Demand and Welfare Analysis in Discrete Choice Models with Social Interactions*

DEBOPAM BHATTACHARYA† PASCALINE DUPAS
University of Cambridge *Stanford University*

SHIN KANAYA
University of Essex and Kyoto University

December 18, 2022

Abstract

Many real-life settings of individual choice involve social interactions, causing targeted policies to have spillover effects. This paper develops novel empirical tools for analyzing demand and welfare effects of policy interventions in binary choice settings with social interactions. Examples include subsidies for health product adoption and vouchers for attending a high-achieving school. We show that even with fully parametric specifications and unique equilibrium, choice data, that are sufficient for counterfactual *demand* prediction under interactions, are *insufficient* for welfare calculations. This is because distinct underlying mechanisms producing the same interaction coefficient can imply different welfare effects and deadweight-loss from a policy intervention. Standard index restrictions imply distribution-free bounds on welfare. We propose ways to identify and consistently estimate the structural parameters and welfare bounds allowing for unobserved group effects that are potentially correlated with observables and are possibly unbounded. We illustrate our results using experimental data on mosquito-net adoption in rural Kenya.

*We are grateful to three anonymous referees, the editor, and Steven Durlauf, James Heckman, Xenia Matschke, GautamTripathi, and seminar participants at the University of Chicago and the University of Luxembourg for helpful feedback. Bhattacharya acknowledges financial support from the ERC consolidator grant EDWEL; the first outline of this project appeared as part b.3 of that research proposal of March 2015. Part of this research was conducted while Kanaya was visiting the Kyoto Institute of Economic Research, Kyoto University (under the Joint Research Program of the KIER), the support and hospitality of which are gratefully acknowledged.

†Address for correspondence: Faculty of Economics, University of Cambridge, CB3 9DD. Phone (+44)7503858289, email: debobhatta@gmail.com

1 INTRODUCTION

Social interaction models – where an individual's payoff from an action depends on aggregate choice – feature prominently in economic and sociological research. In this paper, we address a substantively important issue that has received limited attention within these literatures, viz., how to conduct welfare analysis of economic policy intervention in such settings. Examples include subsidies for adopting a health product and merit-based vouchers for attending a high-achieving school, where the welfare gain of beneficiaries may be accompanied by spillover-led welfare effects on those unable to adopt or move, respectively. Ex-ante welfare analysis of policies is ubiquitous in economic applications, and informs the practical decision of whether to implement the policy in question. Furthermore, common public interventions such as taxes and subsidies are often motivated by efficiency losses resulting from externalities. Therefore, it is important to develop empirical methods for welfare analysis in presence of such externalities, which cannot be done using available tools in the literature. Developing such methods and making them practically relevant also requires one to clarify and extend some aspects of existing empirical models of social interaction.

Literature Review and Contributions: Seminal contributions to the econometrics of social interactions include Manski (1993) for continuous outcomes, and Brock and Durlauf (2001) (henceforth, BD01) for binary outcomes. Bisin, Moro and Topa (2011) discuss some issues related to identification and estimation of structural parameters in choice models with social interaction and multiple equilibria but do not cover welfare analysis. More recently, there has been a surge of research on the related theme of network models, cf. de Paula (2017) who provides a comprehensive review of the relevant literature. On the other hand, the econometric analysis of welfare in standard discrete choice settings, i.e., with heterogeneous consumers but without social spillover, started with Domencich and McFadden (1977), with later contributions by Daly and Zachary (1978), Small and Rosen (1981), and Bhattacharya (2018). The present paper builds on these two separate literatures to examine how social interactions influence welfare effects of policy interventions and the identifiability of such welfare effects from standard choice data. In the context of a logit binary choice model with social interactions, BD01 Sec 3.3 equations (16) and (17) discussed how to infer the sign of the differences between expected ex ante (indirect) utility at each possible equilibria resulting from the policy intervention being studied. This differs from the average of individual compensating variations that restore realized individual utilities to their pre-intervention level which is a money metric, unlike the BD01 measure, and hence can be directly compared with the cost of the intervention, yielding a theoretically justified measure of deadweight loss. Consequently, this measure has received the most attention in the recent literature on applied welfare analysis, cf. Hausman and Newey (2016), Bhattacharya (2015), McFadden and Train (2019). However, in settings involving spillover, we cannot use the methods of the above papers, as they do not allow for individual utilities to be affected by aggregate choices – a feature that has fundamental implications for welfare analysis. Therefore, new methods are required for welfare calculations under spillover, which we develop in the present paper.

Our starting point is a theoretically coherent empirical model where many individuals with some observed and some unobserved attributes interact with each other to produce the aggregate choice in equilibrium before and after the policy intervention. Individual choice data can be used to estimate identifiable parameters of this model in BD01, which can then be used to predict counterfactual *demand*, i.e. equilibrium choice probabilities resulting from a hypothetical price intervention, e.g., a price subsidy. However, we show that unlike counterfactual *demand* estimation, *welfare* effects are generically not identified from standard choice data under interactions, even when utilities and the distribution of unobserved heterogeneity are parametrically specified, equilibrium is unique, and there are no endogeneity concerns. To understand the heuristics behind under-identification, consider the empirical example of evaluating the welfare effect of subsidizing an anti-malarial, insecticide-treated mosquito net. Suppose, under suitable restrictions, we can model choice behavior in this setting via a Brock-Durlauf type social interaction model, and the data can identify the coefficient on the social interaction term. However, this coefficient may reflect an aggregate effect of several distinct mechanisms, viz. (a) a social preference for conforming, (b) learning from others' experiences, (c) a health-concern led desire to protect oneself from mosquitoes deflected from neighbors who adopt a bednet, and (d) desire to free-ride on other users who increase herd-immunity by protecting themselves and/or protect neighbors via the insecticide effects. These distinct mechanisms, with different magnitudes in general, can make the social interaction coefficient positive, but are not separately identifiable from choice data (only their sum is). But they have different implications for welfare if, say, a subsidy is introduced. In particular, if spillovers are all due to preference for social conforming or learning and there is no (perceived) health externality, then as more neighbors buy, a household's perceived utility from buying will increase over and above the gain due to price reduction. At the other extreme, if spillovers are solely due to perceived negative health externality of buyers on non-buyers, then increased purchase by neighbors would lower the utility of a household upon *not* buying via the health-route, but not affect it upon buying since the household is then protected anyway. These different aggregate welfare effects are both consistent with the same positive aggregate social interaction coefficient. This conclusion continues to hold even if eligibility for the subsidy is universal and there are no income effects or endogeneity concerns.

This feature is present in many other choice situations that economists routinely study. For example, merit-based school vouchers for attending a high-achieving school can potentially have a range of possible welfare effects. Aggregate welfare change could be negative if, for example, with high-ability children moving with the voucher the academic quality declines in the resource-poor schools more than the improvement in the selective school via peer effects. In the absence of such negative externalities, aggregate welfare could be positive due to the subsidy-led price decline for voucher users and any positive conforming effects that raise the utility of attending the high-achieving school when more high-ability children also do so. These contradictory welfare implications are compatible with the same positive coefficient on the social interaction term in an individual school choice model.

For standard discrete choice without spillover, Bhattacharya (2015) showed that the choice probability function itself contains *all* the information required for exact welfare analysis. For the special case of quasilinear random utility models with extreme value errors, the popular 'logsum' formula of Small and Rosen (1981) yields average welfare of policy interventions. These results fail to hold in a setting with spillovers because here one *cannot* set the utility from the outside option to zero – an innocuous normalization in standard discrete choice models – since this utility changes as the equilibrium choice-rate changes with the policy intervention. This is in contrast to binary choice *without* spillover, where utility from the outside option, i.e., non-purchase, does not change due to a price change of the inside good.

Nonetheless, under a standard, linear-index specification of utilities, one can calculate distribution-free bounds on average welfare, based solely on choice probability functions. The width of the bounds increases with (i) the extent of net social spillover, i.e. how much the (belief about) average neighborhood choice affects individual choice probabilities, and (ii) the difference in average peer-choice corresponding to realized equilibria before and after the price change. The index structure, which has been universal in the empirical literature on social interactions, leads to dimension reduction that helps identify spillovers effects. We therefore continue to use the index structure as it simplifies our expressions, and comes "for free", because social spillovers cannot in general be identified without such structure anyway. Under stronger and untestable restrictions on the nature of spillover, our bounds can shrink to a singleton, implying point-identification of welfare. Two such restrictions are (a) the effects of an increase in average peer-choice on individual utilities from buying and not buying are exactly equal in magnitude and opposite in sign, or (b) the effect of aggregate choice on either the purchase utility or the non-purchase utility is zero.

A separate identification problem arises when there are, in addition to social interaction, unobserved group-effects that are potentially correlated with observed individual covariates. We address this problem through a novel latent factor structure on the relevant variables, and developing a method of asymptotic analysis where the dimension of parameters, i.e. the group-effects whose magnitude may be unbounded, increases as the number of groups increase.

Empirical Illustration: We illustrate our theoretical results with an empirical example of a hypothetical, targeted public subsidy scheme for anti-malarial bednets. In particular, we use micro-data from a pricing experiment in rural Kenya (Dupas, 2014) to estimate an econometric model of demand for bednets, where spillovers can arise via different channels, including a preference for conformity and perceived negative externality arising from neighbors' use of a bednet. In this setting, we calculate predicted effects of hypothetical income-contingent subsidies on bednet demand and welfare. We perform these calculations by first accounting for social interactions, and then compare these results with what would be obtained if one had ignored these interactions. We find that allowing for (positive) interaction leads to a prediction of lower demand when means-tested eligibility is restricted to fewer households and higher demand when the eligibility criterion is more lenient, relative to ignoring interactions. To illustrate, consider a policymaker debating whether to expand eligibility for the subsidy from 40% to 60% of the population. In the presence of conforming effects, the increase in eligibility will spur more non-eligible to adopt, such that the total demand with spillovers are *larger* than without spillovers. Conversely, if eligibility was cut from 40% to 20%, the drop in adoption would be magnified by conforming effects, such that the total demand with spillovers are *lower* than without spillovers.[1] As for welfare, allowing for social interactions may lead to a welfare *loss* for ineligible households, in turn implying higher deadweight loss from the subsidy scheme, relative to estimates obtained ignoring social spillovers where welfare effects for ineligibles are zero by definition. The resulting net welfare effect, aggregated over both eligibles and ineligibles, admits a large range of possible values including both positive and negative ones, with associated large variation in the implied deadweight loss estimates, all of which are consistent with the *same* coefficient on the social interaction term in the choice probability function.

An implication of these results for applied work is that welfare analysis under spillovers effects requires knowledge of the different channels of spillovers separately, possibly via conducting a 'belief elicitation' survey where subjects are asked the reasons for their actions; knowledge of only the choice probability functions, inclusive of a social interaction term, is insufficient.

Plan of the Paper: The rest of the paper is organized as follows. Section 2 describes the set-up, Section 3 develops the tools for empirical welfare analysis of a price intervention in such models, and associated deadweight loss calculations. Section 4 specifies the stochastic environment and derives the convergence of equilibrium beliefs under I.I.D. unobservables. Section 5 establishes consistency of our estimator. Section 6, describes the context of our empirical application and the data; Section 7 describes the empirical results; Section 8 summarizes and concludes. Technical derivations, formal proofs and additional results are collected in an Appendix.

2 SET-UP

Consider a population of villages indexed by $v \in \{1, \ldots, \bar{v}\}$ and resident households in village v indexed by (v, h), with $h \in \{1, \ldots, N_v\}$. For the purpose of inference discussed later, we will think

[1] The intuition can be understood via a simple example. Suppose the true regression model is $y = \beta_0 + \beta_1 x + u$, where $\beta_1 > 0$. Suppose, to predict y at a value x_0 of x, we ignore the covariate and simply use $\bar{y}$ as the prediction. If $x_0 < \bar{x}$, then the naive prediction $\bar{y} = \beta_0 + \beta_1 \bar{x}$ will be larger than the true value $\beta_0 + \beta_1 x_0$, whereas if $x_0 > \bar{x}$, then the naive prediction $\bar{y}$ will be smaller than the true value $\beta_0 + \beta_1 x_0$.

of these households as a random sample drawn from an infinite superpopulation. The total number of households we observe is $N = \sum_{v=1}^{\bar{v}} N_v$. Each household faces a binary choice between buying one unit of an indivisible good (alternative 1) or not buying it (alternative 0). Its utilities from the two choices are given by $U_1(Y_{vh} - P_{vh}, \Pi_{vh}, \boldsymbol{\eta}_{vh})$ and $U_0(Y_{vh}, \Pi_{vh}, \boldsymbol{\eta}_{vh})$ where the variables Y_{vh}, P_{vh}, and $\boldsymbol{\eta}_{vh}$ denote respectively the income, price, and heterogeneity of household (v, h), and Π_{vh} is household (v, h)'s subjective belief of what fraction of households in her village would choose to buy. The variable $\boldsymbol{\eta}_{vh}$ is privately observed by household (v, h) but is unobserved by the econometrician and other households. The dependence of utilities on Π_{vh} captures social interactions. Below, we will specify how Π_{vh} is formed. Household (v, h)'s choice is described by

$$A_{vh} = 1\left\{U_1\left(Y_{vh} - P_{vh}, \Pi_{vh}, \boldsymbol{\eta}_{vh}\right) \geq U_0\left(Y_{vh}, \Pi_{vh}, \boldsymbol{\eta}_{vh}\right)\right\}, \tag{1}$$

where $1\{\cdot\}$ denotes the indicator function. In the mosquito-net example of our application, one can interpret U_1 and U_0 as expected utilities resulting from differential probabilities of contracting malaria from using and not using the net, respectively.

The utilities, U_1 and U_0, may also depend on other covariates of (v, h). For notational simplicity, we will occasionally write $W_{vh} = (Y_{vh}, P_{vh})^2$, and suppress other covariates for now; additional covariates are used in our empirical implementation in Section 7.

3 WELFARE ANALYSIS

We now lay out the empirical framework for welfare analysis of policy interventions under spillovers. We will assume spillovers are restricted to the village where households reside, hence welfare effects of a policy intervention can be analyzed village by village; so for economy of notation, we drop the (v, h) subscripts except when we account explicitly for village-effects during estimation. Also, we use the same notation π to denote both individual beliefs Π_{vh} entering individual utilities, and the unique equilibrium belief about village take-up rate entering the average demand function. The assumption of a constant (within village) π is justified via Proposition 1 and Proposition 2 in Section 4.1.

In the welfare results derived below, all probabilities and expectations – e.g., mean welfare loss – are calculated with respect to the *marginal* distribution of aggregate unobservables, denoted by $\boldsymbol{\eta}$. In this sense, they are analogous to 'average structural functions' (ASF), introduced by Blundell and Powell (2004). Later, when discussing estimation of the ASF, together with the implied pre- and post-intervention aggregate choice probabilities and average welfare in Section 5.4, we will allude to village-effects explicitly, and show how they are estimated and incorporated in demand and welfare predictions.

[2] All vectors are defined as row vectors.

Define $q_1(p,y,\pi)$ to be the *structural* probability (i.e. average structural function; ASF) of a household choosing option 1 (e.g., buying mosquito-net) when it faces a price of p, has income y and belief π:

$$q_1(p,y,\pi) = \int 1\{U_1(y-p,\pi,\boldsymbol{s}) > U_0(y,\pi,\boldsymbol{s})\}\, dF_{\boldsymbol{\eta}}(\boldsymbol{s}), \tag{2}$$

where $F_{\boldsymbol{\eta}}$ is the cumulative distribution function (CDF) of $\boldsymbol{\eta}$. This probability can be estimated via the conditional probability of purchase given covariates when household level unobservables are uncorrelated with the covariates, as will be assumed in our application. The reason for focusing on the ASF, rather than the purchase probability conditional on covariates is that ultimately, we will be interested in the marginal distribution of welfare in a village resulting from a *potential* price-intervention, e.g. a means-tested subsidy, which is counterfactual.[3]

Linear Index Structure: We now specify the forms of the utility functions. Given a moderate/small number of large peer groups (e.g., there are eleven large villages in our application dataset), it is not easy to consistently estimate the impact of the belief Π_{vh} on the choice probability function *nonparametrically* holding other regressors constant.[4] Accordingly, following Manski (1993), and Brock and Durlauf (2001, 2007), we assume a linear index structure with $\boldsymbol{\eta} = (\eta^0, \eta^1)$ viz. the utilities are given by

$$\begin{aligned} U_0(y,\pi,\boldsymbol{\eta}) &= \delta_0 + \beta_0 y + \alpha_0 \pi + \eta^0, \\ U_1(y-p,\pi,\boldsymbol{\eta}) &= \delta_1 + \beta_1(y-p) + \alpha_1 \pi + \eta^1, \end{aligned} \tag{3}$$

where we assume that $\beta_0 > 0$, $\beta_1 > 0$, i.e., non-satiation in numeraire, and β_1 need not equal β_0, i.e. income effects can be present.[5]

In our empirical setting of anti-malarial bednet (Insecticide-Treated Net; ITN, henceforth) adoption, there are multiple potential sources of interactions (i.e. $\alpha_1, \alpha_0 \neq 0$). The first is a pure

[3] Expressing our results in terms of ASFs help clarify that in general, the object of interest is one, whose identification and consistent estimation may require non-experimental methods if the data at hand are observational. Also, later in the paper, we will allude to village level unobservables i.e. $\eta_{vh} = \xi_v + u_{vh}$, where ξ_v is a village specific unobservable variable (introduced in Section 4). In that context, the object of interest will be the marginal distribution of welfare in each village; thus the relevant distribution to used to compute the ASF will be that of u_{vh} given village specific variables.

[4] This is because Π_{vh} is a constant within a village (as discussed in Section 4.1). In particular, the fixed point constraint, which is a notable feature of the social interaction model, does not help because of dimensionality problems. Indeed, in the fixed point condition: $\pi = \int q_1(p,y,\pi)\, dF_{P,Y}(p,y)$, where the joint CDF $F_{P,Y}(p,y)$ of (P,Y) is identified, the unknown function $q_1(p,y,\pi)$ has more arguments than the identified $F_{P,Y}(p,y)$.

[5] We can also allow for concave income effects by specifying, say,

$$\begin{aligned} U_0(y,\pi,\eta) &= \delta_0 + \beta_0 \ln y + \alpha_0 \pi + \eta^0, \\ U_1(y-p,\pi,\eta) &= \delta_1 + \beta_1 \ln(y-p) + \alpha_1 \pi + \eta^1, \end{aligned}$$

but we wish to keep the utility formulation as simple as possible to highlight the complications in welfare calculations even in the simplest linear utility specification.

preference for conforming; the second is increased awareness of the benefits of a bednet when more villagers use it; the third is the perceived health externality. The medical literature suggests that the *technological* health externality is positive, i.e. as more people are protected, the lower is the malaria burden, but the *perceived* health externality can be negative if households believe that other households' bednet use deflects mosquitoes to unprotected households, but ignore the fact that those deflected mosquitoes are less likely to carry the parasite. Indeed, the implications for adoption are different: under the positive health externality, one would expect free-riding, hence a negative effect of others' adoption on own adoption; under the negative health externality, the correlation would be positive.

In particular, let γ_p denote the conforming plus learning effect, and γ_H denote the health externality. Then it is reasonable to assume that $\alpha_1 \equiv \gamma_p \geq 0$, while $\alpha_0 \equiv \gamma_H - \gamma_p$ could be either negative or positive. It is natural that the conforming/learning/peer effect γ_p affects utilities from buying and non-buying symmetrically, i.e. if Π changes from 0 to π the resulting change in the utility (relative to when Π was 0) from buying and the one from not buying are symmetric and of opposite sign, as is also assumed in BD01, BD07. Further, if a household uses an ITN, then there is no health externality from the neighborhood adoption rate since the household is protected anyway,[6] but if it does not adopt, then there is a net health externality effect γ_H from neighborhood use, which makes the overall effect $\alpha_0 = \gamma_H - \gamma_p$ and in general, there is no exact relationship between $\alpha_0 = \gamma_H - \gamma_p$ and $\alpha_1 = \gamma_p$.[7] Accordingly, we first assume that the perceived *net health* externality is non-positive, and thus $\alpha_1 \geq 0 \geq \alpha_0$, and derive welfare results. In the next subsection, we present the results under the case $\alpha_1 \geq \alpha_0 \geq 0$. Note that the sign of $\alpha = \alpha_1 - \alpha_0$ is identified, and is positive in our data, which rules out $\alpha_0 > \alpha_1 \geq 0$. In the application, we present the bounds separately for $\alpha_1 \geq 0 \geq \alpha_0$ and $\alpha_1 \geq \alpha_0 \geq 0$, and then the union of these.

Given the linear index specification, the structural choice probability of buying at (p, y, π) is given by

$$q_1(p, y, \pi) = F(\underbrace{c_0}_{\delta_1 - \delta_0} + \underbrace{c_1}_{-\beta_1} p + \underbrace{c_2}_{\beta_1 - \beta_0} y + \underbrace{\alpha}_{\alpha_1 - \alpha_0} \pi), \tag{4}$$

where $F(\cdot)$ denotes the marginal distribution function of $\eta^0 - \eta^1$. It is known from Brock and Durlauf (2007) that the structural choice probabilities $F(c_0 + c_1 p + c_2 y + \alpha\pi)$ identify c_0, c_1, c_2 and α, i.e. $(\delta_1 - \delta_0)$, β_0, β_1 and $(\alpha_1 - \alpha_0) = 2\gamma_p - \gamma_H$, up to scale even without knowledge of the probability distribution of $\eta^0 - \eta^1$. In the application, we will consider two different estimates of

[6] We can allow for a smaller health externality, say $\gamma_h < \gamma_H$ when one adopts the bednet. But this does not change the fundamental point about the asymmetric effect of π on the utility from buying and from not buying. So we avoided adding this to save on notation.

[7] An analogous asymmetry is also likely in the school voucher example mentioned in the introduction if the voucher-led 'brain-drain' leads to utility gains and losses of different amounts, e.g., if better teaching resources in the high-achieving school substitute for – or complement – peer effects in a way that is not possible in the resource-poor local school.

the choice probabilities, first ignoring village-specific unobservables and using a standard probit, and then allowing for village-fixed unobservables using a variant of correlated random effects.

The distinct presence of α_1, α_0 makes the model different from standard demand models for binary choice. In the standard case, for the so-called "outside option", i.e. not buying, the utility is normalized to zero. In a social spillovers setting, this cannot be done because that utility depends on the aggregate purchase rate π. As we will see below, in welfare evaluations of a subsidy, α_1 and α_0 appear separately in the expressions for welfare-distributions, but cannot be separately identified from demand data, which can only identify $\alpha \equiv \alpha_1 - \alpha_0$. As a result, point-identification of welfare will in general not be possible. Below, we will consider some *untestable* special cases, under which one obtains point-identification, e.g., with $\alpha_1 \geq 0 \geq \alpha_0$, the interesting special cases are (i) $\alpha_1 = \alpha/2 = -\alpha_0$ (i.e. $\gamma_H = 0$: no health externality and symmetric spillover), which is considered in BD01 for social welfare analysis, (ii) $\alpha_1 = \alpha$, $\alpha_0 = 0$ (i.e. $\gamma_H = \gamma_p$: technological health externality dominates deflection channel and net health externality exactly offsets conforming effect), and (iii) $\alpha_1 = 0$, $\alpha_0 = -\alpha$ ($\gamma_p = 0$ and $\gamma_H = -\alpha$: no conforming effect and deflection channel dominates). Cases (ii) and (iii) will yield respectively the upper and lower bounds on welfare gain for the case $\alpha_1 \geq 0 \geq \alpha_0$. Analogously for the case $\alpha_1 \geq \alpha_0 \geq 0$.

