

Table 8: Parametric Utility with Semi-Parametric Error Specification

Variable	Point Estimate	2.5 Percentile	97.5 percentile
CRRA function parameters (τ)			
HH Size	-1.54	-2.76	3.76
H. Head Education	-2.14	-4.63	0.43
U5	0.95	-2.68	4.54
H. Head Age	-0.30	-1.45	0.19
Malaria Status	1.53	0.57	3.67
Parameters entering Linearly (α)			
Purchase×Community Effects (α_3)	2.53	0.63	4.57
Malaria Status (α_4)	-1.91	-5.10	-0.18
Malaria×HH Size (α_{51})	0.02	-0.41	0.25
Malaria×U5 (α_{52})	-0.22	-1.82	-0.09
Purchase×Malaria (α_6)	1.76	-0.01	6.65

Notes: Estimated on a cluster of 129 villages with 1865 total households. Maximization was carried out using a genetic algorithm and standard errors were computed using subsampling with subsample size set to 103 and 250 replications.

Table 9: Counterfactual Exercises: Maximum Score

Exogenous Shift	Point Estimate	2.5 Percentile	97.5 Percentile
Price of Nets Falls by 50%	24	1	62
Beliefs in net efficacy increase by 50%	2	0	10
Beliefs of Community ownership increase by 50%	18	2	58

Notes: The figures are estimates of the upper bound of the number of households switching to purchase as a consequence of the exogenous shift. Counterfactuals computed using parametric utility and parameter estimates from the semi-parametric error specification. Standard errors were computed using subsampling at the cluster level with subsample size set to 103 and using 250 replications.