Policy Intervention and Welfare Expressions: We start with a situation where the price of the product is p_0 and the value of π is π_0. Now suppose a price subsidy is introduced such that individuals with income less than a threshold τ become eligible to buy the product at price $p_1 < p_0$. This policy will alter the equilibrium adoption rate; suppose the new equilibrium adoption rate changes to π_1, where π_0 and π_1, solve the fixed point conditions:

$$\pi_0 = \int F\left(c_0 + c_1 p_0 + c_2 y + \alpha \pi_0\right) dF_Y\left(y\right), \tag{5}$$

$$\pi_1 = \int \left[1\left\{y \leq \tau\right\} F\left(c_0 + c_1 p_1 + c_2 y + \alpha \pi_1\right) + 1\left\{y > \tau\right\} F\left(c_0 + c_1 p_0 + c_2 y + \alpha \pi_1\right)\right] dF_Y\left(y\right), \tag{6}$$

where F_Y is the CDF of Y_{vh}, and F is defined in (4). Since the price coefficient $c_1 < 0$ and $p_1 > p_0$, therefore if $\alpha > 0$, we have that

$$\begin{aligned} & F\left(c_0 + c_1 p_0 + c_2 y + \alpha \pi\right) \\ & \leq \left[1\left\{y \leq \tau\right\} F\left(c_0 + c_1 p_1 + c_2 y + \alpha \pi\right) + 1\left\{y > \tau\right\} F\left(c_0 + c_1 p_0 + c_2 y + \alpha \pi\right)\right] \end{aligned}$$

for each π, and therefore, the integrand of (5) is smaller for every π than the integrand of (6). If the solutions to (5) and (6) are unique, then the value of π at which (5) holds must be smaller than the value of π where (6) holds. So we shall get $\pi_1 > \pi_0$. This is borne out in our application where sufficient conditions on α for a contraction are satisfied. Under multiple solutions, we can at least say that if $p_1 < p_0$, the smallest solution π_1 to (6) is greater than the smallest solution π_0 to (5).

For given values of π_0 and π_1, we now derive expressions for welfare resulting from the intervention. By "welfare" we mean the compensating variation (CV), viz. what hypothetical income compensation would restore the post-change indirect utility for an individual to its pre-change level. For a subsidy-*eligible* individual, for any potential value of π_1 corresponding to the new equilibrium, the individual compensating variation is the solution S to the equation

$$\max\left\{U_1\left(y+S-p_1,\pi_1,\boldsymbol{\eta}\right),U_0\left(y+S,\pi_1,\boldsymbol{\eta}\right)\right\}=\max\left\{U_1\left(y-p_0,\pi_0,\boldsymbol{\eta}\right),U_0\left(y,\pi_0,\boldsymbol{\eta}\right)\right\}, \quad (7)$$

whereas for a subsidy-*ineligible* individual, it is the solution S to

$$\max\left\{U_1\left(y+S-p_0,\pi_1,\boldsymbol{\eta}\right),U_0\left(y+S,\pi_1,\boldsymbol{\eta}\right)\right\}=\max\left\{U_1\left(y-p_0,\pi_0,\boldsymbol{\eta}\right),U_0\left(y,\pi_0,\boldsymbol{\eta}\right)\right\}. \quad (8)$$

Thus we interpret the CV as measuring utility changes via the value of hypothetical income compensation that would restore utilities to their initial level.[8] Now, since S depends on the unobservables $\boldsymbol{\eta}$, the same price change will produce a *distribution* of welfare effects across individuals; we are interested in calculating that distribution and its functionals such as mean welfare.

The welfare effect of the subsidy can be calculated as described below.

3.1 Welfare Calculation under $\alpha_1 \geq 0 > \alpha_0$

Recall that $\alpha_1 = \gamma_p \geq 0$, $\alpha_0 = \gamma_H - \gamma_p$; thus $\alpha_1 \geq 0 > \alpha_0$ corresponds to the case where either $\gamma_H < 0$, i.e., deflection effect dominates positive health effect in perception, or is positive but smaller than conforming/learning effect.

Welfare for Eligibles ($\alpha_1 \geq 0 > \alpha_0$): The CV for a subsidy-eligible household is given by the solution S to

$$\begin{aligned}&\max\left\{\delta_1+\beta_1\left(y+S-p_1\right)+\alpha_1\pi_1+\eta^1,\delta_0+\beta_0\left(y+S\right)+\alpha_0\pi_1+\eta^0\right\}\\&=\max\left\{\delta_1+\beta_1\left(y-p_0\right)+\alpha_1\pi_0+\eta^1,\delta_0+\beta_0 y+\alpha_0\pi_0+\eta^0\right\}. \end{aligned} \quad (9)$$

The resulting solution S depends on the unobservable heterogeneity η^0 and η^1 and hence we are interested in deriving its distribution and functionals thereof such as mean welfare. Calculating the welfare distribution requires us to compute the CDF of S, i.e. $\Pr\left(S \leq a\right)$ for various values of a (for given (p_0, π_0, p_1, π_1)). Let $f_{\eta^0-\eta^1}(\cdot)$ denote the marginal density function of $\eta^0 - \eta^1$. Then the expression for the CDF of welfare is as follows:

Theorem 1 *Suppose the linear index structure described above holds with* $\beta_0 > 0$, $\beta_1 > 0$ *and* $\alpha_1 \geq 0 \geq \alpha_0$ *with* $\alpha = \alpha_1 - \alpha_0$ *satisfying* $|\alpha| \sup_{e\in\mathbb{R}} f_{\eta^0-\eta^1}(e) < 1$. *Then* $\pi_1 > \pi_0$, *and the*

[8] Note that we do not take account of peer effects of this hypothetical income compensation, which might be an alternative way to define the CV.

distribution of compensating variation for the eligibles, $S = S^{\text{Elig}}$, is given by

$$\Pr\left(S^{\text{Elig}} \le a\right)$$
$$= \begin{cases} 0, & \text{if } a < p_1 - p_0 - \frac{\alpha_1}{\beta_1}(\pi_1 - \pi_0), \\ q_1\left(p_1 - a, y, \pi_0 + \frac{\alpha_1}{\alpha}(\pi_1 - \pi_0)\right), & \text{if } p_1 - p_0 - \frac{\alpha_1}{\beta_1}(\pi_1 - \pi_0) \le a < \frac{\alpha - \alpha_1}{\beta_0}(\pi_1 - \pi_0), \\ 1, & \text{if } a \ge \frac{\alpha - \alpha_1}{\beta_0}(\pi_1 - \pi_0). \end{cases} \tag{10}$$

The proof is provided in Appendix A.1. The condition $|\alpha| \sup_{e \in \mathbb{R}} f_{\eta^0 - \eta^1}(e) < 1$ essentially says that the social interaction parameter is not too large in magnitude, so that (5) and (6) have unique solutions in π_0 and π_1 respectively, whence by the argument following (5) and (6), we have that $\pi_1 > \pi_0$.

Now, note that in the intermediate case in (10), where $a \in [p_1 - p_0 - \frac{\alpha_1}{\beta_1}(\pi_1 - \pi_0), \frac{\alpha_0}{\beta_0}(\pi_0 - \pi_1)]$, $\Pr(S \le a)$ equals

$$q_1(p_1 - a, y, \pi_0 + \frac{\alpha_1}{\alpha}(\pi_1 - \pi_0)) = F(c_0 + \alpha_1(\pi_1 - \pi_0) + c_1(p_1 - a) + c_2 y + \alpha\pi_0). \tag{11}$$

In (11), the intercept $c_0 = \delta_1 - \delta_0$, the slopes $c_1 = -\beta_1, c_2 = \beta_1 - \beta_0$ and $\alpha = \alpha_1 - \alpha_0$ are all identified from conditional choice probabilities; *however α_1 is not identified* and therefore (11) is not point-identified from the structural choice probabilities. However, since $\alpha_1 \in [0, \alpha]$, for each feasible value of $\alpha_1 \in [0, \alpha]$, we can compute a corresponding value of (11), giving us bounds on the welfare distribution.

Note also that the thresholds of a at which the CDF expression changes are also not point-identified for the same reason. However, since $\pi_1 - \pi_0 > 0$ and $\beta_0 > 0$, $\beta_1 > 0$, the interval

$$p_1 - p_0 - \frac{\alpha_1}{\beta_1}(\pi_1 - \pi_0) \le a < \frac{\alpha_0}{\beta_0}(\pi_0 - \pi_1)$$

will translate to the left as α_1 varies from 0 to α.

Remark 1 *Note that the above theorem continues to hold even if the subsidy is universal; we have not used the means-tested nature of the subsidy to derive the result.*

Corollary 1 (Mean Welfare) *From (10), the mean welfare for the eligible is given by*

$$E[S^{\text{Elig}}] = \underbrace{-\int_{p_1 - p_0 - \frac{\alpha_1}{\beta_1}(\pi_1 - \pi_0)}^{0} q_1\left(p_1 - a, y, \pi_0 + \frac{\alpha_1}{\alpha}(\pi_1 - \pi_0)\right) da}_{\text{welfare gain}}$$
$$+ \underbrace{\int_0^{\frac{\alpha - \alpha_1}{\beta_0}(\pi_1 - \pi_0)} \left[1 - q_1\left(p_1 - a, y, \pi_0 + \frac{\alpha_1}{\alpha}(\pi_1 - \pi_0)\right)\right] da}_{\text{welfare loss}}, \tag{12}$$

where the following formula for a random variable X that has finite mean and the CDF, F_X, is used: $E[X] = \int_0^\infty [1 - F_X(x)]dx - \int_{-\infty}^0 F_X(x)\,dx$.

The width of the bounds on (10) and (12), obtained by varying α_1 over $[0, \alpha]$, depends on the extent to which $q_1(\cdot, \cdot, \pi)$ is affected by π, i.e. the extent of social spillover, and also the difference in the realized values π_1 and π_0. For our single-index model, the fixed point restrictions imply that these counterfactual π_1 and π_0 depend on α_1 and α_0 only via $\alpha = \alpha_1 - \alpha_0$ (cf. (5) and (6) above) which is point-identified; thus every potential value of counterfactual demand is point-identified. But given any feasible value of π_1 and π_0, the welfare (12) is not point-identified in general, since α_1 is unknown.

However, given α, the welfare gain in expression (12) is increasing in α_1; i.e., the welfare gain is largest in absolute value when $\alpha_1 = \alpha$ and $\alpha_0 = 0$, and the smallest when $\alpha_1 = 0$ and $\alpha_0 = -\alpha$; conversely for welfare loss. Intuitively, if there is no negative externality from increased π on non-purchasers, then they do not suffer any welfare loss, but purchasers have a welfare gain from both lower price and higher π. Conversely, if all the spillovers are negative, then purchasers still receive a welfare gain via price reduction, but non-purchasers suffer welfare loss due to increased π. Also, note that under quasilinear utilities (i.e., utilities with $\beta_0 = \beta_1$), where income effects are absent, the y drops out of the above expressions, but the same identification problem remains, since α_1 does not disappear. Changing variables $p = p_1 - a$, one can rewrite (12) as

$$E[S^{\text{Elig}}] = \underbrace{-\int_{p_1}^{p_0+\frac{\alpha_1}{\beta_1}(\pi_1-\pi_0)} q_1\left(p, y, \pi_0 + \frac{\alpha_1}{\alpha}(\pi_1 - \pi_0)\right) dp}_{\text{welfare gain}}$$
$$+ \underbrace{\int_{p_1+\frac{\alpha_1-\alpha}{\beta_0}(\pi_1-\pi_0)}^{p_1} \left[1 - q_1\left(p, y, \pi_0 + \frac{\alpha_1}{\alpha}(\pi_1 - \pi_0)\right)\right] dp.}_{\text{welfare Loss}} \tag{13}$$

Note that if $\alpha_1 = 0$, then the first term is the usual consumer surplus capturing the effect of price reduction on consumer welfare; for a positive α_1, the term $\frac{\alpha_1}{\beta_1}(\pi_1 - \pi_0)$ yields the additional effect arising via the conforming channel. Also, if $\alpha_1 = 0$, then the second term, i.e. the welfare loss from not buying, is the largest (given α): this corresponds to the case where all of α is due to the negative externality.

The second term in (13), which represents welfare change caused solely via spillovers and no price change, is still expressed as an integral with respect to price. This is a consequence of the index structure which enables us to express this welfare loss in terms of foregone utility from an equivalent price change.

Special Cases: For the special case of symmetric interactions considered in BD01, where their social welfare is calculated with $\alpha_1 = -\alpha_0$ in (3) (e.g., if $\gamma_H = 0$, i.e. there is no health externality in the health-good example), we have $\frac{\alpha_1}{\alpha} = \frac{-\alpha_0}{-2\alpha_0} = \frac{1}{2}$, and from (13) mean welfare equals:

$$\underbrace{-\int_{p_1}^{p_0+\frac{\alpha}{2\beta_1}(\pi_1-\pi_0)} q_1\left(p, y, \frac{1}{2}(\pi_1 + \pi_0)\right) dp}_{\text{welfare gain}} + \underbrace{\int_{p_1-\frac{\alpha}{2\beta_0}(\pi_1-\pi_0)}^{p_1} \left[1 - q_1\left(p, y, \frac{1}{2}(\pi_1 + \pi_0)\right)\right] dp.}_{\text{welfare loss}} \tag{14}$$

If $\alpha_0 = 0$, and $\alpha = \alpha_1$, i.e. all spillovers are via conforming, mean welfare is given by

$$\underbrace{-\int_{p_1}^{p_0+\frac{\alpha}{\beta_1}(\pi_1-\pi_0)} q_1(p,y,\pi_1)\,dp}_{\text{welfare gain}}; \tag{15}$$

and if, on the other hand, any spillovers are due to perceived health risk, i.e. $\alpha = -\alpha_0$ and $\alpha_1 = 0$, then mean welfare is given by

$$\underbrace{-\int_{p_1}^{p_0} q_1(p,y,\pi_0)\,dp}_{\text{welfare gain}} + \underbrace{\int_{p_1-\frac{\alpha}{\beta_0}(\pi_1-\pi_0)}^{p_1} [1-q_1(p,y,\pi_0)]\,dp}_{\text{welfare loss}}. \tag{16}$$

Expressions (15) and (16) correspond to the upper and lower bounds, respectively, of the overall welfare gain for eligibles.[9]

Welfare for Ineligibles ($\alpha_1 \geq 0 > \alpha_0$): Welfare change for *ineligibles* is measured by the CV defined as the solution S to the equation:

$$\begin{aligned} &\max\left\{\delta_1+\beta_1(y+S-p_0)+\alpha_1\pi_1+\eta^1, \delta_0+\beta_0(y+S)+\alpha_0\pi_1+\eta^0\right\} \\ &= \max\left\{\delta_1+\beta_1(y-p_0)+\alpha_1\pi_0+\eta^1, \delta_0+\beta_0 y+\alpha_0\pi_0+\eta^0\right\}, \end{aligned} \tag{17}$$

which is simply (9) with p_1 replaced by p_0. Therefore, the mean CV is simply (9) with p_1 replaced by p_0.

Corollary 2 *Suppose the linear index structure described above holds with $\beta_0 > 0$, $\beta_1 > 0$, and $\alpha_1 \geq 0 \geq \alpha_0$. Then for each $\alpha_1 \in [0, \alpha]$, the mean welfare for the ineligible, $S = S^{\text{Inelig}}$, is given by*

$$\begin{aligned} E[S^{\text{Inelig}}] = &-\int_{p_0}^{p_0+\frac{\alpha_1}{\beta_1}(\pi_1-\pi_0)} q_1\left(p,y,\pi_0+\frac{\alpha_1}{\alpha}(\pi_1-\pi_0)\right)dp \\ &+\int_{p_0+\frac{\alpha_1-\alpha}{\beta_0}(\pi_1-\pi_0)}^{p_0} \left[1-q_1\left(p,y,\pi_0+\frac{\alpha_1}{\alpha}(\pi_1-\pi_0)\right)\right]dp. \end{aligned} \tag{18}$$

For *ineligibles*, all of the welfare effects come from spillovers, since they experience no price change. In particular, for ineligibles who buy, there is a welfare gain from positive spillovers due to a higher π. For ineligibles who do not buy, there is, however, a potential welfare loss due to increased π. This is why the CV distribution has the support that includes both positive and negative values. The first term in (18) captures the welfare gain resulting from a positive α_1 and higher π; this term would be zero if $\alpha_1 = 0$. The second term in (18) captures the welfare loss also resulting from higher π; this loss would be zero if there are no negative impacts, i.e. $\alpha_0 = 0$. Of course, both

[9] Gautam (2018) obtained point-identified estimates of welfare in parametric discrete choice models with social interactions, purportedly using Dagsvik and Karlstrom's (2005) results for the setting without spillover. Gautam's paper contains no explicit expression for average welfare, but we conjecture that her derivation had implicitly assumed one of the normalizations (14), (15) or (16) under which average welfare is point-identified.

would be zero if $\alpha = 0 = \alpha_1 = \alpha_0$, reflecting the fact that welfare effect on ineligibles would be zero if there is no spillover.

In the three special cases where we have point-identification, viz. (i) $\alpha_1 = -\alpha_0 = \frac{\alpha}{2}$; (ii) $\alpha = \alpha_1$, $\alpha_0 = 0$; and (iii) $\alpha = -\alpha_0$, $\alpha_1 = 0$, mean CV (18) reduces respectively to:

$$\text{(i)}\quad \underbrace{-\int_{p_0}^{p_0+\frac{\alpha}{2\beta_1}(\pi_1-\pi_0)} q_1\left(p, y, \frac{\pi_0+\pi_1}{2}\right) dp}_{\text{welfare gain}} + \underbrace{\int_{p_0-\frac{\alpha}{2\beta_0}(\pi_1-\pi_0)}^{p_0} \left[1 - q_1\left(p, y, \frac{\pi_0+\pi_1}{2}\right)\right] dp}_{\text{welfare loss}};$$

$$\text{(ii)}\quad \underbrace{-\int_{p_0}^{p_0+\frac{\alpha}{\beta_1}(\pi_1-\pi_0)} q_1\left(p, y, \pi_1\right) dp}_{\text{welfare gain}}; \tag{19}$$

$$\text{(iii)}\quad \underbrace{\int_{p_0-\frac{\alpha}{\beta_0}(\pi_1-\pi_0)}^{p_0} \left[1 - q_1\left(p, y, \pi_0\right)\right] dp}_{\text{welfare loss}}. \tag{20}$$

Expressions (19) and (20) correspond to the upper and lower bounds, respectively, of the overall welfare gain for ineligibles, and therefore, the overall bounds generically contain both positive and negative values, since $\alpha \neq 0$.

Deadweight Loss ($\alpha_1 \geq 0 > \alpha_0$): The mean deadweight loss (DWL) can be calculated as the expected subsidy spending less the net welfare gain:

$$\begin{aligned} DWL(y) = {} & \underbrace{1\{y \leq \tau\} \times (p_0 - p_1)\, q_1(p_1, y, \pi_1)}_{\text{subsidy spending}} \\ & \underbrace{-1\{y \leq \tau\} \times \left(\begin{array}{l} \displaystyle\int_{p_1-p_0-\frac{\alpha_1}{\beta_1}(\pi_1-\pi_0)}^{0} q_1\left(p_1 - a, y, \pi_0 + \frac{\alpha_1}{\alpha}(\pi_1 - \pi_0)\right) da \\ \displaystyle -\int_{0}^{\frac{\alpha-\alpha_1}{\beta_0}(\pi_1-\pi_0)} \left[1 - q_1\left(p_1 - a, y, \pi_0 + \frac{\alpha_1}{\alpha}(\pi_1 - \pi_0)\right)\right] da \end{array} \right)}_{\text{welfare gain of eligibles}} \\ & \underbrace{-1\{y > \tau\} \times \left[\begin{array}{l} \displaystyle\int_{p_0}^{p_0+\frac{\alpha_1}{\beta_1}(\pi_1-\pi_0)} q_1\left(p, y, \pi_0 + \frac{\alpha_1}{\alpha}(\pi_1 - \pi_0)\right) dp \\ \displaystyle -\int_{p_0+\frac{\alpha_0}{\beta_0}(\pi_1-\pi_0)}^{p_0} \left[1 - q_1\left(p, y, \pi_0 + \frac{\alpha_1}{\alpha}(\pi_1 - \pi_0)\right)\right] dp \end{array} \right]}_{\text{welfare gain of ineligibles}}.. \end{aligned}$$

The bounds on α_1 then translate into bounds for mean DWL. In particular, if $\alpha_0 = 0$ (so that $\alpha = \alpha_1$), then

$$\begin{aligned} & DWL(y) = 1\{y \leq \tau\} \times (p_0 - p_1)\, q_1(p_1, y, \pi_1) \\ & - 1\{y \leq \tau\} \times \int_{p_1}^{p_0+\frac{\alpha}{\beta_1}(\pi_1-\pi_0)} q_1(p, y, \pi_1)\, dp - 1\{y > \tau\} \times \int_{p_0}^{p_0+\frac{\alpha}{\beta_1}(\pi_1-\pi_0)} q_1(p, y, \pi_1)\, dp. \end{aligned}$$

Therefore, if $\frac{\alpha}{\beta_1}(\pi_1 - \pi_0)$ is sufficiently large, then the mean DWL will be *negative*, i.e. the subsidy will *increase* economic efficiency under positive spillover, as in the standard textbook case. This happens because there is no subsidy expenditure on ineligibles, and yet those ineligibles who buy enjoy a subsidy-induced welfare gain due to positive spillover. Subsidy-eligibles receive an additional welfare gain via positive spillover, over and above the welfare-gain due to reduced price, and it is only the latter that is financed by the subsidy expenditure. In general, the deadweight loss will be lower (more negative) when (i) the positive spillovers (α_1) is larger, (ii) the change in equilibrium adoption ($\pi_1 - \pi_0$) due to the subsidy is greater, and (iii) the price elasticity of demand ($-\beta_1$) is lower – the last effect lowers deadweight loss simply by reducing the substitution effect, even in absence of spillover.

3.2 Mean Welfare under $\alpha_1 \geq \alpha_0 \geq 0$

Recall that $\alpha_1 = \gamma_p \geq 0$, $\alpha_0 = \gamma_H - \gamma_p$; in our application, it holds that $\alpha = \alpha_1 - \alpha_0 > 0$; thus $\alpha_1 > \alpha_0 \geq 0$ corresponds to the case where $\gamma_H > 0$ i.e. insecticide effect dominates deflection effect, is also larger than conforming/learning but less than twice the conforming effect. Note that under this assumption, we must also have $\alpha \leq \alpha_1$.

Welfare for Eligibles ($\alpha_1 \geq \alpha_0 \geq 0$): For subsidy-eligibles, the mean welfare (for given (p_0, π_0, p_1, π_1)) is presented in the following theorem:

Theorem 2 *Suppose the linear index structure described above holds with $\beta_1 \geq \beta_0 > 0$, and $\alpha_1 \geq \alpha_0 \geq 0$. Let $\beta = \beta_1 - \beta_0$ and $\alpha = \alpha_1 - \alpha_0$, which are estimable from the choice probability function, and define*

$$C_1(\alpha_1) := -\int_{p_1 - \frac{\alpha - \alpha_1}{\beta_0}(\pi_1 - \pi_0)}^{p_0 + \frac{\alpha_1}{\beta_1}(\pi_1 - \pi_0)} q_1\left(p, y, \pi_0 + \frac{\alpha_1}{\alpha}(\pi_1 - \pi_0)\right) dp,$$

$$C_2(\alpha_1) := -\int_{p_0 + \frac{\alpha - \alpha_1}{\beta_0}(\pi_1 - \pi_0)}^{p_1 - \frac{\alpha_1}{\beta_1}(\pi_1 - \pi_0)} \left[1 - q_1\left(p, y + p - p_0, \pi_1 - \frac{\alpha_1}{\alpha}(\pi_1 - \pi_0)\right)\right] dp.$$

Then mean welfare for the eligible is given by

$$E\left[S^{\text{Elig}}\right] = \begin{cases} C_1(\alpha_1), & \text{if } \alpha \leq \alpha_1 \leq \frac{\beta_1}{\beta}(\beta_1 - \beta)\frac{p_0 - p_1}{\pi_1 - \pi_0} + \frac{\beta_1}{\beta}\alpha, \\ C_2(\alpha_1), & \text{if } \alpha_1 > \frac{\beta_1}{\beta}(\beta_1 - \beta)\frac{p_0 - p_1}{\pi_1 - \pi_0} + \frac{\beta_1}{\beta}\alpha. \end{cases}$$

The proof is provided in Appendix A.1. Given that α_1 is unknown, this result implies the lower and upper bounds of the mean welfare:

$$LB^{\text{Elig}}_{\alpha_1 \geq \alpha_0 \geq 0} = \min\left\{ \inf_{\alpha_1 \in \left[\alpha, \frac{\beta_1}{\beta}(\beta_1 - \beta)\frac{p_0 - p_1}{\pi_1 - \pi_0} + \frac{\beta_1}{\beta}\alpha\right]} C_1(\alpha_1), \inf_{\alpha_1 \in \left[\frac{\beta_1}{\beta}(\beta_1 - \beta)\frac{p_0 - p_1}{\pi_1 - \pi_0} + \frac{\beta_1}{\beta}\alpha, \infty\right)} C_2(\alpha_1) \right\}, \quad (21)$$

$$UB^{\text{Elig}}_{\alpha_1 \geq \alpha_0 \geq 0} = \max\left\{ \sup_{\alpha_1 \in \left[\alpha, \frac{\beta_1}{\beta}(\beta_1 - \beta)\frac{p_0 - p_1}{\pi_1 - \pi_0} + \frac{\beta_1}{\beta}\alpha\right]} C_1(\alpha_1), \sup_{\alpha_1 \in \left[\frac{\beta_1}{\beta}(\beta_1 - \beta)\frac{p_0 - p_1}{\pi_1 - \pi_0} + \frac{\beta_1}{\beta}\alpha, \infty\right)} C_2(\alpha_1) \right\}. \quad (22)$$

Thus, allowing for both $\alpha_1 \geq \alpha_0 \geq 0$ and $\alpha_1 \geq 0 \geq \alpha_0$ yields the wider bounds on mean welfare for the eligible to:

$$LB^{\text{Elig}} = \min\left\{LB^{\text{Elig}}_{\alpha_1\geq\alpha_0\geq 0},\ LB^{\text{Elig}}_{\alpha_1\geq 0\geq\alpha_0}\right\}, \tag{23}$$

$$UB^{\text{Elig}} = \max\left\{UB^{\text{Elig}}_{\alpha_1\geq\alpha_0\geq 0},\ UB^{\text{Elig}}_{\alpha_1\geq 0\geq\alpha_0}\right\}. \tag{24}$$

where $LB^{\text{Elig}}_{\alpha_1\geq 0\geq\alpha_0}$ and $UB^{\text{Elig}}_{\alpha_1\geq 0\geq\alpha_0}$ are defined as expressions (20) and (19), respectively. Since we expect $\beta_1 > \beta_0$ (also borne out by the empirical results), $C_2(\alpha_1)$ will tend to $-\infty$ as $\alpha_1, \alpha_0 \to \infty$. Therefore, as $\alpha_1, \alpha_0 \to \infty$, the integrand in $C_2(\alpha_1)$ will tend to 1 and $LB^{\text{Elig}}_{\alpha_1\geq\alpha_0\geq 0}$ in (21) will tend to $-\infty$ whereas the $UB^{\text{Elig}}_{\alpha_1\geq\alpha_0\geq 0}$ in (22) will remain bounded. Therefore, the lower bound on welfare gain will be finite and the upper bound infinite under $\alpha_1 \geq \alpha_0 \geq 0$.

Welfare for Ineligibles ($\alpha_1 \geq \alpha_0 \geq 0$): For subsidy-ineligibles, the mean welfare is obtained simply by replacing p_1 by p_0 in the expressions for eligibles:

Corollary 3 *Suppose the linear index structure described above holds with $\beta_1 \geq \beta_0 > 0$, and $\alpha_1 \geq \alpha_0 \geq 0$. Define*

$$D_1(a_1) := -\int_{p_0+\frac{\alpha_1-\alpha}{\beta_0}(\pi_1-\pi_0)}^{p_0+\frac{\alpha_1}{\beta_1}(\pi_1-\pi_0)} q_1\left(p, y, \pi_0 + \frac{\alpha_1}{\alpha}(\pi_1-\pi_0)\right) dp,$$

$$D_2(\alpha_1) := -\int_{p_0+\frac{\alpha-\alpha_1}{\beta_0}(\pi_1-\pi_0)}^{p_0-\frac{\alpha_1}{\beta_1}(\pi_1-\pi_0)} \left[1 - q_1\left(p, y+p-p_0, \pi_1 - \frac{\alpha_1}{\alpha}(\pi_1-\pi_0)\right)\right] dp.$$

Then, the mean welfare for the ineligible is given by

$$E\left[S^{\text{Inelig}}\right] = \begin{cases} D_1(\alpha_1), & \text{if } \alpha \leq \alpha_1 \leq \frac{\beta_1}{\beta_1-\beta_0}\alpha, \\ D_2(\alpha_1), & \text{if } \frac{\beta_1}{\beta_1-\beta_0}\alpha < \alpha_1 < \infty. \end{cases}$$

From the above results, it follows that allowing for $\alpha_1 > \alpha_0 \geq 0$ in addition to the possibility $\alpha_1 \geq \alpha_0 \geq 0$ widens the overall bounds for mean welfare of the ineligible from (19) and (20) to

$$LB^{\text{Inelig}} = \min\left\{\int_{p_0-\frac{\alpha}{\beta_0}(\pi_1-\pi_0)}^{p_0} \{1 - q_1(p,y,\pi_0)\}\, dp,\ \inf_{\alpha_1>\frac{\beta_1}{\beta_1-\beta_0}\alpha} D_2(\alpha_1),\ \inf_{\alpha\leq\alpha_1\leq\frac{\beta_1}{\beta_1-\beta_0}\alpha} D_1(\alpha_1)\right\}, \tag{25}$$

$$UB^{\text{Inelig}} = \max\left\{-\int_{p_0}^{p_0+\frac{\alpha}{\beta_1}(\pi_1-\pi_0)} q_1(p,y,\pi_1)\, dp,\ \sup_{\alpha_1>\frac{\beta_1}{\beta_1-\beta_0}\alpha} D_2(\alpha_1),\ \sup_{\alpha\leq\alpha_1\leq\frac{\beta_1}{\beta_1-\beta_0}\alpha} D_1(\alpha_1)\right\}. \tag{26}$$

The deadweight loss expressions are analogous to those for the case with $\alpha_1 \geq 0 \geq \alpha_0$ and not repeated here.

4 STOCHASTIC ENVIRONMENT AND EQUILIBRIUM BELIEFS

Incomplete-Information Setting: In this section, we formulate interactions of households as an incomplete-information Bayesian game, whose stochastic structure will be laid out below. In each village v, each of the N_v households is provided the opportunity to buy the product at a researcher-specified price P_{vh} randomly varied across households. They have incomplete information in that each household (v, h) knows her own variables $(A_{vh}, W_{vh}, \boldsymbol{\eta}_{vh})$ but does not know the values of all the variables $W_{\tilde{v}k}, \boldsymbol{\eta}_{\tilde{v}k}, A_{\tilde{v}k}$ for every household $k \neq h$ selected in the experiment.

We assume households have 'consistent beliefs' in accordance with the standard Bayes-Nash setting, i.e., each (v, h)'s belief is formed as

$$\Pi_{vh} = \frac{1}{N_v - 1} \sum\nolimits_{1 \leq k \leq N_v;\ k \neq h} E[A_{vk} | \mathcal{I}_{vh}], \tag{27}$$

where A_{vk} is given in (1) and $E[\cdot\ | \mathcal{I}_{vh}]$ is the conditional expectation computed through the probability law that governs all the relevant variables given (v, h)'s information set $\mathcal{I}_{vh}$ that includes $(W_{vh}, \boldsymbol{\eta}_{vh})$. The explicit form of (27) *in equilibrium* is investigated in the next subsection.

Each household (v, h) is solely concerned with behavior of other households in the same village v. Thus the econometrician observes $\bar{v}$ games ($\bar{v} = 11$ in our application), each with 'many' households. To formalize our model as a Bayesian game, given the form of (27), U_1 and U_0 are to be interpreted as expected utilities. This is possible when the underlying von Neumann-Morgenstern utility indices u_1 and u_0 satisfy

$$U_1 (Y_{vh} - P_{vh}, \Pi_{vh}, \boldsymbol{\eta}_{vh}) = E[u_1(Y_{vh} - P_{vh}, \tfrac{1}{N_v - 1} \sum\nolimits_{1 \leq k \leq N_v;\ k \neq h} A_{vk}, \boldsymbol{\eta}_{vh}) | \mathcal{I}_{vh}],$$

i.e., u_1 is linear in the second argument; U_0 and u_0 satisfy an analogous relationship. This will hold in particular when utilities have a linear index structure as in Manski (1993) and Brock and Durlauf (2001, 2007). We have already presented our linear specifications of U_1 and U_0 in (3), but these are further elaborated below in this section and in Section 5.

Unobserved Heterogeneity: We assume that unobserved heterogeneity $\{\boldsymbol{\eta}_{vh}\}_{v=1}^{N_v}$ $(v = 1, \ldots \bar{v})$ takes the following form:

$$\boldsymbol{\eta}_{vh} = \boldsymbol{\xi}_v + \boldsymbol{u}_{vh}, \tag{28}$$

where $\boldsymbol{\xi}_v$ stands for a village-specific vector of variables that are common to all members in the vth village and $\boldsymbol{u}_{vh}$ represents an individual specific variable. Let (d_v, e_v) be an underlying vector of village-specific variables such that d_v is a common factor affecting both the unobservable $\boldsymbol{\xi}_v$ and the observable covariates W_{vh}, and e_v affects only $\boldsymbol{\xi}_v$ with $\boldsymbol{\xi}_v$ fully determined by (d_v, e_v), i.e., $\boldsymbol{\xi}_v = \boldsymbol{\xi}(d_v, e_v)$.[10] Each household in village v is assumed to know (d_v, e_v), the functional

[10] The need to separate d_v and e_v will become clear below in the context of identification of model parameters in presence of unobserved group-effects.

form $\boldsymbol{\xi}(\cdot)$, and thus $\boldsymbol{\xi}_v$, while $\boldsymbol{u}_{vh}$ is a purely private variable known only to individual (v, h). None of $\{(d_v, e_v)\}$, $\{\boldsymbol{\xi}_v\}$, and $\{\boldsymbol{u}_{vh}\}$ is observable to the econometrician. Denote household (v, h)'s information set by

$$\mathcal{I}_{vh} = (W_{vh}, \boldsymbol{u}_{vh}, d_v, e_v). \tag{29}$$

We now impose the following conditions on the probabilistic law for these variables:

C1 $\{(W_{vh}, \boldsymbol{u}_{vh}, d_v, e_v)\}_{h=1}^{N_v}$, $v = 1, \ldots, \bar{v}$, are independent across v.

Assumption **C1** says that variables in village v are independent of those in village $\tilde{v}(\neq v)$.

C2 (i) For each v, the sequence $\{(W_{vh}, \boldsymbol{u}_{vh})\}_{h=1}^{N_v}$ is I.I.D. conditionally on (d_v, e_v). (ii) $\{\boldsymbol{u}_{vh}\}_{h=1}^{N_v}$ is independent of $\{W_{vh}\}_{h=1}^{N_v}$ conditionally on (d_v, e_v).

The conditional I.I.D.-ness imposed in **C2 (i)** leads to equi-dependence within each village, i.e., $\text{Cov}\left[\boldsymbol{\eta}_{vh}, \boldsymbol{\eta}_{vk}\right] = \text{Cov}\left[\boldsymbol{\eta}_{v\tilde{h}}, \boldsymbol{\eta}_{v\tilde{k}}\right] (\neq 0)$ for any $h \neq k$ and $\tilde{h} \neq \tilde{k}$. Further, each household (v, h)'s unobservable $\boldsymbol{u}_{vh}$ is not useful for predicting another household (v, k)'s variables and behavior, and therefore her belief Π_{vh} (27) is reduced to the average of the *unconditional* expectations (as formally shown in Proposition 1) below. This condition rules out spatial correlation in unobservables which, if present, would complicate the analysis in a non-trivial way by making a household's belief a function of its privately known variables.

C2 (ii) is the exogeneity condition. This allows for identification and consistent estimation of model parameters. In the context of the field experiment in our empirical exercise, this exogeneity condition can be interpreted as saying that realization of unobserved heterogeneity is independent of how researchers have selected the sample. Note that the exogeneity condition is *conditional* on (d_v, e_v), and it does *not* exclude correlation of $\boldsymbol{u}_{vh}$ and $W_{vh} = (P_{vh}, Y_{vh})$ in the *unconditional* sense. In our application, prices P_{vh} are randomly assigned to individuals by researchers and thus P_{vh} and $\boldsymbol{u}_{vh}$ are independent both unconditionally and conditionally.

Note that under the (28) introduced later, we compute the ASF (2) using the marginal distribution of $\boldsymbol{u}_{vh}$ conditionally on $\boldsymbol{\xi}_v$ in later sections (see also Footnote 3).

4.1 Equilibrium Beliefs

We now investigate the forms of households' beliefs defined in (27). We show that under **C2**, the high-level assumption in BD01 that beliefs, corresponding to our Π_{vh}, are constant and symmetric across all households in the same village can be formalized in our incomplete-information game setting via the specification of a Bayes-Nash equilibrium.

Proposition 1 *Suppose that Conditions **C1** and **C2** are common knowledge in the Bayesian game described above. Then, for any $k \neq h$ in village v with (d_v, e_v),*

$$E[A_{vk}|\mathcal{I}_{vh}] = E[A_{vk}|d_v, e_v],$$

where the information set $\mathcal{I}_{vh}$ is defined in (29).

The proof of Proposition 1 is provided in Appendix A.2. Note that this proposition does not utilize any *equilibrium* condition. It simply confirms, formally, the intuitive statement that (v,h)'s own variables are not useful to predict other (v,k)'s behavior A_{vk}. Given this result, we can write the belief Π_{vh} (defined in (27)) as

$$\Pi_{vh} = \bar{\Pi}_{vh}, \tag{30}$$

where

$$\bar{\Pi}_{vh} = \bar{\Pi}_{vh}(d_v, e_v) := \tfrac{1}{N_v - 1}\sum\nolimits_{1 \leq k \leq N_v;\ k \neq h} E[A_{vk}|d_v, e_v],$$

and $\bar{\Pi}_{vh}$ is a function of (d_v, e_v) and independent of (v,h)-specific variables, $(W_{vh}, \boldsymbol{u}_{vh})$; for notational simplicity, we suppress the dependence of $\bar{\Pi}_{vh}$ on (d_v, e_v) from now on.

Beliefs *in equilibrium* solve the system of N_v equations:

$$\bar{\Pi}_{vh} = \tfrac{1}{N_v - 1}\sum\nolimits_{1 \leq k \leq N_v;\ k \neq h} E_v \left[1 \left\{ \begin{array}{c} U_1(Y_{vk} - P_{vk}, \bar{\Pi}_{vk}, \boldsymbol{\eta}_{vk}) \\ \geq U_0(y, \bar{\Pi}_{vk}, \boldsymbol{\eta}_{vk}) \end{array} \right\} \right], \quad h = 1, \ldots, N_v, \tag{31}$$

where $E_v[\cdot]$ denotes the conditional expectation operator given (d_v, e_v) (i.e., $E[\cdot|d_v, e_v]$). BD01 focus on equilibria with constant and symmetric beliefs. Using our notation above, we say that (constant) beliefs are symmetric when $\bar{\Pi}_{vh} = \bar{\Pi}_{vk}$ for any $h, k \in \{1, \ldots, N_v\}$ (for each v). When Brock and Durlauf's framework is interpreted as a Bayesian game, one can justify their focus on constant and symmetric beliefs under conditions laid out in Proposition 2 below.

To establish this proposition, define for each v, given (d_v, e_v), a function $m_v : [0,1] \to [0,1]$ as

$$m_v(r) := E_v\left[1\left\{U_1(Y_{vh} - P_{vh}, r, \boldsymbol{\xi}_v + \boldsymbol{u}_{vh}) \geq U_0\left(Y_{vh}, r, \boldsymbol{\xi}_v + \boldsymbol{u}_{vh}\right)\right\}\right]; \tag{32}$$

note that $m_v(r)$ is independent of individual index h under the conditional I.I.D. assumption given (d_v, e_v). Then the following characterization of beliefs holds:

Proposition 2 *Suppose that the same conditions hold as in Proposition 1 and the function $m_v(\cdot)$ defined in (32) is a contraction, i.e., for some $\rho \in (0,1)$,*

$$|m_v(r) - m_v(\tilde{r})| \leq \rho |r - \tilde{r}| \quad \textit{for any } r, \tilde{r} \in [0,1]. \tag{33}$$

Then, a solution $(\bar{\Pi}_{v1}, \ldots, \bar{\Pi}_{vN_v})$ of the system of N_v equations in (31) uniquely exists and is given by symmetric beliefs, i.e.,

$$\bar{\Pi}_{vh} = \bar{\Pi}_{vk} \quad \textit{for any } h, k \in \{1, \ldots, N_v\}.$$

The proof is given in Appendix A.2. Propositions 1-2 show that, given the (conditional) I.I.D. and contraction conditions, the equilibrium is characterized through

$$\Pi_{vh} = \bar{\pi}_v \quad \text{for any } h = 1, \ldots, N_v,$$

for some constant $\bar{\pi}_v := \bar{\pi}_v(d_v, e_v) \in [0,1]$ within each village (given (d_v, e_v)). This implies that the *beliefs can be consistently estimated by the sample average of* A_{vk} *over village* v, which is exploited in our empirical study.

The contraction condition (33) holds when the social interactions coefficient α is not large (in our linear index specification). In Section 5.5 below, we will provide sufficient conditions for the contraction and equilibrium uniqueness, as well as explain additional procedures that are needed for estimation and counterfactual analysis when multiplicity of equilibria may arise.

5 ECONOMETRIC SPECIFICATION, IDENTIFICATION AND ESTIMATION

Taking the belief variable Π_{vh} in the linear-index choice probability function $q_1(\cdot)$ to be the (limit of) observed fraction of usage in each village i.e. $\bar{\pi}_v = E_v[A_{vh}]\,(= \lim_{N_v\to\infty}\sum_{h=1}^{N_v} A_{vh}/N_v)$ as justified in Propositions 1-2, the index coefficients can be estimated semiparametrically using say, Bhattacharya (2008). However, unobserved village-effects may confound the consistency of these estimates; we overcome this by using a correlated random effects (CRE, henceforth) probit approach to estimate $q_1(\cdot)$, which is derived from a factor structure on the covariates and the village-effects, as follows.

5.1 Village Effects Specification

Our data for the application come from eleven different villages with an average of 195 households per village. It is plausible that utilities from using and from not using an ITN are affected by village-specific unobservable characteristics (i.e., $\xi_v = \xi_v^1 - \xi_v^0$ introduced in (35)), such as the chance of contracting malaria when not using an ITN. Recall the linear utility structure (3) from Section 3. Given this, together with the unobserved heterogeneity specification in (28), $\boldsymbol{\eta}_{vh} = \boldsymbol{\xi}_{vh} + \boldsymbol{u}_{vh}$, we model

$$
\begin{aligned}
U_0(Y_{vh}, \Pi_{vh}, \boldsymbol{\eta}_{vh}) &= \delta_0 + \beta_0 Y_{vh} + \alpha_0 \Pi_{vh} + \underbrace{\xi_v^0 + u_{vh}^0}_{\eta_{vh}^0}, \\
U_1(Y_{vh} - P_{vh}, \Pi_{vh}, \boldsymbol{\eta}_{vh}) &= \delta_1 + \beta_1 (Y_{vh} - P_{vh}) + \alpha_1 \Pi_{vh} + \underbrace{\xi_v^1 + u_{vh}^1}_{\eta_{vh}^1},
\end{aligned} \tag{34}
$$

where $\boldsymbol{\xi}_v = (\xi_v^0, \xi_v^1)$ and $\boldsymbol{u}_{vh} = (u_{vh}^0, u_{vh}^1)$ denote village and individual specific characteristics, respectively, both of which are unobservable. Therefore,

$$
\begin{aligned}
&U_1(Y_{vh} - P_{vh}, \Pi_{vh}, \boldsymbol{\eta}_{vh}) - U_0(Y_{vh}, \Pi_{vh}, \boldsymbol{\eta}_{vh}) \\
&= (\delta_1 - \delta_0) - \beta_1 P_{vh} + (\beta_1 - \beta_0) Y_{vh} + (\alpha_1 - \alpha_0)\Pi_{vh} + \underbrace{\xi_v^1 - \xi_v^0}_{\xi_v} + \underbrace{(u_{vh}^1 - u_{vh}^0)}_{\varepsilon_{vh}} \\
&\equiv c_0 + c_1 P_{vh} + c_2 Y_{vh} + \alpha \Pi_{vh} + \xi_v + \varepsilon_{vh},
\end{aligned} \tag{35}
$$

where ε_{vh} is assumed to have zero mean and unit variance for scale and location normalization.

Non-identification of the village effects ξ_v: Brock and Durlauf (2007) discussed difficulties of estimating social interactions models in presence of group-specific unobservables and presented a non-identification result (their Proposition 2). To see this in our context, consider constant beliefs, $\Pi_{vh} = \bar{\pi}_v$ (justified in Propositions 1-2). Since ξ_v is village specific and many observations per village are available, we can estimate village specific intercepts γ_v by regression of take-up A_{vh} on price and income $W_{vh} = (P_{vh}, Y_{vh})$ that vary across households h within village v, together with village dummies, i.e.,

$$\Pr\left(A_{vh} = 1 | W_{vh} = w; d_v, e_v\right) = F_\varepsilon(w\boldsymbol{c}' + \underbrace{c_0 + \alpha\bar{\pi}_v + \xi_v}_{\gamma_v}), \tag{36}$$

where the left-hand side (LHS) is computed under the conditional law given (d_v, e_v), and $F_\varepsilon(\cdot)$ is the CDF of $-\varepsilon_{vh}$.[11]

The realized ξ_v is a constant within each village; thus, $\xi_1, \ldots, \xi_{\bar{v}}$ and the universal constant c_0 cannot be separately identified and we reparametrize $\bar{\xi}_v := c_0 + \xi_v$. For each v and each realized ξ_v, the LHS of (36) is identifiable as a function of w; thus, under a parametric specification of $F_\varepsilon(\cdot)$ together with the exogeneity condition **C2** (ii) and a rank condition for covariates (stated below), $(\boldsymbol{c}, \gamma_1, \ldots, r_{\bar{v}})$ is also identified. The identified coefficients $\gamma_1, \ldots, \gamma_{\bar{v}}$ on the village dummies therefore satisfy the equations $\gamma_v = \alpha\pi_v + \bar{\xi}_v \equiv c_0 + \xi_v$ $(v = 1, \ldots, \bar{v})$.[12] However, even in the reparametrized equations, there are as many $\bar{\xi}_v$ as there are γ_v, so that we have $\bar{v}$ equations with $\bar{v} + 1$ unknowns $\bar{\xi}_1, \ldots, \bar{\xi}_{\bar{v}}$, and α, which are needed for policy and counterfactual analysis but cannot be separately identified.

5.2 Factor Structure and Correlated Random Effects Modelling

We surmount non-identification of ξ_v by an approximate version of the Mundlak-Chamberlain correlated random effects (CRE) structure, cf. Section 15.8.2 of Wooldridge (2010), which is routinely used as a reasonable middle ground between fixed and random effects in the panel econometrics literature. While the CRE device is typically intended for short panels, our setting here may be seen like a "long panel" in that each village is supposed to have its own effect that is shared by a large number of households (note that our dataset does not have a panel structure but consists of several cross sectional datasets). To have our specification consistent with the long-panel-like setting and Section 4 (in particular, **C2**), we consider the following factor structure for the observable

[11] Recall that $\varepsilon_{vh}\left(= u_{vh}^1 - u_{vh}^0\right)$ is assumed to be independent of W_{vh} conditionally on (d_v, e_v) in **C2**; and ξ_v is determined by (d_v, e_v). Below, it is further assumed that ε_{vh} is jointly independent of W_{vh} and (d_v, e_v).

[12] In the application, we a run a probit of A_{vh} on covariates W_{vh} (P_{vh}, Y_{vh}, and other variables) and a dummy for each village which corresponds to converting these conditional moments to a set of unconditional ones.

covariate W_{vh} and the village specific variable ξ_v,

$$W_{vh} = d_v + \tau_{vh} \text{ and } \xi_v = d_v\boldsymbol{\delta}' + e_v, \tag{37}$$

where d_v is a vector of "factor" variables (with the same dimension as W_{vh}) that are common in W_{vh} and ξ_v, τ_{vh} is the covariate specific, idiosyncratic component that is assumed to have zero mean (for location normalization) and is defined through $\tau_{vh} := W_{vh} - d_v$, $\boldsymbol{\delta}$ is a (row) vector of constant coefficients on the factor, and e_v is a village specific variable that affects only ξ_v.[13]

We assume each household in village v knows the realization of (d_v, e_v), while all the right-hand-side components in (37) are unobservable to researchers. Let $\bar{W}_v := (1/N_v)\sum_{v=1}^{N_v} W_{vh}$. Then, we can write $d_v = \bar{W}_v - (1/N_v)\sum_{v=1}^{N_v}\tau_{vh}$. Plugging this into the second equation in (37), we can write

$$\xi_v = \bar{W}_v\boldsymbol{\delta}' + e_v + o_p(1), \tag{38}$$

for each N_v, which follows from $(1/N_v)\sum_{v=1}^{N_v}\tau_{vh} = O_p(1/\sqrt{N_v}) = o_p(1)$ by a standard central limit theorem. We note that (38) is a reduced-form representation for each (sufficiently large) N_v derived from the structural assumption (37). We further assume that the error term satisfies

$$e_v \perp \left(\{(W_{vh}, \varepsilon_{vh})\}_{h=1}^{N_v}, d_v\right) \text{ and } e_v \sim N\left(0, (\sigma_e^*)^2\right), \tag{39}$$

for each v, where we note that $\{e_v\}_{v=1}^{\bar{v}}$ is I.I.D. under **C1** and (39) (we denote by σ_e^* the *true* standard deviation parameter; and subsequently, $*$ is often used to denote true parameters).[14] In standard short-panel cases, a distributional assumption is directly imposed on group effects, say, $\xi_v | \{W_{vh}\}_{h=1}^{N_v} \sim N(\bar{W}_v\boldsymbol{\delta}', (\sigma_e^*)^2)$ (see Wooldridge, 2010, p. 615); in our setting, this conditional normality of ξ_v holds in an approximate sense with a small order $o_p(1)$ term in (38).[15] We further assume that

$$\varepsilon_{vh} \Big| \left(\{W_{vh}\}_{h=1}^{N_v}, d_v\right) \sim N(0,1), \tag{40}$$

which is analogous to specifications in Chamberlain (1980) and Wooldridge (2010). Putting all of this together, we can write

$$A_{vh} = 1\{W_{vh}\boldsymbol{c}' + c_0 + \alpha\bar{\pi}_v + d_v\boldsymbol{\delta}' + e_v + \varepsilon_{vh} \geq 0\} \text{ for each } (v,h)$$

[13] Note that one component W_{vh} is the price P_{vh} faced by the household, which is randomized across households. The corresponding component of d_v in (37) is the average price within the village and its coefficient in $\boldsymbol{\delta}$ is set as zero in our application (as the randomized price does not capture village specific features).

[14] Brock and Durlauf (2007) have also considered a (linear) restriction on the group effects similar to (38) (see their Section 4.1.2 and Assumption L.1) and argue that it may help partial identification.

[15] The original CRE model, the so-called Mundlak-Chamberlain device, is not derived from a factor structure as in (37); we do not know of any other paper that considers a CRE model as a reduced form derived from some factor structure, which can be thought of as a separate contribution of the present paper. Our derivation of the approximate CRE model makes the households' information structure transparent which is required for constructing an econometric framework consistent with the game structure and **C2**. If we directly imposed (38) as is done in standard short panel contexts, it would be difficult to see which parts of ξ_v should be known to households and to interpret the conditional i.i.d.-ness in **C2** given the village specific variables.

and compute the conditional probability as

$$\begin{aligned}\Pr\left(A_{vh}=1|W_{vh}=w;d_v\right) &= \int F_{\varepsilon}\left(w\boldsymbol{c}'+c_0+\alpha\bar{\pi}_v+d_v\boldsymbol{\delta}'+e\right)\phi_{\sigma_e^2}\left(e\right)de \\ &= F_{\varepsilon+e}\left(w\boldsymbol{c}'+c_0+\alpha\bar{\pi}_v+d_v\boldsymbol{\delta}'\right) \\ &= F_{\varepsilon+e}\left(w\boldsymbol{c}'+c_0+\alpha\bar{\pi}_v+\bar{W}_v\boldsymbol{\delta}'\right)+o_p\left(1\right), \end{aligned} \quad (41)$$

where the probability on the LHS is computed under the conditional law given d_v (i.e., it is with respect to the distribution of $-\left(\varepsilon_{vh}+e_v\right)$); $F_{\varepsilon+e}$ is the CDF of $-(\varepsilon_{vh}+e_v)\sim N(0,1+\sigma_e^2)$, and last equality holds since $d_v=\bar{W}_v+o_p\left(1\right)$; and (41) can be shown to hold uniformly over (v,h), w, and $(\boldsymbol{c},c_0,\alpha,\boldsymbol{\delta})$, under compactness of the parameter space and Condition **CR2 (ii)**, imposed in the next subsection. Denote by Φ the CDF of $N\left(0,1\right)$. Then, the leading term on the right-hand side (RHS) of (41) can be written as

$$\Phi\left(\frac{w\boldsymbol{c}'+c_0+\alpha\bar{\pi}_v+\bar{W}_v\boldsymbol{\delta}'}{\sqrt{1+(\sigma_e^*)^2}}\right)\equiv\Phi(w\bar{\boldsymbol{c}}'+\bar{c}_0+\bar{\alpha}\bar{\pi}_v+\bar{W}_v\bar{\boldsymbol{\delta}}'). \quad (42)$$

For calculating the LHS of (36), e_v is treated as a part of the parameter γ_v; in contrast, the LHS of (41) is calculated with respect to the distribution of the unobservable $\varepsilon_{vh}+e_v$ across all households over all villages. Both the probabilities in (36) and (41) concern the same outcome variable A_{vh} but they differ in conditioning variables. The former probability can be consistently estimated within each village as $N_v\to\infty$ for each v, while consistent estimation of the latter requires $\bar{v}\to\infty$ (in addition to $N_v\to\infty$) since village specific effects e_v have to be averaged out to match the probability (41) computed as the integral of e_v via its approximation (42).

Putting all this together, our estimation steps are as follows:

1. First run a probit of A_{vh} on W_{vh}, $\bar{W}_v$ and $\hat{\pi}_v\equiv(1/N_v)\sum_{v=1}^{N_v}A_{vh}$ corresponding to (42) to obtain estimates of $\bar{\boldsymbol{c}},\bar{c}_0,\bar{\alpha},\bar{\boldsymbol{\delta}}$;

2. Then run a probit of A_{vh} on W_{vh} and village dummies corresponding to (36) and obtain estimates $\boldsymbol{c}',\gamma_1,\gamma_2,...\gamma_{\bar{v}}$;

3. Estimate σ_e^* by the ratio of the price coefficient in the former to that in the latter probit;

4. Estimate c_0 via $\bar{c}_0\times\sqrt{1+(\sigma_e^*)^2}$ and α via $\bar{\alpha}\times\sqrt{1+(\sigma_e^*)^2}$;

5. From (36), estimate $\xi_v=\gamma_v-c_0-\alpha\hat{\pi}_v$.

These are all the quantities we need for empirical calculation of welfare expressions outlined in Section 3. In the empirical application below, the parameters $\bar{\boldsymbol{c}},\bar{c}_0,\bar{\alpha},\bar{\boldsymbol{\delta}}$ are estimated via pseudo-MLE by running an ordinary probit regression of A_{vh} on $W_{vh},\bar{\pi}_v$ and $\bar{W}_v$.

Thus to summarize, it follows from Brock and Durlauf's (2007) arguments, outlined above, that identification of village specific parameters, ξ_v, is in general impossible in the presence of social interaction effects. We overcome this through our CRE condition (38) which imposes more structure on ξ_v and letting the number of groups, i.e. $\bar{v} \to \infty$, as formally stated in the next subsection and the proof of consistency in Appendix A.3. As such, this is a new finding for social-interactions models. Note that if e_v is non-stochastic (i.e., $\sigma_e^* = 0$ and $\xi_v = \bar{W}_v \boldsymbol{\delta}' + o_p(1)$, instead of (38)), the above scheme using two probit regressions leads to identification and consistent estimation without the many-village assumption of $\bar{v} \to \infty$.[16]

5.3 Estimation and Consistency

Now we discuss consistency of the estimation procedure outlined in the previous subsection. We focus on the consistency of the first probit (36), the setting of which is non-standard under the CRE structure and the many-village asymptotics $\bar{v} \to \infty$; in contrast, the setting of the second probit (41) or (42) can be analyzed in the same way as in Hahn and Kuersteiner (2011), and a detailed dicussion of its consistency is omitted.[17]

For verification of consistency, we assume that the number of households in each village can be written as

$$N_v = r_v N_0, \tag{43}$$

where $r_v \in (\underline{r}, \bar{r})$ is a constant that is independent of N_0 and $\bar{v}$ with $0 < \underline{r} \le \bar{r} < \infty$ (i.e., r_v is uniformly bounded from below and above), and let $N = \sum_{v=1}^{\bar{v}} N_v$ is the total number of households in all villages combined. This assumption means that all $N_1, \ldots, N_{\bar{v}}$ grow at the same rate, so that none of villages is asymptotically negligible.

Comparing the two probabilities in (42), we can identify/estimate all the parameters γ_v^*, $(\boldsymbol{c}^*, c_0^*, \alpha^*, \boldsymbol{\delta}^*)$, and σ_e^*, which allows us to obtain estimates of $\xi_1, \ldots \xi_{\bar{v}}$. Consistent estimation of these parameters can be achieved through the following two probit regressions.[18] First, a probit of A_{vh} on W_{vh} and village dummies allows us to obtain estimates

$$(\hat{\boldsymbol{c}}, \hat{\gamma}_1, \ldots, \hat{\gamma}_{\bar{v}}) = \underset{\boldsymbol{c} \in \Upsilon_1;\ (\gamma_1, \ldots, \gamma_{\bar{v}}) \in \Upsilon(\bar{v}) \times \cdots \times \Upsilon(\bar{v})}{\operatorname{argmax}} \hat{Q}(\boldsymbol{c}, \gamma_1, \ldots, \gamma_{\bar{v}}),$$

where the objective function $\hat{Q}$ is

$$\hat{Q}(\boldsymbol{c}, \gamma_1, \ldots, \gamma_{\bar{v}}) = \frac{1}{N} \sum_{v=1}^{\bar{v}} \sum_{h=1}^{N_v} \mathcal{L}_{vh}(\boldsymbol{c}, \gamma_v), \tag{44}$$

[16] This identification/estimation scheme of CRE models using two probit regressions appears new, which allows us to recover the standard deviation σ_e^* (which is not typically identified in standard short-panel cases; see e.g. p. 617 of Wooldridge, 2010) and further all the realized values of $e_1, \ldots, e_{\bar{v}}$.

[17] Our second probit setting is even simpler than Hahn and Kuersteiner's in that the number of parameters do not increase $N_v \to \infty$ or $\bar{v} \to \infty$. A notable difference is that the objective function $\hat{R}$ incurs some approximation error $o_p(1)$ by using (42) instead of (41); but given the uniformity of the $o_p(1)$ as stated, this error can be negligible for the consistency discussion.

[18] Note that our practical estimation procedure exploits the fixed point restriction by an iteration process (discussed in Section 5.5). It is slightly more complicated than the procedure outlined here; but the substance of our identification arguments does not change between the two procedures; our exposition here is based on the simpler procedure.

$$\mathcal{L}_{vh}\left(\boldsymbol{c},\gamma_v\right) := A_{vh}\log\Phi\left(W_{vh}\boldsymbol{c}'+\gamma_v\right)+\left(1-A_{vh}\right)\log\left(1-\Phi\left(W_{vh}\boldsymbol{c}'+\gamma_v\right)\right), \tag{45}$$

Υ_1 is a compact set in $\mathbb{R}^{d_W}$, and $\Upsilon(\bar{v})$ is a compact interval on $\mathbb{R}$ that may grow as $\bar{v}\to\infty$ (specified in (69) in Appendix A.3). Second, via a second probit of A_{vh} on $(W_{vh},1,\hat{\pi}_v,\bar{W}_v)$, we can estimate the coefficients $(\bar{\boldsymbol{c}}^*,\bar{c}_0^*,\bar{\alpha}^*,\bar{\boldsymbol{\delta}}^*)$ through

$$(\widehat{\bar{\boldsymbol{c}}},\widehat{\bar{c}}_0,\widehat{\bar{\alpha}},\widehat{\bar{\boldsymbol{\delta}}}) = \underset{(\bar{\boldsymbol{c}},\bar{c}_0,\bar{\alpha},\bar{\boldsymbol{\delta}})\in\Upsilon_2}{\operatorname{argmax}}\ \hat{R}(\bar{\boldsymbol{c}},\bar{c}_0,\bar{\alpha},\bar{\boldsymbol{\delta}}),$$

where the objective function $\hat{R}$ is defined as

$$\begin{aligned}\hat{R}(\bar{\boldsymbol{c}},\bar{c}_0,\bar{\alpha},\bar{\boldsymbol{\delta}}) = \frac{1}{N}\sum\nolimits_{v=1}^{\bar{v}}\sum\nolimits_{h=1}^{N_v}\Big[&A_{vh}\log\Phi(W_{vh}\bar{\boldsymbol{c}}'+\bar{c}_0+\bar{\alpha}\hat{\pi}_v+\bar{W}_v\bar{\boldsymbol{\delta}}')\\ &+(1-A_{vh})\log(1-\Phi(W_{vh}\bar{\boldsymbol{c}}'+\bar{c}_0+\bar{\alpha}\hat{\pi}_v+\bar{W}_v\bar{\boldsymbol{\delta}}'))\Big],\end{aligned}$$

and Υ_2 is a compact set in $\mathbb{R}^{2d_W+1}$.[19] Then, we can recover an estimate of $\hat{\sigma}_e^2$ through a ratio of the first (or any other) components of $\hat{\boldsymbol{c}}$ and $\widehat{\bar{\boldsymbol{c}}}$, which yields $\sqrt{1+\hat{\sigma}_e^2}$ and further $(\hat{c}_0,\hat{\alpha},\hat{\boldsymbol{\delta}}) = (\widehat{\bar{c}}_0,\widehat{\bar{\alpha}},\widehat{\bar{\boldsymbol{\delta}}})\times\sqrt{1+\hat{\sigma}_e^2}$. Finally, given these estimates, we can compute

$$\hat{\xi}_v = \hat{\gamma}_v - \hat{c}_0 - \hat{\alpha}\hat{\pi}_v$$

for each v, which then allows us to calculates the welfare estimates of Section 3.

The proof of consistency for the first probit is involved due to the CRE structure and the formal steps are provided in Appendix A.3. The key substantive assumption delivering consistency is as follows:

CR1 (i) For each v, let $\lambda_v^{\min}$ be the minimum of the eigenvalues of $E_{\boldsymbol{\omega}_v}[(W_{vh},1)'(W_{vh},1)]$, which is a square (real symmetric) matrix of order d_W+1, where d_W is the dimension of W_{vh}. Then, $\inf_{v\geq 1}\lambda_v^{\min}>0$. (ii) The covariates and unobservables $(W_{vh},\xi_v,\varepsilon_{vh})$ satisfy (37), (39), and (40). (iii) Let $\bar{W}_v^* := \underset{N_v\to\infty}{\operatorname{plim}}\frac{1}{N_v}\sum_{h=1}^{N_v}W_{vh}(=E_{\boldsymbol{\omega}_v}[W_{vh}])$ for each v (the existence of $\bar{W}_v^*$ is supposed), and

$$\bar{\mathbb{W}} := \begin{bmatrix} 1 & \bar{\pi}_1 & \bar{W}_1^* \\ \vdots & & \vdots \\ 1 & \bar{\pi}_{\bar{v}} & \bar{W}_{\bar{v}}^* \end{bmatrix},$$

which is a $\bar{v}\times(2+d_W)$ matrix. Then, suppose that $\bar{\mathbb{W}}$ is of rank $2+d_W$.

Condition (i) of **CR1** allows us to identify $(\boldsymbol{c}^*,\gamma_v^*)$ for each v as the maximizer of $Q_v(\boldsymbol{c},\gamma_v) = E_{\boldsymbol{\omega}_v}[\mathcal{L}_{vh}(\boldsymbol{c},\gamma_v)]$ whose empirical analogue $\hat{Q}_v$ defined in (67) is a constituent of the objective $\hat{Q}$ defined in (44); see the expression (68) in Appendix A.3. The condition on the uniform lower bound

[19] By the results in Section 4.1, we have $\bar{\pi}_v = E[A_{vh}|d_v,e_v]$ in the equilibrium, which can be consistently estimated by an average within each village, $\hat{\pi}_v = \frac{1}{N_v}\sum_{h=1}^{N_v}A_{vh}$.

of $\lambda_v^{\min}$ together with (43) guarantees the identification/consistency of all γ_v^* when $\bar{v} \to \infty$. (iii) of **CR1** is used to verify identification of the parameters in the second probit, $(\bar{\boldsymbol{c}}^*, \bar{c}_0^*, \bar{\alpha}^*, \bar{\boldsymbol{\delta}}^*)$, where we note that by the definition in (36) and the reduced form expression of ξ_v in (38), we can write the true village-specific effect $\gamma_v^* = c_0^* + \alpha^* \bar{\pi}_v + \bar{W}_v^* \boldsymbol{\delta}' + e_v$.

The formal consistency statement is expressed in the following proposition:

Proposition 3 *Suppose that Conditions **C1**, **C2**, **CR1 (i)-(ii)**, and the technical condition **CR2** (stated in Appendix A.3), Specifications (39) and (40) hold and that $\bar{v}$ and N_0 satisfy Assumption (43) with*

$$\bar{v}^{4(\sigma_e^*)^2} (\log \bar{v})^{3+4(\sigma_e^*)^2} (\log N_0) \Big/ N_0 \to 0 \quad (as\ N_0 \to \infty). \tag{46}$$

Then, as $N_0 \to \infty$ and $\bar{v} \to \infty$,

$$\|\hat{\boldsymbol{c}} - \boldsymbol{c}^*\| \xrightarrow{p} 0 \quad and \quad \max_{v \in \{1,\ldots,\bar{v}\}} |\hat{\gamma}_v - \gamma_v^*| \xrightarrow{p} 0.$$

The proof is provided in Appendix A.3.

Verification of this proposition for the first probit is not trivial. This is because (I) given the asymptotic assumption $\bar{v} \to \infty$, required for the consistency in the second probit, the number of parameters tends to infinity; and (II) each parameter $\gamma_v^* = c_0^* + \alpha^* \bar{\pi}_v + \bar{W}_v^* \boldsymbol{\delta}^{*\prime} + e_v$ includes a realization of $e_v \sim N\left(0, \sigma_e^2\right)$ and thus the maximum of realized $|\gamma_1^*|, \ldots, |\gamma_{\bar{v}}^*|$ grows with positive probability as $\bar{v} \to \infty$ since e_v has unbounded support $(-\infty, \infty)$.

Several previous papers on panel models have considered a setting like (I), such as Hahn and Kuersteiner (2011) and Fernández-Val and Weidner (2016). However, in these papers, the "growing *magnitude* of parameters" as (II) is not allowed for, i.e., typically, all parameters are supposed to be in a fixed compact set.[20]

These problems, in particular (II), make it hard to establish the identification of $(\boldsymbol{c}^*, \gamma_v^*)$. However, we can overcome this by showing that given the I.I.D. $\{e_v\}_{v \geq 1}$, the maximum of $|e_1|, \ldots, |e_{\bar{v}}|$ is bounded by $\sqrt{4(\sigma_e^*)^2 \log[\bar{v}\,(\log \bar{v})^t]}$ almost surely for any $t > 1/2$ (Lemma 1). This result allows us to restrict possible support of each γ_v^* as a compact set that grows (as $\bar{v} \to \infty$); and within this support, if $||\,(\boldsymbol{c}, \gamma_v) - (\boldsymbol{c}^*, \gamma_v^*)\,|| > \epsilon_1$, we can always find some constant $C_Q > 0$ such that $Q_v\left(\boldsymbol{c}^*, \gamma_v^*\right) - Q_v\left(\boldsymbol{c}, \gamma_v\right) \geq C_Q[\bar{v}(\log \bar{v})]^{-2(\sigma_e^*)^2}$ almost surely (shown in Lemma 2), which means the identification of $(\boldsymbol{c}^*, \gamma_v^*)$ as the unique maximizer of Q_v. If the support were not restricted, the lower bound of the difference between $Q_v\left(\boldsymbol{c}^*, \gamma_v^*\right)$ and $Q_v\left(\boldsymbol{c}, \gamma_v\right)$ would be zero as $\bar{v} \to \infty$, which would complicate identification and consistency.[21]

[20] This is explicitly assumed in Hahn and Kuersteiner's Condition 4, while cases like (II) have to be typically excluded by Fernández-Val and Weidner's Assumption 4.1 (v) (the presence of uniform bounds $b_{\min}$ and $b_{\max}$ for the derivative of their objective function).

[21] To see this, let $||\boldsymbol{c} - \boldsymbol{c}^*|| > \epsilon_1$ and $\gamma_v = \gamma_v^*$, for example. Then, for any $\boldsymbol{c} \neq \boldsymbol{c}^*$, $[Q_v\left(\boldsymbol{c}^*, \gamma_v^*\right) - Q_v\left(\boldsymbol{c}, \gamma_v^*\right)] \to 0$ as $|\gamma_v^*| \to \infty$, which holds for some v as $\bar{v} \to \infty$ since $\gamma_v^* = c_0^* + \alpha^* \bar{\pi}_v + \bar{W}_v^* \boldsymbol{\delta}^{*\prime} + e_v$ includes the normally distributed

The rate condition (46) requires $\bar{v}$ to grow slower than N_0 in particular when the variance of e_v is large, which is reasonable in the context of our empirical application, where $\bar{v} = 11$ may be regarded as small relative to $N_v = r_v N_0$ which is 195 on average.[22] It guarantees that the difference between $Q_v(\boldsymbol{c}^*, \gamma_v^*)$ and $Q_v(\boldsymbol{c}, \gamma_v)$ is larger than that between $\hat{Q}_v$ and Q_v, justifying the maximizer of $\hat{Q}_v$ as an estimator, implying the consistency.

To see how our two-step probit performs in finite samples, we implemented a small Monte Carlo exercise that is reported in Appendix A.5 and shows reassuring results for magnitudes of sample size resembling ours.

5.4 Calculation of Predicted Demand and Welfare

In order to calculate our welfare-related quantities, we need to estimate the structural choice probabilities $q_1(p, y, \pi)$ and the equilibrium values of the choice probabilities, π_0 and π_1, in the pre and post intervention situations. To do this we will assume that the unobservables $\varepsilon_{vh}\left(= u_{vh}^1 - u_{vh}^0\right)$ are independent of price and income, conditional on unobserved village-effects.[23] Note that prices in our data are randomly assigned, so the endogeneity concern is solely regarding income. Under income endogeneity, Bhattacharya (2018) had discussed interpretation of welfare distributions as conditional on income (see our discussion at the end of Section 7).

Welfare Calculations: Once we have estimates of the structural choice probabilities from the parametric model above, we can proceed with welfare calculation in presence of social spillovers and unobserved group-effects, as follows. Consider an initial situation where everyone faces the unsubsidized price p_0, so that the predicted take-up rate π_{0v} in village v solves:

$$\pi_{0v} = \int \Phi\left(c_0 + c_1 p_0 + c_2 y + \alpha \pi_{0v} + \xi_v\right) dF_Y^v(y), \tag{47}$$

where $F_Y^v(y)$ is the CDF of income Y_{vh} in village v. Section 5.2 above outlines the calculations of all parameters including the ξ_v's appearing in (47).

Now consider a policy induced price regime p_0 for ineligibles (with wealth larger than a) and p_1 for the eligible (with wealth less than or equal to a). Then the resulting usage $\pi_1 = \pi_{1v}$ in village

variable e_v. Note that for large $|\gamma_v^*|$, both $\Phi\left(W_{vh}(\boldsymbol{c}^*)' + \gamma_v^*\right)$ and $\Phi\left(W_{vh}\boldsymbol{c}' + \gamma_v^*\right)$ (the normal CDF's) are very close to 1 or 0 regardless of $\boldsymbol{c} \neq \boldsymbol{c}^*$ (i.e., variation of $\Phi(\cdot)$ is tiny in the tail region); thus, the difference between $Q_v(\boldsymbol{c}^*, \gamma_v^*)$ and $Q_v(\boldsymbol{c}, \gamma_v^*)$ is very small, which are computed through these CDF's.

[22] Note that regardless of the rate condition (46), for the first probit, the magnitude of $\bar{v}$ does not directly affect estimation precision of $(\hat{\boldsymbol{c}}, \hat{\gamma}_v)$ (up to first order), whose convergence rate is $1/\sqrt{N_v}$. In contrast, the rate condition matters for the second probit, for which the integration with respect to e_v has to be approximated by the sum over $e_1, \ldots, e_{\bar{v}}$.

[23] $q_1(p, y, \pi)$ defined in (4) as the probability computed with respect to the distribution of $\eta^0 - \eta^1\left(= \eta_{vh}^0 - \eta_{vh}^1\right)$. But given the specification of $\boldsymbol{\eta}_{vh} = \left(\eta_{vh}^0, \eta_{vh}^1\right)$ in Sections 4.1-5, it should be now interpreted as the one with respect to the distribution of ε_{vh} (conditionally on the village-fixed effects (d_v, e_v) or ξ_v), i.e., the probability (41) as a function of $w = (p, y)$ and $\bar{\pi}_v = \pi$.

v is obtained via solving the fixed point π_{1v} in the equation:

$$\pi_{1v} = \int \begin{bmatrix} 1\{y \le \tau\}\Phi(c_0 + c_1 p_1 + c_2 y + \alpha \pi_{1v} + \xi_v) \\ +1\{y > \tau\}\Phi(c_0 + c_1 p_0 + c_2 y + \alpha \pi_{1v} + \xi_v) \end{bmatrix} dF_Y^v(y). \tag{48}$$

For fixed (p_0, p_1), the right-hand sides of the above fixed point equations (47) and (48), viewed as functions of π_{0v} and π_{1v} respectively, are a map from $[0,1]$ to $[0,1]$ (π_{0v} and π_{1v} are probabilities taking values in $[0,1]$). By the continuity of Φ and Brouwer's fixed point theorem, there is at least one solution in π_{0v} and π_{1v}, respectively, implying "coherence". However, there may be multiple solutions, and then our welfare expressions would have to be applied separately for each feasible pair of values (π_{0v}, π_{1v}) (see our discussion on the equilibrium multiplicity in Section 5.5). Note that even if the solutions to (47) and (48) are unique, our expressions in Theorem 1 and Corollary 2 imply that welfare distributions are still not point-identified.

Finally, mean welfare effect of the policy change in village v can be calculated as

$$\mathcal{W}_v = \int \left[1\{y \le \tau\}\mathcal{W}_v^{\text{Elig}}(y) + 1\{y > \tau\}\mathcal{W}_v^{\text{Inelig}}(y)\right] dF_Y^v(y), \tag{49}$$

where $\mathcal{W}_v^{\text{Elig}}(y)$ and $\mathcal{W}_v^{\text{Inelig}}(y)$ are mean welfare at income y in village v, calculated from (10) for the eligible and (25) for the ineligible, respectively, using π_{0v} and π_{1v} as the predicted take-up probabilities in village v (analogous to π_0 and π_1 in (10) and (25)), $\alpha_1 \in [0, \alpha]$ as above).

5.5 Equilibrium Existence and Uniqueness

In this section, we present sufficient conditions of unique equilibrium and then discuss multiplicity of equilibria as well as its implication for our demand and welfare estimation. Note that given the parametric model, our equilibrium condition (or fixed point restriction) takes the form:

$$\pi_v = \int \Phi\left(w\boldsymbol{c}' + c_0 + \alpha\pi_v + \xi_v\right) d\tilde{F}_W(w) \quad (v = 1, \ldots, \bar{v}), \tag{50}$$

where $\tilde{F}_W(\cdot)$ is a CDF of $W_{vh} = (P_{vh}, Y_{vh})$. Some different $\tilde{F}_W$ has to be used, depending on the context. For example, on the RHS of (47), $\tilde{F}_W$ corresponds to the distribution that gives a point mass for $P_{vh} = p_0$ (when considering a counterfactual analysis with this p_0) and $Y_{vh} \sim F_Y^v(y)$ (the marginal distribution of the observable variable Y_{vh}); and on the RHS of (48), a different $\tilde{F}_W$ representing the new subsidy scheme is used.

As stated in the previous subsection, existence of a solution to (50) follows from Brouwer's fixed point theorem. It is also clear that if $\alpha \le 0$, the solution is unique. On the other hand, if $\alpha > 0$, a contraction condition is sufficient for uniqueness. The contraction condition (33) in Proposition 2 can be verified on a case by case basis. In particular, for the linear index model, it is easy to see that the condition for contraction is

$$|\alpha| \sup_{e \in \mathbb{R}} f_\varepsilon(e) < 1,$$

where α denotes the social interaction term, and $f_\varepsilon(\cdot)$ denotes the probability density of $-\varepsilon_{vh}$. In a probit specification in which ε_{vh} is the standard normal variable, $\sup_{e\in\mathbb{R}} f_\varepsilon(e) = 1/\sqrt{2\pi}$ and thus we require $|\alpha| < \sqrt{2\pi}(\simeq 2.506)$ and for a logit specification, $\sup_{e\in\mathbb{R}} f_\varepsilon(e) = 1/4$, and thus $|\alpha| < 4$. We check that our probit estimate satisfies this condition in our application.

Note that the contraction condition (33) is not necessary for uniqueness. That is, if a solution $(\bar{\Pi}_{v1}, \ldots, \bar{\Pi}_{vN_v})$ to the system of equations (31) is unique and $m_v(\cdot)$ (defined in (32)), which also depends on the distribution of covariates, has a unique fixed point (i.e., a solution to $r = m_v(r)$ is unique), the uniqueness for the equilibrium solution holds. We have imposed (33) as it is a convenient condition that guarantees uniqueness equilibrium solution; it is also typically easy to verify in applications.

The Maximum Number of Equilibrium Solutions: The variable w in (50) is multivariate but its RHS can be written as $\int \Phi(q + c_0 + \alpha\pi_v + \xi_v)\, dF^v_{W\boldsymbol{c}'}(q)$ in terms of the integral with respect to the univariate variable $W_{vh}\boldsymbol{c}'$, using its CDF $F^v_{W\boldsymbol{c}'}$ and support $[\underline{q}_v, \bar{q}_v]$, where existence of the finite endpoints of the support of $W_{vh}\boldsymbol{c}'$ is guaranteed under Condition **CR2** (ii) (provided in Appendix A.3). Then, applying the mean value theorem for Stieltjes integrals, we can find some $q_v \in [\underline{q}_v, \bar{q}_v]$ such that

$$\begin{aligned}\int_{\underline{q}_v}^{\bar{q}_v} \Phi(q + c_0 + \alpha\pi_v + \xi_v)\, dF^v_{W\boldsymbol{c}'}(q) &= \Phi(q_v + c_0 + \alpha\pi_v + \xi_v) \int_{\underline{q}_v}^{\bar{q}_v} dF^v_{W\boldsymbol{c}'}(q) \\ &= \Phi(q_v + c_0 + \alpha\pi_v + \xi_v).\end{aligned}$$

Therefore, for each v, the fixed point restriction can be re-written as

$$\pi_v = \Phi(q_v + c_0 + \alpha\pi_v + \xi_v) \quad \text{for each } v. \tag{51}$$

Here, by shape properties of the standard normal CDF $\Phi(x)$ (e.g., its derivative is the normal density $\phi(x)$ with the two inflection points, -1 and 1), we can see that (51) has at most three solutions ($\pi_v = 0$ or 1 cannot be a solution since each value of Φ is on $(0, 1)$). In particular, a continuum of solutions cannot exist since $\Phi(x)$ does not have a linear part on any interval in the real line. This is summarized via the following proposition whose proof is also evident from the above discussion.

Proposition 4 *For each v, the maximum number of (equilibrium) solutions to (50) is three.*

This is analogous to Proposition 2 of BD01 for the logit distribution case without covariates W_{vh}. The number of equilibria is determined by the value of α as well as those of q_v, c_0, and (in particular) the unobserved group effects ξ_v. We now discuss implications of equilibrium multiplicity in estimation.

Preference and Demand Estimation under Multiple Equilibria: Our estimation involves maximum likelihood in nonlinear models with strategic interactions; thus, it is useful to recall Hahn and Moon (2010) who consider estimation of game theoretic models possibly with multiple equilibria under a panel setting, i.e., observations from many markets are obtained repeatedly over several time periods. These authors interpret unobservables affecting equilibrium selection as an unobserved fixed effect, assuming that which equilibrium is selected in one group is fully characterized by each unobserved fixed effect, which may be correlated with observed characteristics. Then they show that equilibrium multiplicity is unlikely to be a problem in panel settings when the number of equilibria is finite which, in the panel terminology, is equivalent to the fixed effect having finitely many support points. However, this result requires that the number of equilibria be constant across parameters and covariates, which is a strong restriction.

In contrast, in our setting, each village/group effect can be interpreted as a part of players' preference parameters and it does not fully determine which equilibrium arises. That is, under the same value of group/village effect ξ_v, we may observe different equilibria in village v. Given the knowledge of all the preference parameters and the distribution of all the covariates and error variables, one can determine the number of possible equilibria and possible values of beliefs by investigating all solutions of the fixed point equation (47), but cannot in advance see which equilibrium would be realized. We note that our model setting is not equipped with any equilibrium selection mechanism (just like thus in Brock and Durlauf's). In our setup, given $\bar{\pi}_v$ and ξ_v, we do not need to solve the equilibrium system to predict each player's behavior, which is determined by $A_{vh} = 1\{W_{vh}\boldsymbol{c}' + c_0 + \alpha\bar{\pi}_v + \xi_v + \varepsilon_{vh} \geq 0\}$, and preference parameters can be estimated without exploiting the equilibrium fixed point condition.[24] This is possible since (A) the only objects that are endogenously determined in equilibrium are $\bar{\pi}_v$ $(v = 1, \ldots, \bar{v})$, which can be identifiable as $E_v[A_{vh}]$ and thus consistently estimated in our "large market" setting with a large number of players in each village (as discussed in Section 4.1); and (B) the group effects parameters ξ_v can also be consistently estimated under the correlated random effects structure.

These features of our (and Brock and Durlauf's) modeling allow us to avoid intrinsically difficult problems caused by the equilibrium multiplicity. In particular, in any equilibrium realization, the same preference parameters (that are invariant under different equilibria) can be identified and thus consistently estimated. As a further illustration, consider a case in which there are three equilibria, i.e., the fixed point equation (47) has three solutions, $\bar{\pi}_v^H$, $\bar{\pi}_v^M$, and $\bar{\pi}_v^L$, where we let $\bar{\pi}_v^H > \bar{\pi}_v^M > \bar{\pi}_v^L$ and call each of equilibria as H, M, or L. Then, depending on $t \in \{H, M, L\}$, we have a different discrete choice model:

$$A_{vh} = 1\{W_{vh}\boldsymbol{c}' + c_0 + \alpha\pi_v^t + \xi_v + \varepsilon_{vh} \geq 0\}.$$

Note that the outcome variable changes depending on which equilibrium arises (i.e., one can write $A_{vh} = A_{vh}^t$); and thus, for each equilibrium t, π_v^t can be consistently estimated by $\frac{1}{N_v}\sum_{h=1}^{N_v} A_{vh}$. By

[24] This is quite different from the so-called two step estimation approach (typically used in the empirical industrial organization game literature) as in Hotz and Miller (1993) and Pesendorfer and Schmidt-Dengler (2008), in which equilibrium conditions provide the basis of identification.

plugging in the estimated version of $\pi_v = \pi_v^t$, we can construct objective functions (i.e., likelihood functions) to be maximized, based on which we can consistently estimate the preference parameters regardless of the realized equilibrium $t \in \{H, M, L\}$. As for consistency, the preference parameters can be identified as the unique maximizers of the limits of the objective functions. In particular, for our *first* probit regression (cf. Section 5.3), the choice probability is

$$\Pr\left(A_{vh}^t = 1 | W_{vh} = w; d_v, e_v\right) = \Phi(w(\boldsymbol{c}^*)' + \gamma_v^{*t}),$$

under equilibrium t, where $\gamma_v^{*t} (= c_0^* + \alpha^* \bar{\pi}_v^t + \xi_v)$ depends on which equilibrium has occurred, and the (limit) objective function (under equilibrium t) is

$$Q_v\left(\boldsymbol{c}, \gamma_v\right) = E_v\left[A_{vh}^t \log \Phi\left(W_{vh}\boldsymbol{c}' + \gamma_v\right) + \left(1 - A_{vh}^t\right) \log\left(1 - \Phi\left(W_{vh}\boldsymbol{c}' + \gamma_v\right)\right)\right],$$

where $E_v[\cdot] = E[\cdot | d_v, e_v]$. Then, given the true parameter $(\boldsymbol{c}^*, \gamma_v^{t*}) (\neq (\boldsymbol{c}, \gamma_v))$

$$\begin{aligned}
&Q_v\left(\boldsymbol{c}^*, \gamma_v^{*t}\right) - Q_v\left(\boldsymbol{c}, \gamma_v\right) \\
&= -E_v\left[\Phi(W_{vh}(\boldsymbol{c}^*)' + \gamma_v^{t*}) \log\left(\frac{\Phi\left(W_{vh}\boldsymbol{c}' + \gamma_v\right)}{\Phi\left(W_{vh}\boldsymbol{c}^{*\prime} + \gamma_v^{t*}\right)}\right)\right. \\
&\left. + \left(1 - \Phi(W_{vh}(\boldsymbol{c}^*)' + \gamma_v^{t*})\right) \log\left(\frac{1 - \Phi\left(W_{vh}\boldsymbol{c}' + \gamma_v\right)}{1 - \Phi\left(W_{vh}\boldsymbol{c}^{*\prime} + \gamma_v^{t*}\right)}\right)\right] \\
&> -E_v \log\left\{\Phi\left(W_{vh}\boldsymbol{c}' + \gamma_v\right) + 1 - \Phi\left(W_{vh}\boldsymbol{c}' + \gamma_v\right)\right\} = -\log\{1\} = 0,
\end{aligned}$$

where the equality has used the law of iterated expectation and correct specification assumption (i.e., $E_v\left[A_{vh}^t | W_{vh}\right] = \Phi(W_{vh}(\boldsymbol{c}^*)' + \gamma_v^{t*})$), and the strict inequality follows from Jensen's inequality, the strict convexity of $-\log(\cdot)$, and the rank condition on $(W_{vh}, 1)$ (**CR1** (ii)).[25] That is, we have

$$Q_v\left(\boldsymbol{c}^*, \gamma_v^{t*}\right) > Q_v\left(\boldsymbol{c}, \gamma_v\right) \text{ for any } \left(\boldsymbol{c}, \gamma_v\right) \neq \left(\boldsymbol{c}^*, \gamma_v^{t*}\right).$$

Thus, $\left(\boldsymbol{c}^*, \gamma_v^{t*}\right)$ is identified as the unique maximizer of $Q_v(\cdot, \cdot)$; in particular, the same $\boldsymbol{c}^*$ is always identified under any of equilibrium t, while the identified γ_v^{t*} depends on t, which corresponds to our specification of $\gamma_v^{t*} = c_0^* + \alpha^* \bar{\pi}_v^t + \xi_v$ (including the equilibrium object $\bar{\pi}_v^t$). The identification argument for the *second* probit under the CRE structure is analogous (so details are omitted here): under any equilibrium t, the same $(\bar{\boldsymbol{c}}^*, \bar{c}^*, \bar{\alpha}^* \bar{\boldsymbol{\delta}}^*)$ is obtained as the unique maximizer of the limit of $\hat{R}(\bar{\boldsymbol{c}}, \bar{c}_0, \bar{\alpha}, \bar{\boldsymbol{\delta}})$. That is, given the CRE structure, the same group effects ξ_v can be identified in any equilibrium t through the procedure outlined in Section 5.3. Thus our estimation procedure need not to use the equilibrium fixed point restriction and is robust to equilibrium multiplicity.

In our empirical application, we use an iterative estimator that exploits the equilibrium fixed point restriction as in Pastorello, Patilea, and Renault (2003), and Dominitz and Sherman (2005).

[25] This identification argument is standard (as in Newey and McFadden, 1994, Example 1.2 on page 2125). While we believe that this inequality for each v is useful for illustrating identification under the equilibrium multiplicity, it is not sufficient for consistency when $\bar{v} \to \infty$ i.e., Proposition 3. For verification of the proposition, we have derived uniform lower bound of $Q_v\left(\boldsymbol{c}^*, \gamma_v^{*t}\right) - Q_v\left(\boldsymbol{c}, \gamma_v\right)$ (Lemma 2 in the Appendix).

This estimator is more efficient under correct specification than the estimator that does not use (50). For this iterative estimation, the contraction property of the fixed point mapping (implying unique equilibrium) is key, the sufficient condition for which is a "small α". Through preliminary investigation i.e. checking estimates obtained without exploiting the fixed-point condition (50), we have confirmed that the estimate of α is small, so that the contraction condition is satisfied.

Counterfactual Welfare Estimation under Equilibrium Multiplicity: As discussed above, we do not need to solve the equilibrium condition (50) for estimation of preference parameters and the ξ_v's (we do use the equilibrium conditions to predict the counterfactual π resulting from the policy experiment), and are therefore not affected by the multiplicity. However, when predicting counterfactual outcomes, we need to solve the equilibrium fixed point condition, and find solutions π_v's in the counterfactual scenario, e.g., a hypothetical subsidy rule to buy an ITN. Given the already estimated structural parameter values, the solutions π_v of the fixed point equation (50) in the counterfactual scenario can be computed for each v. If equilibrium multiplicity is anticipated, e.g., when the estimated α is larger than the threshold for contraction, we can compute these multiple solutions for each $v \in \{1, \ldots, \bar{v}\}$ through eye-balling since the number of solutions is at most three. This can be done by drawing a graph of the LHS of (50) and checking points on the graph that intersect the 45 degree line (or the graph for a least squares objective function as in (53), presented in Figure 2). The number of villages is eleven in our application and eye-balling is not difficult. Then, given multiple solutions, a bound for average welfare can be computed for each solution, and one can report multiple (at most three) bounds of them or a single union of the multiple intervals.

6 EMPIRICAL CONEXT AND DATA

Our empirical application concerns the provision of anti-malarial bednets. Malaria is a life-threatening parasitic disease transmitted from human to human through mosquitoes. In 2019, an estimated 229 million cases of malaria occurred worldwide, with 90% of the cases in sub-Saharan Africa (WHO, 2017). The main tool for malaria control in sub-Saharan Africa is the use of insecticide treated bednets. Regular use of a bednet reduces overall child mortality by around 18 percent and reduces morbidity for the entire population (Lengeler, 2004). However, at \$6 or more a piece, bednets are unaffordable for many households, and to palliate the very low coverage levels observed in the mid-2000s, public subsidy schemes were introduced in numerous countries in the last 15 years. Our empirical exercise is designed to evaluate such subsidy schemes not just in terms of their effectiveness in promoting bednet adoption, but also their impact on individual welfare and deadweight loss. Based on our discussion in Section 4, we focus on two main sources of spillover, viz. (a) a preference for conformity, and (b) a concern that mosquitoes will be deflected to oneself when neighbors protect themselves. Both will generate a positive effect of the aggregate adoption rate on one's own adoption decision, but they have different implications for the welfare impact of a price subsidy policy.

Experimental Design: We exploit data from a 2007 randomized bednet pricing experiment conducted in eleven villages of Western Kenya, where malaria is transmitted year-round. In each village, a list of 150 to 200 households was compiled from school registers, and households on the list were randomly assigned to a price at which they could purchase a long-lasting ITN, a new, highly effective type of antimalarial bednet. After the random assignment had been performed in office, trained enumerators visited each sampled household to administer a baseline survey. At the end of the interview, the household was given a voucher for one long-lasting ITN at the randomly assigned price level. The amount of subsidy (for those who received any) varied from 40% to 100% of the market price in two villages, and from 40% to 90% in the remaining 9 villages; there were 22 corresponding final prices faced by households, ranging from 0 to 300 Ksh (US $5.50), where 300 Ksh would be the non-subsidized sale price. Vouchers could be redeemed within three months at participating local retailers.

Data: We use data on bednet adoption as observed from coupon redemption and verified acquisition through a follow-up survey. We also use data on baseline household characteristics measured during the baseline survey. The three main baseline characteristics we consider are wealth (the combined value of all durable and animal assets owned by the household); the number of children under ten years old; and the education level of the female head of household.

Nonparticipating Households: While all households in a given village were potentially interacting, our sample does not cover all village members. This can potentially cause a problem since selected households might interact with non-selected ones. However, at the time of the experiment, non-selected households did not have the opportunity to buy a long-lasting ITN, so the outcome variable A for such households is zero, whose conditional expectations are zero as well. Thus, in our specification, even if we allow for interactions among all the village members, it is easy to do the necessary adjustments in the empirics, viz. replace Π_{vh} in (27) by

$$\check{\Pi}_{vh} = \left(\tfrac{N_v - 1}{\check{N}_v - 1}\right)\tfrac{1}{N_v - 1}\sum\nolimits_{1 \le k \le N_v;\ k \ne h} E[A_{vk}|\mathcal{I}_{vh}] = \left(\tfrac{N_v - 1}{\check{N}_v - 1}\right)\Pi_{vh}, \qquad (52)$$

where $\check{N}_v$ equals the total number of households in the village, and N_v those participating in the game. In our empirical setting, this ratio is about 0.8 for each village, and we apply this adjustment throughout the empirical analysis.

7 EMPIRICAL SPECIFICATION AND RESULTS

We work with the linear index structure (34), where Y_{vh} is taken to be the household wealth, P_{vh} is the experimentally set price faced by the household, Π_{vh} is the observed average adoption in the village. We also use additional controls, denoted by Z_{vh} below, that can potentially affect preferences and therefore the take-up of bednet, viz. presence of children under the age of ten, and years of education of the oldest female member of the household. A village-specific variable that could affect adoption is the extent of malaria exposure risk in the village. We measure this in our data from the response to the question: "Did anyone in your household have malaria in the past month?". Summary statistics are reported in Table 1, and their village averages are shown in Table 2, for each of the eleven villages in the data.

Our first set of results correspond to taking $F(\cdot)$ to be the probit CDF of $\eta_{vh} = \eta_{vh}^0 - \eta_{vh}^1$ (as in (4)), i.e. with no village-effects), and then our main results use the correlated random effects probit model that accounts for village-level unobservables. The marginal effects at mean are presented in Table 3, corresponding to both the probit model without village-effects and the CRE probit model that accounts for village-level unobservables. The fact that the price elasticity is very similar in the two specifications is expected, since price was exogenously assigned in the experiment, so accounting for village-fixed unobservables has no impact on the marginal effect.

It is evident from the table that the demand coefficient is negative and significantly different from zero (the averaged price elasticity is -0.12), and that bednet adoption in the village has a significant positive association with private adoption, conditional on price and other household characteristics, i.e. $\alpha > 0$ in our notation above. The social interaction coefficient α is 2.2 for the probit, which is less than the upper bound for the fixed point map to be a contraction (see discussion in Section 5.5). The effect of children is negative, likely reflecting that households with children had already invested in other anti-malarial steps, e.g., had bought a less effective traditional bednet prior to the experiment.

Next, we consider a hypothetical subsidy rule, where those with wealth less than τ are eligible to get the bednet for 50 KSh (83% subsidy), whereas those with wealth larger than τ get it for the price of 250 KSh (17% subsidy). Based on our preferred CRE probit model (which is used for all subsequent results, unless mentioned otherwise), we plot the predicted aggregate take-up of bednets corresponding to different income thresholds τ. In Figure 1, for each threshold τ, we plot the fraction of households eligible for a subsidy on the horizontal axis, and the predicted fraction choosing the bednet on the vertical axis, based on coefficients obtained by including (solid) and excluding (small dash) the spillover effect. The 45 degree line (large dash), showing the fraction eligible for the subsidy, is also plotted in the same figure for comparison.

It is evident from Figure 1 that ignoring spillovers leads to over-estimation of adoption at lower thresholds and underestimation at higher thresholds of eligibility. This happens because ignoring a covariate (here π) with positive impact on the outcome in prediction amounts to "smoothing" over values of π.

Having obtained these (uncompensated) effects, we now turn to calculating the demand and the mean compensating variation for a hypothetical subsidy scheme. We consider an initial situation where everyone faces a price of 250 KSh for the bednet, and a final situation where a bednet is offered for 50 KSh to households with wealth less than $\tau = 8000$ KSh (about the 27th percentile of the wealth distribution), and for the price of 250 KSh to those with wealth above that. The demand results are reported in Table 4, and the welfare results in Table 5. We perform these calculations village-by-village, and then aggregate across villages. To calculate these numbers, we first predict the bednet adoption when everyone is facing a price of 250 KSh, and then when eligibles face a price of 50 KSh and the rest stay at 250 KSh, giving us the equilibrium values of π_0 and π_1, respectively. In all such calculations with our data, we always detected a single solution to the fixed point π (i.e. a unique equilibrium) as can be seen from Figure 2, where we plot the squared difference between the RHS and the LHS of an empirical version of the fixed point equation (6) (with the additional covariate Z_{vh}), i.e.

$$\left[\pi_1 - \int \left[1\left\{y \leq \tau\right\} \hat{q}_1\left(p_1, y, z, \pi_1\right) + 1\left\{y > \tau\right\} \hat{q}_1\left(p_0, y, z, \pi_1\right)\right] d\hat{F}_{Y,Z}\left(y, z\right)\right]^2 \tag{53}$$

on the vertical axis, and π_1 on the horizontal axis, separately for each of the eleven villages, where $\hat{q}_1\left(p, y, z, \pi\right)$ is the predicted demand (choice probability) function at $\left(p, y, z, \pi\right)$. The globally convex nature of each objective function is evident from Figure 1. The minima are relatively close to each other around 0.15, except village 7 and 10, where it is larger. As for π_0, which minimizes $\left[\pi_0 - \int \hat{q}_1\left(p_1, y, z, \pi_0\right) d\hat{F}_{Y,Z}\left(y, z\right)\right]^2$, the objective function is also convex with a minimizing value close to zero in every village, reflecting that very few households would buy at this high price. These predicted values of π_0 and π_1 are used as inputs into the prediction of demand as the structural choice probability (2) and welfare as per Theorem 1 and Corollary 2.

The first row of Table 4 shows the pre-subsidy predicted demand by subsidy eligibility. In the second row, we calculate the predicted effect of the subsidy on demand, and break that up by the own price effect (Row 2) and the spillover effect (Row 3). The own effect is obtained by changing the price in accordance with the subsidy but keeping the village demand equal to the pre-subsidy value; the spillover effect is the difference between the overall effect and the own effect. It is clear that spillover effects on both eligibles and ineligibles are large in magnitude. In particular, the spillovers effect raises demand for ineligibles by an amount that nearly equals its pre-subsidy level.

In Table 5, we report welfare calculations with standard errors computed via the simple non-parametric bootstrap where households were resampled within each village in each bootstrap replication. In the first row, we report the welfare gain of the subsidy rule for eligibles, first assuming no spillovers and using a probit model without village-effects. In this case, we simply use the results of Bhattacharya (2015) to calculate the (point-identified) CV for eligibles as the price changes from 250 KSh to 50 KSh. This yields the value of welfare gain to be 52.589 KSh. As there is no spillover, the welfare change of ineligibles is zero by definition, and therefore the net welfare gain is simply the fraction eligible (0.27) times the CV for eligibles. This is reported in the third column of Table 5. The case with spillovers under probit and assuming $\alpha_1 \geq 0 \geq \alpha_0$ are reported in the second panel of Row 1 using (the negatives of) (13), (15) and (16) for eligibles, and using (18), (19) and (20) for ineligibles.

The 2nd-4th row present analogous results using CRE probit to control for fixed effects; the 2nd row does this for $\alpha_1 \geq 0 \geq \alpha_0$; the third row for $\alpha_1 \geq \alpha_0 \geq 0$, using a large upper limit of α_1 (and concurrently $\alpha_0 = \alpha_1 - \alpha > 0$) to proxy $\alpha_1 \nearrow \infty$, (cf. (21) and (22) above). Finally, the 4th row presents the overall bounds by taking union of the previous two cases.

Under $\alpha_1 \geq 0 \geq \alpha_0$, both specifications imply that ineligibles can suffer a large welfare *loss* due to the subsidy. This is because the subsidy facilitates usage for solely the eligibles, raising the equilibrium usage π in the village, but the ineligibles keep facing the high price, and thus a lower utility from not buying because π is now higher and $\alpha_0 \leq 0$. However, the few ineligibles who buy, despite the high price, get some welfare increase from a rise in the adoption rate, that explains the small upper bound corresponding to the case $\alpha_0 = 0$. The overall welfare gain aggregated over eligibles and ineligibles is reported in the column headed “Net Welfare Gain”.

Deadweight Loss: To compute the deadweight loss, we subtract the net welfare from the predicted subsidy expenditure. The latter equals the amount of subsidy (200 KSh) times the demand at the subsidized price 50 KSh of the eligibles. Thus the expression for DWL is given by

$$D = \int \begin{bmatrix} 200 \times 1\{y \leq \tau\} q_1(50, y, z, \pi_1) \\ -1\{y \leq \tau\} \mu^{\text{Elig}}(y, z, \pi_1, \pi_0) - 1\{y > \tau\} \mu^{\text{Inelig}}(y, z, \pi_1, \pi_0) \end{bmatrix} dF(y, z),$$

where y denotes wealth, z denotes other covariates, $q_1(50, y, z, \pi_1)$ denotes predicted demand at price 50 KSh including the effect of spillover, and μ^{Elig} and μ^{Inelig} refer to welfare *gain* for eligibles and ineligibles, respectively. Ignoring spillovers leads to the point-identified deadweight loss

$$D = \int \left[200 \times 1\{y \leq \tau\} \times q_1^{\text{No-spillover}}(50, y, z) - 1\{y \leq \tau\} \mu^{\text{No-spillover}}(y, z)\right] dF(y, z).$$

For the case $\alpha_1 \geq \alpha_0 \geq 0$, in the last-but-one row of Table 5, there is no welfare loss for anyone, since all spillover is positive, which explains the negative deadweight loss lower bound, i.e. an efficiency from subsidizing a positive externality.

These welfare and DWL numbers support the overall conclusion that accounting for spillovers can lead to much lower estimates of net welfare gain from the subsidy program and higher deadweight loss. Some of this difference arises from potential welfare loss suffered by ineligibles that is missed upon assuming no spillover, and some from the impact of including spillovers terms on the prediction of counterfactual purchase-rates (cf. Fig 1). Furthermore, the two cases $\alpha_1 \geq 0 \geq \alpha_0$ and $\alpha_1 \geq \alpha_0 \geq 0$, which are both consistent with the observed $\alpha = \alpha_1 - \alpha_0 > 0$, yield vastly different bounds on welfare, resulting in wide *overall* bounds on net welfare gain and deadweight loss that include zero (cf. last row of Table 5), which is the key substantive point of this paper.

Endogeneity: Price variation is exogenous in our application, since price was varied randomly by the experimenter. Indeed, it is still possible that wealth Y is correlated with $\boldsymbol{\eta}$, the unobserved determinants of bednet purchase (even conditionally on the village specific effects). However, experimental variation in price P implies also that P is independent of $\boldsymbol{\eta}$, given Y. Consequently, one can invoke the argument presented in Bhattacharya (2018, Section 3.1), and interpret the estimated choice probabilities and the corresponding welfare numbers as conditional on y, and then integrating with respect to the marginal distribution of y. This overcomes the problem posed by potentially endogenous income.

8 CONCLUSION

This paper develops tools for economic demand and welfare analysis in binary choice models with social interactions. The key finding is that under interactions, *welfare* distributions resulting from policy changes such as a price subsidy are generically not point-identified for given values of counterfactual aggregate demand, unlike the case *without* spillovers. This is true even when utility functions and distribution of unobserved heterogeneity are fully parametrized and there is a unique equilibrium. Non-identification results from the inability of standard choice data to distinguish between *different* underlying latent mechanisms, e.g. conforming motives, consumer learning, negative externalities etc., which produce the same aggregate social interaction coefficient, but have different welfare implications depending on which mechanism dominates. This feature is endemic to many practical settings that economists study, including the health-product adoption case examined here. Another prominent example is school-choice, where merit-based vouchers to attend a fee-paying selective school can create negative externalities by lowering the academic quality of the free local school via increased departure of high-achieving students. The resulting welfare implications cannot be calculated based solely on a Brock-Durlauf style empirical model of individual school-choice inclusive of a social interaction term. This is in contrast to models *without* social interaction, where choice probability functions have been shown to contain all the information required for welfare analysis. Nonetheless, we show that under standard linear index restrictions, welfare distributions can be bounded. Under some special and empirically untestable cases e.g. exactly symmetric spillovers effects or absence of negative externalities, these bounds shrink to point-identified values. Next, we develop methods of identification and consistent estimation for the structural utility parameters, required for prediction of counterfactual outcome and welfare bounds, when there is unobserved group-level heterogeneity possibly correlated with observable covariates. This is achieved via a novel latent factor modelling of unobserved group-effects and observed covariates, and developing a method of asymptotic analysis where the dimension of nuisance parameters, i.e. the group-effects whose magnitude may be unbounded, increases as the number of groups increase.

We apply our methods to an empirical setting of adoption of anti-malarial bednets, using data from a pricing experiment by Dupas (2014) in rural Kenya. We find that accounting for spillovers provides different predictions for demand and welfare resulting from hypothetical, means-tested subsidy rules. In particular, with *positive* interaction effects, predicted demand when including spillovers are *lower* for less generous eligibility criteria, compared to demand predicted by ignoring spillovers. At more generous eligibility thresholds, the conclusion reverses. As for welfare, if negative health externalities are present, then subsidy-ineligibles can suffer welfare loss due to increased use by subsidized buyers in the neighborhood; if solely conforming effects are present and there is no health-related externality, then welfare can improve.

The implication of these results for applied work is that under social interactions, welfare analysis of potential interventions requires more information regarding individual channels of spillovers than knowledge of solely the choice probability functions inclusive of a social interaction term. Belief-eliciting surveys, recording the reasons behind the subjects' actions, can provide a potential solution.

References

[1] BHATTACHARYA, D. (2008), A Permutation-based estimator for monotone index models. Econometric Theory 24, 795-807.

[2] BHATTACHARYA, D. (2015), Nonparametric welfare analysis for discrete choice. Econometrica 83, 617-649.

[3] BHATTACHARYA, D. (2018), Empirical welfare analysis for discrete choice: Some general results. Quantitative Economics 9, 571-615.

[4] BISIN, A., MORO, A. and TOPA, G., (2011), The empirical content of models with multiple equilibria in economies with social interactions (No. w17196). National Bureau of Economic Research.

[5] BLUNDELL, R. & POWELL, J. (2004), Endogeneity in nonparametric and semiparametric regression Models, in Advances in Economics and Econometrics, Cambridge University Press, Cambridge, UK.

[6] BROCK, W.A. & DURLAUF, S.N. (2001), Discrete choice with social spillover. The Review of Economic Studies 68, 235-60.

[7] BROCK, W.A. & DURLAUF, S.N. (2007), Identification of binary choice models with social interactions. Journal of Econometrics 140, 52-75.

[8] CHAMBERLAIN, G. (1980), Analysis of covariance with qualitative data. The Review of Economic Studies 47, 225-238.

[9] DALY, A. & ZACHARY, S. (1978), Improved multiple choice models, in Determinants of travel choice, 335-357.

[10] DOMINITZ, J. & SHERMAN R.P. (2005), Some convergence theory for iterative estimation procedures with an application to semiparametric estimation. Econometric Theory 21, 838-863.

[11] DUPAS, P. (2014), Short-run subsidies and long-run adoption of new health products: Evidence from a field experiment. Econometrica 82, 197-228.

[12] FERNANDEZ-VAL, I. & WEIDNER, M. (2016), Individual and time effects in nonlinear panel models with large N, T, Journal of Econometrics 192, 291-312.

[13] GAUTAM, S. (2018), Quantifying welfare effects in the presence of externalities: An ex-ante evaluation of a sanitation intervention, Mimeo.

[14] HAUSMAN, J.A, & NEWEY W. (2016), Individual heterogeneity and welfare. Econometrica 84, 1225-48.

[15] HAHN, J. & KUERSTEINER, G. (2011), Bias reduction for dynamic nonlinear panel models with fixed effects, Econometric Theory 27, 1152-1191.

[16] HAHN, J. & MOON, H.R. (2010), Panel data models with finite number of multiple equilibria, Econometric Theory 26, 863-881.

[17] HOTZ, V.J. & MILLIER, R.A. (1993), Conditional choice probabilities and the estimation of dynamic models. The Review of Economic Studies 60, 497-529.

[18] LENGELER, C. (2004), Insecticide-treated bed nets and curtains for preventing malaria. The Cochrane Library.

[19] MANSKI, C.F. (1993), Identification of endogenous social effects: The reflection problem. The Review of Economic Studies 60, 531-542.

[20] McFADDEN, D. & TRAIN, K., (2019), Welfare economics in product markets. Working paper, University of California, Berkeley.

[21] NEWEY, W.K. & McFADDEN, D. (1994), Large sample estimation and hypothesis testing, Ch. 36 in Handbook of Econometrics, Volume IV, (Ed. by R.F. Engle & D.L. McFadden)

[22] PESENDORFER, M. and SCHMIDT-DENGLER, P. (2008), Asymptotic least squares estimator for dynamic games. The Reviews of Economics Studies 75, 901-928.

[23] PASTORELLO, S., PATILEA, V. & RENAULT, E. (2003), Iterative and recursive estimation in structural nonadaptive models. Journal of Business & Economic Statistics 21, 449-509.

[24] RUST, J. (1987), Optimal replacement of GMC bus engines: An Empirical model of Harold Zurcher. Econometrica 55, 999-1033.

[25] DE PAULA, A. (2017), Econometrics of network models. In Advances in economics and econometrics: Theory and applications, eleventh world congress (pp. 268-323), Cambridge University Press.

[26] SMALL, K. & ROSEN, H. (1981). Applied welfare economics with discrete choice models. Econometrica 49, 105-130.

[27] WHO (2020), World malaria report 2020, Geneva, World Health Organization, Licence: CCBY-NC-SA 3.0 IGO.

[28] WOOLDRIDGE, J.M. (2010), Econometric analysis of cross section and panel data. MIT press.

A Appendix

A.1 Proofs of Welfare Results

Proof of Theorem 1. The condition $|\alpha| \sup_{e\in\mathbb{R}} f_{\eta^0-\eta^1}(e) < 1$ guarantees that the maps on the RHS of (5) and (6) are contractions by exactly the same argument as in Subsection 5.5, and therefore by Bruower's fixed point theorem, the solutions to (5) and (6) are unique. Hence it follows by the argument following (5) and (6) that $\pi_1 > \pi_0$.

Next, note that since $\beta_0, \beta_1 > 0$, the LHS of (9) is strictly increasing in S, so the condition $S \le a$ is equivalent to

$$\begin{aligned}
&\max\left\{\delta_1 + \beta_1(y + a - p_1) + \alpha_1\pi_1 + \eta^1, \delta_0 + \beta_0(y+a) + \alpha_0\pi_1 + \eta^0\right\} \\
&\ge \max\left\{\delta_1 + \beta_1(y - p_0) + \alpha_1\pi_0 + \eta^1, \delta_0 + \beta_0 y + \alpha_0\pi_0 + \eta^0\right\}. \qquad (54)
\end{aligned}$$

If $a < p_1 - p_0 - \frac{\alpha_1}{\beta_1}(\pi_1 - \pi_0) < 0$, then each term on the LHS of (54) is smaller than the corresponding term on the RHS. If $a \ge \frac{\alpha_0}{\beta_0}(\pi_0 - \pi_1) > 0$, then each term on the LHS is larger than the corresponding term on the RHS.[26] This gives us the support of S:

$$\Pr(S \le a) = \begin{cases} 0, & \text{if } a < p_1 - p_0 - \frac{\alpha_1}{\beta_1}(\pi_1 - \pi_0), \\ 1, & \text{if } a \ge \frac{\alpha_0}{\beta_0}(\pi_0 - \pi_1). \end{cases}$$

Now consider the intermediate case where

$$a \in [\underbrace{p_1 - p_0 - \frac{\alpha_1}{\beta_1}(\pi_1 - \pi_0)}_{<0}, \underbrace{\frac{\alpha_0}{\beta_0}(\pi_0 - \pi_1)}_{>0}).$$

In this case, the first term on LHS of (54) is larger than first term on RHS for all η_1, and the second term on LHS of (54) is smaller than the second term on the RHS for all η_0, and thus (54) is equivalent to

$$\begin{aligned}
&\delta_1 + \beta_1(y + a - p_1) + \alpha_1\pi_1 + \eta^1 \ge \delta_0 + \beta_0 y + \alpha_0\pi_0 + \eta^0 \\
\Leftrightarrow\ &\delta_1 + \beta_1(y + a - p_1) + \alpha_1\pi_0 + \alpha_1(\pi_1 - \pi_0) + \eta^1 \ge \delta_0 + \beta_0 y + \alpha_0\pi_0 + \eta^0. \qquad (55)
\end{aligned}$$

Thus, for any given α_1, we have that the probability of (55) reduces to

$$\begin{aligned}
&F(c_0 + \alpha_1(\pi_1 - \pi_0) + c_1(p_1 - a) + c_2 y + \alpha\pi_0) \\
&= q_1(p_1 - a, y, \pi_0 + \frac{\alpha_1}{\alpha}(\pi_1 - \pi_0)). \qquad (56)
\end{aligned}$$

■

[26] Note that the above reasoning also helps establish existence of a solution to (9). We know from above that for $S < p_1 - p_0 - \frac{\alpha_1}{\beta_1}(\pi_1 - \pi_0)$, the LHS of (9) is strictly smaller than the RHS, and for $S \ge \frac{\alpha_0}{\beta_0}(\pi_0 - \pi_1)$, the LHS of (9) is strictly larger than the RHS. By continuity, and the intermediate value theorem, it follows that there must be at least one S where (9) holds with equality.

Proof of Theorem 2. We have from (54) that $\Pr(S \leq a)$ equals

$$\begin{aligned} &\max\left\{\delta_1 + \beta_1 (y + a - p_1) + \alpha_1 \pi_1 + \eta^1, \delta_0 + \beta_0 (y + a) + \alpha_0 \pi_1 + \eta^0\right\} \\ &\geq \max\left\{\delta_1 + \beta_1 (y - p_0) + \alpha_1 \pi_0 + \eta^1, \delta_0 + \beta_0 y + \alpha_0 \pi_0 + \eta^0\right\}. \end{aligned}$$

Now, there are 2 cases to consider. If $p_1 - p_0 - \frac{\alpha_1}{\beta_1}(\pi_1 - \pi_0) < -\frac{\alpha_0}{\beta_0}(\pi_1 - \pi_0)$, then $Pr(S \leq a)$ reduces to (10). However, because the support is entirely negative now (since $-\frac{\alpha_0}{\beta_0}(\pi_1 - \pi_0) \leq 0$), mean welfare is given by $E(S) = -\int_{p_1 - p_0 - \frac{\alpha_1}{\beta_1}(\pi_1 - \pi_0)}^{\frac{\alpha - \alpha_1}{\beta_0}(\pi_1 - \pi_0)} F_S(a)\, da$ which equals

$$\begin{aligned} &-\int_{p_1 - p_0 - \frac{\alpha_1}{\beta_1}(\pi_1 - \pi_0)}^{\frac{\alpha - \alpha_1}{\beta_0}(\pi_1 - \pi_0)} q_1\left(p_1 - a, y, \pi_0 + \frac{\alpha_1}{\alpha}(\pi_1 - \pi_0)\right) da \\ &= -\int_{p_1 - \frac{\alpha - \alpha_1}{\beta_0}(\pi_1 - \pi_0)}^{p_0 + \frac{\alpha_1}{\beta_1}(\pi_1 - \pi_0)} q_1\left(p, y, \pi_0 + \frac{\alpha_1}{\alpha}(\pi_1 - \pi_0)\right) dp. \end{aligned}$$

The 2nd case is where $p_1 - p_0 - \frac{\alpha_1}{\beta_1}(\pi_1 - \pi_0) > -\frac{\alpha_0}{\beta_0}(\pi_1 - \pi_0)$, then for $a \in [-\frac{\alpha_0}{\beta_0}(\pi_1 - \pi_0), p_1 - p_0 - \frac{\alpha_1}{\beta_1}(\pi_1 - \pi_0)]$, we have that $S^{\text{Elig}} \leq a$ is equivalent to

$$\delta_0 + \beta_0 (y + a) + \alpha_0 \pi_1 + \eta^0 \geq \delta_1 + \beta_1 (y - p_0) + \alpha_1 \pi_0 + \eta^1,$$

whose probability equals

$$\begin{aligned} \eta^0 - \eta^1 &\geq \delta_1 - \delta_0 + (\beta_1 - \beta_0) y - \beta_1 p_0 - \beta_0 a + \alpha_1 \pi_0 - \alpha_0 \pi_1 \\ &= c_0 + c_1 (p_0 + a) + c_2 (y + a) + \alpha_1 \pi_0 + (\alpha - \alpha_1) \pi_1 \\ &= 1 - q_1\left(p_0 + a, y + a, \pi_1 - \frac{\alpha_1}{\alpha}(\pi_1 - \pi_0)\right). \end{aligned}$$

Thus we get

$$\begin{aligned} &\Pr\left(S^{\text{Elig}} \leq a\right) \\ &= \begin{cases} 0, & \text{if } a < -\frac{\alpha_0}{\beta_0}(\pi_1 - \pi_0), \\ 1 - q_1\left(p_0 + a, y + a, \pi_1 - \frac{\alpha_1}{\alpha}(\pi_1 - \pi_0)\right), & \text{if } -\frac{\alpha_0}{\beta_0}(\pi_1 - \pi_0) \leq a < p_1 - p_0 - \frac{\alpha_1}{\beta_1}(\pi_1 - \pi_0), \\ 1, & \text{if } a \geq p_1 - p_0 - \frac{\alpha_1}{\beta_1}(\pi_1 - \pi_0). \end{cases} \end{aligned} \tag{57}$$

Using that $E(S) = -\int_{-\frac{\alpha_0}{\beta_0}(\pi_1 - \pi_0)}^{p_1 - p_0 - \frac{\alpha_1}{\beta_1}(\pi_1 - \pi_0)} F_S(a)\, da$, we get from (57) that

$$\begin{aligned} E\left(S^{\text{Elig}}\right) &= -\int_{-\frac{\alpha_0}{\beta_0}(\pi_1 - \pi_0)}^{p_1 - p_0 - \frac{\alpha_1}{\beta_1}(\pi_1 - \pi_0)} \left(1 - q_1\left(p_0 + a, y + a, \pi_1 - \frac{\alpha_1}{\alpha}(\pi_1 - \pi_0)\right)\right) da \\ &= -\int_{p_0 + \frac{\alpha - \alpha_1}{\beta_0}(\pi_1 - \pi_0)}^{p_1 - \frac{\alpha_1}{\beta_1}(\pi_1 - \pi_0)} \left(1 - q_1\left(p, y + p - p_0, \pi_1 - \frac{\alpha_1}{\alpha}(\pi_1 - \pi_0)\right)\right) dp. \end{aligned} \tag{58}$$

■

A.2 Proofs of Equilibrium Results

C2 (i) For each v, the sequence $\{(W_{vh}, \boldsymbol{u}_{vh})\}_{h=1}^{N_v}$ is I.I.D. conditionally on (d_v, e_v). (ii) $\{\boldsymbol{u}_{vh}\}_{h=1}^{N_v}$ is independent of $\{W_{vh}\}_{h=1}^{N_v}$ conditionally on (d_v, e_v).

Proof of Proposition 1. For notational simplicity, we write $\boldsymbol{\omega}_v := (d_v, e_v)$ in this proof. By the definition in (27), $\Pi_{vk} = \frac{1}{N_v-1}\sum_{1\le j\le N_v;\, j\ne k} E[A_{vj}|\mathcal{I}_{vk}]$. Since this is the conditional expectations given $\mathcal{I}_{vk} = (W_{vk}, \boldsymbol{u}_{vk}, \boldsymbol{\omega}_v)$, we can write (v,k)'s belief as

$$\Pi_{vk} = g_{vk}(W_{vk}, \boldsymbol{u}_{vk}, \boldsymbol{\omega}_v),$$

using a function $g_{vk}(\cdot)$ which may depend on each index (v,k) but is non-random. Thus, plugging this expression of Π_{vk} into $A_{vk} = 1\{U_1(Y_{vk} - P_{vk}, \Pi_{vk}, \boldsymbol{\eta}_{vk}) \ge U_0(Y_{vk}, \Pi_{vk}, \boldsymbol{\eta}_{vk})\}$, we can write

$$A_{vk} = f_{vk}(W_{vk}, \boldsymbol{u}_{vk}, \boldsymbol{\omega}_v), \tag{59}$$

for some non-random function $f_{vk}(\cdot)$, where $W_{vk} = (Y_{vk}, P_{vk})$.

By **C2**, we have the two conditional independence restrictions: $(\boldsymbol{u}_{vh}, \boldsymbol{u}_{vk}) \perp W_{vh}|\boldsymbol{\omega}_v$and $\boldsymbol{u}_{vh} \perp \boldsymbol{u}_{vk}|\boldsymbol{\omega}_v$. These imply that

$$\text{“}\boldsymbol{u}_{vk} \perp W_{vh}|\,(\boldsymbol{u}_{vh}, \boldsymbol{\omega}_v)\ \ \&\ \boldsymbol{u}_{vk} \perp \boldsymbol{u}_{vh}|\boldsymbol{\omega}_v\text{”}\ \Leftrightarrow\ \boldsymbol{u}_{vk} \perp (W_{vh}, \boldsymbol{u}_{vh})|\boldsymbol{\omega}_v, \tag{60}$$

where we have used the following conditional independence relation: for random objects Q, R, and S,

$$\text{“}Q \perp R|\,(S, \boldsymbol{\omega}_v)\ \ \&\ Q \perp S|\boldsymbol{\omega}_v\text{” is equivalent to “}Q \perp (R,S)|\boldsymbol{\omega}_v\text{”}, \tag{61}$$

which is applied with $Q = \boldsymbol{u}_{vk}$, $R = W_{vh}$, and $S = \boldsymbol{u}_{vh}$. By the same argument, **C2** implies that

$$\begin{aligned} & (W_{vk}, W_{vh}) \perp (\boldsymbol{u}_{vk}, \boldsymbol{u}_{vh})|\boldsymbol{\omega}_v \ \&\ W_{vk} \perp W_{vh}|\boldsymbol{\omega}_v \\ \Rightarrow\quad & W_{vk} \perp (\boldsymbol{u}_{vk}, \boldsymbol{u}_{vh})|(W_{vh}, \boldsymbol{\omega}_v)\ \&\ W_{vk} \perp W_{vh}|\boldsymbol{\omega}_v, \end{aligned}$$

which is equivalent to

$$W_{vk} \perp (W_{vh}, \boldsymbol{u}_{vk}, \boldsymbol{u}_{vh})|\boldsymbol{\omega}_v. \tag{62}$$

Below, we denote by $E_v[\cdot]$ the conditional expectation operator given $\boldsymbol{\omega}_v$ ($= (d_v, e_v)$) and $E_v[\cdot|B] \equiv E[\cdot|\boldsymbol{\omega}_v, B]$. Given the above, we have

$$\begin{aligned} E[A_{vk}|\mathcal{I}_{vh}] &= E_v[f_{vk}(W_{vk}, \boldsymbol{u}_{vk}, \boldsymbol{\omega}_v)|W_{vh}, \boldsymbol{u}_{vh}] \\ &= \int E_v[f_{vk}(W_{vk}, \tilde{u}, \boldsymbol{\omega}_v)|W_{vh}, \boldsymbol{u}_{vh}, \boldsymbol{u}_{vk} = \tilde{u}]dF_{\boldsymbol{u}}^v(\tilde{u}|\boldsymbol{\omega}_v) \\ &= \int E_v[f_{vk}(W_{vk}, \tilde{u}, \boldsymbol{\omega}_v)]dF_{\boldsymbol{u}}^v(\tilde{u}|\boldsymbol{\omega}_v) \\ &= E_v[f_{vk}(W_{vk}, \boldsymbol{u}_{vk}, \boldsymbol{\omega}_v)] = E[A_{vk}|\boldsymbol{\omega}_v], \end{aligned}$$

where the first equality uses (59), the second and third equalities follow from (60) and (62), respectively, the fourth equality holds since $W_{vk} \perp \boldsymbol{u}_{vk}|\boldsymbol{\omega}_v$, completing the proof. ■

Proof of Proposition 2. Let

$$\bar{\pi}_{vk} = \bar{\pi}_{vk}(d_v, e_v) := E[A_{vk}|d_v, e_v] \text{ for } h = 1, \ldots, N_v, \tag{63}$$

where henceforth we suppress the dependence of $\bar{\pi}_{vk}$ on (d_v, e_v) for notational simplicity. By Proposition 1 and (30), we have

$$\Pi_{vh} = \bar{\Pi}_{vh} = \tfrac{1}{N_v - 1} \sum\nolimits_{1 \le k \le N_v;\ k \ne h} \bar{\pi}_{vk}. \tag{64}$$

Given these, we can write

$$\bar{\pi}_{vh} = E\left[1\left\{ \begin{array}{c} U_1(Y_{vh} - P_{vh}, \tfrac{1}{N_v-1}\sum_{1 \le k \le N_v;\ k \ne h} \bar{\pi}_{vk}, \boldsymbol{\eta}_{vh}) \\ \ge U_0(y, \tfrac{1}{N_v-1}\sum_{1 \le k \le N_v;\ k \ne h} \bar{\pi}_{vk}, \boldsymbol{\eta}_{vh}) \end{array} \right\} \middle| d_v, e_v \right], \quad h = 1, \ldots, N_v. \tag{65}$$

We can easily see that if a symmetric solution to the system of N_v equations in (65) exists uniquely, then that of (31) in terms of $\{\bar{\Pi}_{vh}\}_{h=1}^{N_v}$ also exists uniquely (vice versa; note that $\bar{\pi}_{vh} = \sum_{k=1}^{N_v} \bar{\Pi}_{vk} - (N_v - 1)\,\bar{\Pi}_{vh}$ by (64)). Therefore, we investigate (65).

Corresponding to (65), define an N_v-dimensional vector-valued function of $\boldsymbol{r} = (r_1, r_2, \ldots, r_{N_v}) \in [0,1]^{N_v}$ as

$$\mathcal{M}_v(\boldsymbol{r}) := \left(m_v(\tfrac{1}{N_v-1}\sum\nolimits_{k \ne 1} r_k), \ldots, m_v(\tfrac{1}{N_v-1}\sum\nolimits_{k \ne N_v} r_k) \right),$$

where we write $\sum_{1 \le k \le N_v;\ k \ne h} = \sum_{k \ne h}$, and the metric in the domain and range spaces of $\mathcal{M}_v$ is defined as

$$||\boldsymbol{s} - \tilde{\boldsymbol{s}}||_\infty := \max_{1 \le h \le N_v} |s_h - \tilde{s}_h|,$$

for any $\boldsymbol{s} = (s_1, \ldots, s_{N_v})$, $\tilde{\boldsymbol{s}} = (\tilde{s}_1, \ldots, \tilde{s}_{N_v}) \in [0,1]^{N_v}$ (note that both the spaces are taken to be $[0,1]^{N_v}$). Given these definitions of $\mathcal{M}_v(\boldsymbol{r})$ and the metric, we can easily show that the contraction property of $m_v(\cdot)$ carries over to $\mathcal{M}_v(\cdot)$, i.e.,

$$\|\mathcal{M}_v(\boldsymbol{r}) - \mathcal{M}_v(\tilde{\boldsymbol{r}})\|_\infty \le \rho\, \|\boldsymbol{r} - \tilde{\boldsymbol{r}}\|_\infty,$$

which implies that there exists a unique solution $\boldsymbol{r}^*$ to the (N_v-dimensional) vector-valued equation:

$$\boldsymbol{r} = \mathcal{M}_v(\boldsymbol{r}). \tag{66}$$

Now, consider the following scalar-valued equation $r = m_v(r)$. By the contraction property (33), it has a unique solution. Denote this solution by $\bar{r}^* \in [0,1]$. By the definition of $\mathcal{M}_v(\cdot)$, the vector $\bar{\boldsymbol{r}}^* = (\bar{r}^*, \ldots, \bar{r}^*) \in [0,1]^{N_v}$ must be a solution to (66). Then, by the uniqueness of the solution to (66), this $\bar{\boldsymbol{r}}^*$ must be a unique solution, which is a set of symmetric beliefs. The proof is therefore complete. ■

A.3 Proof of Consistency

To verify the consistency of our first probit estimator, we introduce the following functions for each $v \in \{1, \ldots, \bar{v}\}$:

$$\hat{Q}_v(\boldsymbol{c}, \gamma_v) := \frac{1}{N_v} \sum\nolimits_{v=1}^{N_v} \mathcal{L}_{vh}(\boldsymbol{c}, \gamma_v) \quad \text{and} \quad Q_v(\boldsymbol{c}, \gamma_v) := E_{\boldsymbol{\omega}_v}[\mathcal{L}_{vh}(\boldsymbol{c}, \gamma_v)], \tag{67}$$

where

$$\mathcal{L}_{vh}(\boldsymbol{c}, \gamma_v) = A_{vh} \log \Phi\left(W_{vh}\boldsymbol{c}' + \gamma_v\right) + (1 - A_{vh}) \log\left(1 - \Phi\left(W_{vh}\boldsymbol{c}' + \gamma_v\right)\right).$$

Given this definition of $\hat{Q}_v$, we can write the objective function of the first probit estimator as

$$\hat{Q}(\boldsymbol{c}, \gamma_1, \ldots, \gamma_{\bar{v}}) = \frac{1}{N} \sum\nolimits_{v=1}^{\bar{v}} \sum\nolimits_{h=1}^{N_v} \mathcal{L}_{vh}(\boldsymbol{c}, \gamma_v) = \frac{1}{\bar{v}} \sum\nolimits_{v=1}^{\bar{v}} \frac{\bar{v}N_v}{N} \hat{Q}_v(\boldsymbol{c}, \gamma_v). \tag{68}$$

Note that each of the weights $\bar{v}N_v/N$ does not degenerate. This is because, under our assumption that $N_v = r_v N_0$ with $\underline{r} \le r_v \le \bar{r}$ $(0 < \underline{r} \le \bar{r} < \infty)$ for any v and $N = \sum_{v=1}^{\bar{v}} r_v N_0$, we have

$$0 < \underline{r}/\bar{r} \le \frac{\bar{v}N_v}{N} \le \bar{r}/\underline{r} < \infty.$$

Thus, each of $\hat{Q}_1, \ldots, \hat{Q}_{\bar{v}}$ has a non-negligible contribution to $\hat{Q}$ relative to others, so that the consistency of each of $\hat{\gamma}_1, \ldots, \hat{\gamma}_{\bar{v}}$ can be guaranteed.

The objective function $\hat{Q}(\boldsymbol{c}, \gamma_1, \ldots, \gamma_{\bar{v}})$ as well as its limit when $N_0 \to \infty$ and $\bar{v} \to \infty$ is strictly concave. Thus, in practice, there is no need to specify parameter spaces for $(\boldsymbol{c}, \gamma_1, \ldots, \gamma_{\bar{v}})$. However, for a rigorous theoretical derivation of the uniform convergence of $\hat{Q}_v$ to Q_v and identification of $(\boldsymbol{c}^*, \gamma_v^*)$ based on the limit Q_v, it is convenient to specify the parameter spaces. We let $\boldsymbol{\Upsilon_1}$ be a fixed compact set on $\mathbb{R}^{d_W}$ in which the true parameter $\boldsymbol{c}^*$ lies and $\Upsilon(\bar{v})$ be a compact interval on $\mathbb{R}$ defined as

$$\Upsilon(\bar{v}) := \left\{ |\gamma| \le s \;\middle|\; s = \sqrt{4(\sigma_e^*)^2 \log[\bar{v} (\log \bar{v})^t]} \right\}, \tag{69}$$

which grows when $\bar{v} \to \infty$, where $t \in (1/2, 1)$ is an arbitrary constant, introduced in Lemma 1. Note that we can treat $\{\gamma_v^*\}_{v=1}^{\bar{v}}$ as if these are parameters to be estimated since the law of covariates is given conditionally on $\boldsymbol{\omega}_v = (d_v, e_v)$ in **C2** and the law of $\{e_v\}_{v=1}^{\bar{v}}$ is independent of the rest of variables as supposed in **C1** and (39).

We impose the following additional conditions:

CR2 (i) Let $\gamma_v^* = c_0^* + \alpha^* \bar{\pi}_v + d_v \boldsymbol{\delta}^{*\prime} + e_v$ (the village-specific unobservable variable as defined in (36) and (37)). Suppose that the village specific factors $\{d_v\}$ satisfy $\sup_{v \ge 1} \|d_v\| \le C_d \in (0, \infty)$. (ii) The observable variables $\{W_{vh}\}$ satisfy $\sup_{\bar{v} \ge 1,\, h \ge 1} \|W_{vh}\| \le C_W \in (0, \infty)$. $\{e_v\}_{v=1}^{\bar{v}}$ is independent of $\{\left(\{W_{vh}\}_{h=1}^{N_v}, \varepsilon_{vh}, d_v\right)\}_{v=1}^{\bar{v}}$. (iii) $\bar{v}$ tends to ∞ with $\log \bar{v}$ being at most of polynomial order of N_0 as $N_0 \to \infty$ (i.e. there exists some $\bar{\kappa} > 0$ such that $\log \bar{v} \le N_0^{\bar{\kappa}}$ for any large N_0), where N_0 is introduced in (43).

(i)-(ii) of **CR2** suppose the uniform boundedness of the village specific factors d_v and covariates W_{vh}. While these may be easily relaxed at the cost of some additional conditions and notational complexity, we maintain them for simplicity. Condition **(iii)** on the growth rate of $\bar{v}$ is required for uniform convergence of $\hat{Q}_v$ to Q_v, which is a mild condition. It anyway has to be satisfied under the rate condition (46) for the consistency in Proposition 3.

Given these conditions, we use the following three lemmas based on which the consistency result, i.e. Proposition 3, is verified (proofs of the lemmas are provided in the next subsection):

Lemma 1 *Suppose that* $\{e_v\}_{v\geq 1}$ *is I.I.D. with* $N\left(0,(\sigma_e^*)^2\right)$ *with* $\sigma_e^* > 0$ *(as implied by* ***C1*** *and (39)) and* ***CR2 (i)*** *holds. Then, for each sample point* $\omega \in \bar{\Omega}$ *with* $\Pr\left[\bar{\Omega}\right] = 1$*, there exists some sufficiently large* $K\,(= K(\omega) \in [2,\infty))$ *such that for any* $\bar{v} \geq K$*,*

$$\max_{1\leq v\leq \bar{v}} |\gamma_v^*| \leq \sqrt{4(\sigma_e^*)^2 \log[\bar{v}\,(\log \bar{v})^t]},$$

where $t > 1/2$ *is an arbitrary constant that is independent of* ω *and* $\bar{v}$*.*

This lemma derives the almost sure uniform bound of $|\gamma_{\bar{v}}^*|$ as $\bar{v} \to \infty$, which can be understood as follows: For each sample point $\omega \in \bar{\Omega}$, we have an infinite sequence of realized real numbers $\{e_1, e_2, \dots\} = \{e_1(\omega), e_2(\omega), \dots\}$ and thus its corresponding sequence of $\{\gamma_1^*, \gamma_2^*, \dots\}$. For this realized sequence, we can always find some number $K(\omega)$ such that if $\bar{v} \geq K(\omega)$, then none of $\{|\gamma_1^*|, |\gamma_2^*|, \dots, |\gamma_{\bar{v}}^*|\}$ is larger than $\sqrt{4(\sigma_e^*)^2 \log[\bar{v}\,(\log \bar{v})^t]}$. This upper bound depends on the standard error σ_e^* of e_v: we have a smaller maximum of $|\gamma_v^*|$ when the variation σ_e^* is smaller. While the support of γ_v^* is unbounded, which consists of a normally distributed error e_v, Lemma 1 allows us to restrict its parameter space as in (69). The following two lemmas effectively use this restricted-parameter-space result for γ_v^*:

Lemma 2 (Identification of $(\mathbf{c}^*, \gamma_1^*, \dots, \gamma_{\bar{v}}^*)$**)** *Suppose that* $\{e_v\}_{v\geq 1}$ *is I.I.D. with* $N\left(0,(\sigma_e^*)^2\right)$ *and* ***CR1 (i)-(ii)*** *and* ***CR2*** *hold. Let* $Q_v(\boldsymbol{c},\gamma_v)$ *be the function defined in (67). Then, for each* $\epsilon_1 > 0$*, there exists some constant* $C_Q > 0$ *(independent of* $\bar{v}$*) such that as* $\bar{v} \to \infty$*,*

$$\inf_{1\leq v\leq \bar{v}} \left[Q_v(\boldsymbol{c}^*, \gamma_v^*) - \sup_{(\boldsymbol{c},\gamma_v)\in \Upsilon_{\boldsymbol{c}} \times \Upsilon(\bar{v});\ \|(\boldsymbol{c},\gamma_v)-(\boldsymbol{c}^*,\gamma_v^*)\|\geq \epsilon_1} Q_v(\boldsymbol{c},\gamma_v) \right] \geq C_Q [\bar{v} (\log \bar{v})]^{-2(\sigma_e^*)^2} \tag{70}$$

holds with probability 1*.*

Lemma 3 (Uniform convergence) *Suppose that* $\{e_v\}_{v\geq 1}$ *is I.I.D. with* $N\left(0,(\sigma_e^*)^2\right)$ *and* ***C2, CR1****, and* ***CR2*** *hold. Let* $\hat{Q}_v$ *and* Q_v *be functions defined in (67). Then, as* $N_0 \to \infty$*,*

$$\max_{1\leq v\leq \bar{v}} \sup_{(\boldsymbol{c},\gamma_v)\in \Upsilon_1 \times \Upsilon(\bar{v})} \left| \hat{Q}_v(\boldsymbol{c},\gamma_v) - Q_v(\boldsymbol{c},\gamma_v) \right| = O_p\Big(\sqrt{(\log \bar{v})^3 (\log N_0) \big/ N_0}\Big).$$

Lemma 2 shows the identification of $(\boldsymbol{c}^*, \gamma_v^*)$ through the objective function Q_v. For each v, we have $Q_v(\boldsymbol{c}^*, \gamma_v^*) - Q_v(\boldsymbol{c}, \gamma_v) > 0$ for any $(\boldsymbol{c}, \gamma_v) \neq (\boldsymbol{c}^*, \gamma_v^*)$, which is a standard result easily shown under the rank condition on $(W_{vh}, 1)$ and the probit specification; however, if the domain of γ_v^* were not restricted, the uniform lower bound $Q_v(\boldsymbol{c}^*, \gamma_v^*) - Q_v(\boldsymbol{c}, \gamma_v)$ over v would be zero, which would not allow us to show the consistency result.

Lemma 3 derives the uniform convergence of $\hat{Q}_v(\boldsymbol{c}, \gamma_v)$. This is based on the Bernstein exponential inequality for I.I.D. sequences. The restriction of the parameter space $\Upsilon(\bar{v})$ as in (69), which is compact for each $\bar{v}$, can also facilitate the verification of the uniform convergence.

Given these results, we are now ready to prove consistency:

Proof of Proposition 3. Note that $\hat{Q}(\boldsymbol{c}, \gamma_1, \ldots, \gamma_{\bar{v}})$ can be written as in (68) and this is a weighted average of $\hat{Q}_v(\boldsymbol{c}, \gamma_v)$ with uniformly non-degenerate weights, $\frac{\bar{v}N_v}{N} \in \left[\frac{\underline{r}}{\bar{r}}, \frac{\bar{r}}{\underline{r}}\right]$ for any v. Let $\epsilon_1 > 0$ be an arbitrary constant and $\epsilon_2 = \epsilon_2(\bar{v}) := \frac{C_Q}{2}[\bar{v}(\log \bar{v})]^{-2(\sigma_e^*)^2} > 0$, where $C_Q > 0$ is a constant given in Lemma 2. We look at

$$\begin{aligned}
&\sup_{\boldsymbol{c} \in \Upsilon_{\boldsymbol{c}};\ \forall v,\ \gamma_v \in \Upsilon(\bar{v}) \text{ s.t. } \exists v,\ |(\boldsymbol{c}, \gamma_v) - (\boldsymbol{c}^*, \gamma_v^*)| > \epsilon_1} \hat{Q}(\boldsymbol{c}, \gamma_1, \ldots, \gamma_{\bar{v}}) \\
&\leq \sup_{\boldsymbol{c} \in \Upsilon_{\boldsymbol{c}};\ \forall v,\ \gamma_v \in \Upsilon(\bar{v}) \text{ s.t. } \exists v,\ |\gamma_v - \gamma_v^*| > \epsilon_1} \frac{1}{\bar{v}} \sum\nolimits_{v=1}^{\bar{v}} \frac{\bar{v}N_v}{N} Q_v(\boldsymbol{c}, \gamma_v) + O_p\left(\sqrt{(\log \bar{v})^3 (\log N_0) \big/ N_0}\right) \\
&< \frac{1}{\bar{v}} \sum\nolimits_{v=1}^{\bar{v}} \frac{\bar{v}N_v}{N} Q_v(\boldsymbol{c}^*, \gamma_v^*) - C_Q[\bar{v}(\log \bar{v})]^{-2(\sigma_e^*)^2} + O_p\left(\sqrt{(\log \bar{v})^3 (\log N_0) \big/ N_0}\right) \\
&\leq \frac{1}{\bar{v}} \sum\nolimits_{v=1}^{\bar{v}} \frac{\bar{v}N_v}{N} \hat{Q}_v(\boldsymbol{c}^*, \gamma_v^*) - C_Q[\bar{v}(\log \bar{v})]^{-2(\sigma_e^*)^2} + O_p\left(\sqrt{(\log \bar{v})^3 (\log N_0) \big/ N_0}\right) \\
&\leq \frac{1}{\bar{v}} \sum\nolimits_{v=1}^{\bar{v}} \frac{\bar{v}N_v}{N} \hat{Q}_v(\boldsymbol{c}^*, \gamma_v^*) - \epsilon_2 \leq \frac{1}{\bar{v}} \sum\nolimits_{v=1}^{\bar{v}} \frac{\bar{v}N_v}{N} \hat{Q}_v(\hat{\boldsymbol{c}}, \hat{\gamma}_v) - \epsilon_2,
\end{aligned}$$

where the first and third inequalities follow from the uniform convergence of $\hat{Q}_v$ to Q_v (derived in Lemma 3); the second follows from Lemma 2, the fourth follows from the definition of ϵ_2. and the rate condition in (46) (i.e. $\sqrt{(\log \bar{v})^3 (\log N_0) \big/ N_0}$ is much smaller than $C_Q[\bar{v}(\log \bar{v})]^{-2(\sigma_e^*)^2}$); and the last is due to the definition of $(\hat{\boldsymbol{c}}, \hat{\gamma}_1, \ldots, \hat{\gamma}_{\bar{v}})$, which is the maximizer of $\hat{Q}$. Now, we have verified that for any $(\boldsymbol{c}, \gamma_1, \ldots, \gamma_v) \in \Upsilon_{\boldsymbol{c}} \times \Upsilon_\gamma^{\bar{v}} \times \cdots \times \Upsilon_\gamma^{\bar{v}}$ with $|(\boldsymbol{c}, \gamma_v) - (\boldsymbol{c}^*, \gamma_v^*)| > \epsilon_1$ for some v, it holds that $\hat{Q}(\boldsymbol{c}, \gamma_1, \ldots, \gamma_{\bar{v}}) < \frac{1}{\bar{v}} \sum_{v=1}^{\bar{v}} \frac{\bar{v}N_v}{N} \hat{Q}_v(\hat{\boldsymbol{c}}, \hat{\gamma}_v) = \hat{Q}(\hat{\boldsymbol{c}}, \hat{\gamma}_1, \ldots, \hat{\gamma}_{\bar{v}})$ with probability approaching 1. This implies that $\hat{\gamma}_v$ has to satisfy $|(\hat{\boldsymbol{c}}, \hat{\gamma}_v) - (\boldsymbol{c}^*, \gamma_v^*)| \leq \epsilon_1$ for any v with probability approaching 1, leading to the desired result of the uniform convergence of $(\hat{\boldsymbol{c}}, \hat{\gamma}_v)$, completing the proof . ■

A.4 Proofs of Auxiliary Lemmas for Consistency

Proof of Lemma 1.. We first derive a tail probability bound of e_v: for any $y > 0$,

$$\Pr[|e_v| > y] = \Pr[|e_v| / \sigma_e^* > y / \sigma_e^*] \leq 2 \frac{1}{\sqrt{2\pi}(y/\sigma_e^*)} \exp\left\{-\frac{(y/\sigma_e^*)^2}{2}\right\},$$

where the inequality follows from a tail bound of the standard normal CDF: $1-\Phi(x) \leq \frac{1}{x}\phi(x)$ for any $x>0$ (p. 112 of Karatzas and Shreve, 1991). This implies that

$$\Pr\left[\max_{1\leq v\leq \bar{v}} |e_v| > y\right] \leq \sum\nolimits_{v=1}^{\bar{v}} \Pr\left[|e_v| > y\right] \leq \bar{v}\frac{\sigma_e^*}{\sqrt{\pi/2}y} \exp\left\{-\frac{(y/\sigma_e^*)^2}{2}\right\}.$$

Letting $y=\sqrt{4(\sigma_e^*)^2 \log[\bar{v}(\log\bar{v})^{t_0}]}$ for an arbitrary $t_0>1/2$, we have for $\bar{v}\geq 2$,

$$\Pr\left[\max_{1\leq v\leq \bar{v}} |e_v| > \sqrt{4(\sigma_e^*)^2 \log[\bar{v}(\log\bar{v})^{t_0}]}\right] \leq \frac{1}{\sqrt{\pi/2}\sqrt{4\log[\bar{v}(\log\bar{v})^{t_0}]}} \frac{1}{\bar{v}(\log\bar{v})^{t_0}}$$

and thus

$$\sum\nolimits_{\bar{v}=2}^{\infty} \Pr\left[\max_{1\leq v\leq \bar{v}} |e_v| > \sqrt{4(\sigma_e^*)^2 \log[\bar{v}(\log\bar{v})^{t_0}]}\right] < \infty,$$

which holds since $\sum_{m=2}^{\infty}\frac{1}{m(\log m)^p}<\infty$ for any $p>1$. By the Borel-Cantelli lemma, this implies that the event " $\max_{1\leq v\leq \bar{v}} |e_v| \geq \sqrt{4(\sigma_e^*)^2 \log[\bar{v}(\log\bar{v})^{t_0}]}$" may happen at most finitely many times. Thus, for each sample point $\omega\in\bar{\Omega}$ with $\Pr\left[\bar{\Omega}\right]=1$, there exists some (sufficiently large) $K\,(=K(\omega)\geq 2)$ such that for any $\bar{v}\geq K$,

$$\max_{1\leq v\leq \bar{v}} |e_v| \leq \sqrt{4(\sigma_e^*)^2 \log[\bar{v}(\log\bar{v})^{t_0}]}. \tag{71}$$

Define $C_\gamma^0 := |c_0^*|+|\alpha^*|+C_d\|\boldsymbol{\delta}^*\|$ (C_d is the upper bound of d_v introduced in **CR2 (i)**). Then, we have $|\gamma_v^*| = \left|c_0^*+\alpha^*\bar{\pi}_v+d_v\boldsymbol{\delta}^{*\prime}+e_v\right| \leq C_\gamma^0+|e_v|$. For another arbitrary constant $t>t_0\,(>1/2)$,

$$\max_{1\leq v\leq \bar{v}} |\gamma_v^*| \leq C_\gamma^0 + \sqrt{4(\sigma_e^*)^2 \log[\bar{v}(\log\bar{v})^{t_0}]} \leq \sqrt{4(\sigma_e^*)^2 \log[\bar{v}(\log\bar{v})^{t}]},$$

for sufficiently large $\bar{v}$ (if necessary, we may re-define K introduced for (71) so that (71) and this inequalities hold simultaneously). Now, the proof of Lemma 1 is complete. ■

Proof of Lemma 2. For notational simplicity, let $\vartheta_v := (\boldsymbol{c},\gamma_v)$ and $Q_v(\boldsymbol{c},\gamma_v)=Q_v(\vartheta_v)$. We also write $\vartheta_v^* := (\boldsymbol{c}^*,\gamma_v^*)$ and define $\rho_{vh}(\vartheta_v^*,\vartheta_v) := W_{vh}(\boldsymbol{c}^*)' - W_{vh}\boldsymbol{c}' + \gamma_v^* - \gamma_v$. Noting the "single index" structure of the model, we can write $\mathcal{L}_{vh}(\boldsymbol{c},\gamma_v) = \mathcal{L}_{vh}(\boldsymbol{c}^*, -\rho_{vh}(\boldsymbol{c}^*,\boldsymbol{c})+\gamma_v^*)$. Then, using the Taylor expansion,

$$\mathcal{L}_{vh}(\boldsymbol{c}^*,\gamma_v^*) - \mathcal{L}_{vh}(\boldsymbol{c},\gamma_v^*) = \rho_{vh}(\vartheta_v^*,\vartheta_v)\int_0^1 \partial_\gamma\mathcal{L}_{vh}(\boldsymbol{c}^*, -(1-\lambda)\rho_{vh}(\vartheta_v^*,\vartheta_v)+\gamma_v^*)\,d\lambda.$$

For computing the partial derivative $\partial_\gamma\mathcal{L}_{vh}$, we define a function κ_0 as

$$\kappa_0(x) := \frac{\phi(x)}{\Phi(x)[1-\Phi(x)]}.$$

Then,

$$\begin{aligned} Q_v(\vartheta_v^*) - Q_v(\vartheta_v) &= E_{\boldsymbol{\omega}_v}\left[\rho_{vh}(\boldsymbol{c}^*,\boldsymbol{c})\int_0^1 \partial_\gamma\mathcal{L}_{vh}(\boldsymbol{c}^*, -(1-\lambda)\rho_{vh}(\vartheta_v^*,\vartheta_v)+\gamma_v^*)\,d\lambda\right] \\ &= E_{\boldsymbol{\omega}_v}\Big[|\rho_{vh}(\vartheta_v^*,\vartheta_v)|\int_0^1 \kappa_0\left(W_{vh}(\boldsymbol{c}^*)' - (1-\lambda)\rho_{vh}(\vartheta_v^*,\vartheta_v)+\gamma_v^*\right) \\ &\quad\times\left|\Phi(W_{vh}(\boldsymbol{c}^*)'+\gamma_v^*) - \Phi(W_{vh}(\boldsymbol{c}^*)' - (1-\lambda)\rho_{vh}(\vartheta_v^*,\vartheta_v)+\gamma_v^*)\right|d\lambda\Big], \end{aligned} \tag{72}$$

where the absolute value signs can be given to $\rho_{vh}(\vartheta_v^*, \vartheta_v)$ and the last component of the integrand since $\kappa_0(x) > 0$ for any x and $\Phi(\cdot)$ is monotone (i.e., if $\rho_{vh}(\vartheta_v^*, \vartheta_v) > 0$,

$$\Phi(W_{vh}(\boldsymbol{c}^*)' + \gamma_v^*) - \Phi(W_{vh}(\boldsymbol{c}^*)' - (1-\lambda)\rho_{vh}(\vartheta_v^*, \vartheta_v) + \gamma_v^*) \geq 0 \quad \text{for any } \lambda \in [0,1];$$

and if $\rho_{vh}(\vartheta_v^*, \vartheta_v) < 0$, the inequality is reversed).

For deriving a lower bound of (72), we use the following inequalities:

$$\kappa_0(x) = \frac{\phi(x)}{\Phi(x)[1-\Phi(x)]} \geq \phi(1) \quad \text{for any } x \in \mathbb{R}; \tag{73}$$

$$|\Phi(y) - \Phi(x)| \geq |y - x|\,\phi(\max\{|y|, |x|\}) \quad \text{for any } y, x \in \mathbb{R}, \tag{74}$$

where $\phi(1) = (1/\sqrt{2\pi})\exp\{-1/2\}$, whose proofs are provided below. Using the uniform lower bound (73),

$$\begin{aligned}
\text{the RHS of (72)} \geq\ & E_{\boldsymbol{\omega}_v}\Big[|\rho_{vh}(\vartheta_v^*, \vartheta_v)|\,\phi(1) \\
& \times \int_0^{1/2} \left|\Phi\left(W_{vh}(\boldsymbol{c}^*)' + \gamma_v^*\right) - \Phi(W_{vh}(\boldsymbol{c}^*)' - (1-\lambda)\rho_{vh}(\vartheta_v^*, \vartheta_v) + \gamma_v^*)\right| d\lambda\Big] \\
\geq\ & E_{\boldsymbol{\omega}_v}\Big[|\rho_{vh}(\vartheta_v^*, \vartheta_v)|\,\phi(1) \\
& \times \left|\Phi\left(W_{vh}(\boldsymbol{c}^*)' + \gamma_v^*\right) - \Phi(W_{vh}(\boldsymbol{c}^*)' - \tfrac{1}{2}\rho_{vh}(\vartheta_v^*, \vartheta_v) + \gamma_v^*)\right|\Big] \\
\geq\ & E_{\boldsymbol{\omega}_v}\Big[|\rho_{vh}(\vartheta_v^*, \vartheta_v)|\,\phi(1) \\
& \times \left|\tfrac{1}{2}\rho_{vh}(\vartheta_v^*, \vartheta_v)\right|\phi\left(\max\left\{\left|W_{vh}(\boldsymbol{c}^*)' + \gamma_v^*\right|, \left|W_{vh}(\boldsymbol{c}^*)' - \tfrac{1}{2}\rho_{vh}(\vartheta_v^*, \vartheta_v) + \gamma_v^*\right|\right\}\right)\Big], \\
\geq\ & \frac{\phi(1)}{2}E_{\boldsymbol{\omega}_v}\left[|\rho_{vh}(\vartheta_v^*, \vartheta_v)|^2\right]\phi\left(C_W||\boldsymbol{c}^*|| + \frac{\epsilon_1}{2}(C_W+1) + |\gamma_v^*|\right),
\end{aligned} \tag{75}$$

where the second inequality holds since Φ is monotone and the difference between $W_{vh}(\boldsymbol{c}^*)' + \gamma_v^*$ and $W_{vh}(\boldsymbol{c}^*)' - (1-\lambda)\rho_{vh}(\vartheta_v^*, \vartheta_v) + \gamma_v^*$ is minimized at $\lambda = 1/2$ over $[0, 1/2]$; the third equality uses $|\Phi(y) - \Phi(x)| \geq |y - x|\,\phi(\max\{|y|, |x|\})$ in (73) with $y = W_{vh}(\boldsymbol{c}^*)' + \gamma_v^*$ and $x = W_{vh}(\boldsymbol{c}^*)' - \frac{1}{2}\rho_{vh}(\vartheta_v^*, \vartheta_v) + \gamma_v^*$ (note that $\phi(s)$ is decreasing for $s \geq 0$); and the last equality holds since

$$\max\{|y|, |x|\} \leq C_W||\boldsymbol{c}^*|| + \frac{\epsilon_1}{2}(C_W+1) + |\gamma_v^*|.$$

Recalling properties of quadratic forms (and noting that all vectors are defined as row vectors), we have

$$\begin{aligned}
E_{\boldsymbol{\omega}_v}\left[|\rho_{vh}(\vartheta_v^*, \vartheta_v)|^2\right] &= (\vartheta_v^*, \vartheta_v)\,E_{\boldsymbol{\omega}_v}\left[(W_{vh}, 1)'(W_{vh}, 1)\right](\vartheta_v^*, \vartheta_v)' \\
&\geq \left(\inf\nolimits_{v\geq 1}\lambda_v^{\min}\right)\|\vartheta_v^* - \vartheta_v\|^2 \geq \left(\inf\nolimits_{v\geq 1}\lambda_v^{\min}\right)\epsilon_1^2,
\end{aligned}$$

where $\lambda_v^{\min}$ is the minimum eigenvalue of the symmetric matrix $E_{\boldsymbol{\omega}_v}[W_{vh}'W_{vh}]$. From these, we can obtain

$$Q_v(\vartheta_v^*) - Q_v(\vartheta_v) \geq \frac{\phi(1)}{2}\left(\inf_{v\geq 1}\lambda_v^{\min}\right)\epsilon_1^2\phi\left(C_W||\boldsymbol{c}^*|| + \frac{\epsilon_1}{2}(C_W+1) + |\gamma_v^*|\right).$$

Noting that this lower bound is independent of $\vartheta_v = (\boldsymbol{c}, \gamma_v)$ and using the almost sure bound of $|\gamma_v^*|$ (Lemma 1 with $t \in (1/2, 1)$), we can obtain

$$\begin{aligned}&\inf_{1\le v\le \bar{v}} \left[Q_v\left(\vartheta_v^*\right) - \sup_{\vartheta_v \in \Upsilon_{\boldsymbol{c}} \times \Upsilon(\bar{v});\ \|\vartheta_v - \vartheta_v^*\| \ge \epsilon_1} Q_v\left(\vartheta_v\right) \right] \\ &\ge \frac{\phi\left(1\right)}{2} \left(\inf_{v\ge 1} \lambda_v^{\min} \right) \epsilon_1^2 \inf_{|\gamma_v^*| \le \sqrt{4(\sigma_e^*)^2 \log[\bar{v}(\log \bar{v})^t]}} \phi\left(C_W ||\boldsymbol{c}^*|| + \frac{\epsilon_1}{2}\left(C_W + 1\right) + |\gamma_v^*| \right),\end{aligned}$$

where the inequality holds almost surely for (sufficiently) large $\bar{v}$. We can also derive lower bound of the last component:

$$\begin{aligned}&\inf_{|\gamma_v^*| \le \sqrt{4(\sigma_e^*)^2 \log[\bar{v}(\log \bar{v})^t]}} \phi\left(C_W ||\boldsymbol{c}^*|| + \frac{\epsilon_1}{2}\left(C_W + 1\right) + |\gamma_v^*| \right) \\ &= \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{1}{2} \left| C_W ||\boldsymbol{c}^*|| + \frac{\epsilon_1}{2}\left(C_W + 1\right) + \sqrt{4(\sigma_e^*)^2 \log[\bar{v}\left(\log \bar{v}\right)^t]} \right|^2 \right\} \\ &\ge \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{1}{2} \left| \sqrt{4(\sigma_e^*)^2 \log[\bar{v}\left(\log \bar{v}\right)]} \right|^2 \right\} = \frac{1}{\sqrt{2\pi}} [\bar{v}(\log \bar{v})]^{-2(\sigma_e^*)^2},\end{aligned}$$

where the last equality holds for sufficiently large $\bar{v}$ (since $\log[\bar{v}(\log \bar{v})]$ is much larger than $\log[\bar{v}(\log \bar{v})^t]$). From these, we can obtain the desired result:

$$\text{the LHS of (70)} \ge C_Q [\bar{v}(\log \bar{v})]^{-2(\sigma_e^*)^2}$$

with $C_Q = C_Q(\epsilon_1) = \frac{\phi(1)}{2\sqrt{2\pi}} \left(\inf_{v\ge 1} \lambda_v^{\min} \right) \epsilon_1^2$.

It remains to verify two inequalities (73) and (74).

Proof of inequality (73): Using a bound of the standard normal CDF (p. 112 of Karatzas and Shreve, 1991): $\int_x^\infty \phi\left(u\right) du \le \frac{1}{x}\phi\left(x\right)$ for any $x > 0$, we can derive

$$1 - \Phi\left(x\right) \le \frac{1}{x}\phi\left(x\right) \text{ for } x > 0; \quad \Phi\left(x\right) \le \frac{1}{|x|}\phi\left(|x|\right) \text{ for } x < 0$$

through the symmetry of normal distributions. To avoid explosive behavior of these upper bounds when x is close to 0, we use slightly modified bounds as follows:

$$1 - \Phi\left(x\right) \le 1\left\{0 \le x \le 1\right\} + \frac{1}{x}\phi\left(x\right) 1\left\{1 < x\right\} \quad \text{for } x > 0; \tag{76}$$

$$\Phi\left(x\right) \le 1\left\{-1 \le x \le 0\right\} + \frac{1}{|x|}\phi\left(x\right) 1\left\{x < -1\right\} \quad \text{for } x < 0, \tag{77}$$

which also covers the case with $x = 0$. Since $\Phi\left(x\right) \in (0, 1)$, we have

$$\frac{\phi\left(x\right)}{\Phi\left(x\right)\left[1 - \Phi\left(x\right)\right]} \ge \begin{cases} \phi\left(x\right) & \text{if } |x| \le 1, \\ |x|^2 & \text{if } x > 1, \end{cases}$$

which, together with (76) and (77), leads to

$$\begin{aligned}\frac{\phi\left(x\right)}{\Phi\left(x\right)\left[1 - \Phi\left(x\right)\right]} &\ge 1\left\{x \le 1\right\}\phi\left(1\right) + 1\left\{|x| > 1\right\} x^2 \\ &\ge \phi\left(1\right) \quad \text{for any } x \in \mathbb{R},\end{aligned}$$

where the first inequality holds since $\phi(x) \geq \phi(1) = \frac{1}{\sqrt{2\pi}}\exp\{-\frac{1}{2}\}$ for any $|x| \leq 1$ and the second holds since $1 > \phi(1)$.

Proof of inequality (74): By the Taylor expansion, we have

$$|\Phi(y) - \Phi(x)| = |y - x| \int_0^1 \phi(x + \lambda(y - x))\, d\lambda. \tag{78}$$

Consider an interval defined as

$$\{x + \lambda(y - x) \mid \lambda \in [0, 1]\}.$$

This interval is a connected subset on $\mathbb{R}$; thus it may contain zero, be in the negative region ($x, y < 0$), or be in the positive region ($x, y > 0$). In each of these cases, by the shape of $\phi(z)$ (having a unique peak at $z = 0$, being symmetric, decreasing in the region $z > 0$, and increasing $z < 0$), we have for any $\lambda \in [0, 1]$,

$$\phi(x + \lambda(y - x)) \geq \phi(\max\{|y|, |x|\}) \quad \text{for any } \lambda \in [0, 1],$$

which, together with (78), implies the desired result (74). The proof of Lemma 2 is now complete. ■

Proof of Lemma 3. Let $\bar{\Upsilon}(\bar{v}) := \Upsilon_1 \times \Upsilon(\bar{v})$. This set $\bar{\Upsilon}(\bar{v})$ is compact for each $\bar{v}$ and thus it can be divided into $M(\epsilon_2)$ subsets, $\bar{\Upsilon}^1(\bar{v}), \bar{\Upsilon}^2(\bar{v}), \ldots, \Upsilon^{M(\epsilon_2)}(\bar{v})$, such that $\|(\boldsymbol{c}, \gamma_v) - (\tilde{\boldsymbol{c}}, \tilde{\gamma}_v)\| < \epsilon_2$ whenever $(\boldsymbol{c}, \gamma_v)$ and $(\tilde{\boldsymbol{c}}, \tilde{\gamma}_v)$ are in the same subset. Let $(\boldsymbol{c}^{(j)}, \gamma_v^{(j)})$ be some point in $\bar{\Upsilon}^j(\bar{v})$ for each $j \in \{1, 2, \ldots, M(\epsilon_2)\}$. Since Υ_1 is a compact subset of $\mathbb{R}^{d_W}$ and $\Upsilon(\bar{v})$ is a compact interval on $\mathbb{R}$ that may grow with the rate of $\sqrt{\log[\bar{v}(\log \bar{v})^t]}$, the number of subsets that cover $\bar{\Upsilon}(\bar{v})$ can be bounded as follows:

$$M(\epsilon_2) \leq O(\epsilon_2^{-d_W}) \times O(\epsilon_2^{-1}\sqrt{\log[\bar{v}(\log \bar{v})^t]}) = O(\epsilon_2^{-(d_W+1)}\sqrt{\log[\bar{v}(\log \bar{v})^t]}).$$

Then,

$$\begin{aligned}
&\max_{1 \leq v \leq \bar{v}} \sup_{(\boldsymbol{c}, \gamma_v) \in \bar{\Upsilon}(\bar{v})} \left|\hat{Q}_v(\boldsymbol{c}, \gamma_v) - Q_v(\boldsymbol{c}, \gamma_v)\right| \\
&\leq \max_{1 \leq v \leq \bar{v}} \max_{j \in \{1,2,\ldots,M(\epsilon_2)\}} \sup_{(\boldsymbol{c}, \gamma_v) \in \bar{\Upsilon}^j(\bar{v})} \left|\hat{Q}_v(\boldsymbol{c}, \gamma_v) - Q_v(\boldsymbol{c}, \gamma_v)\right| \\
&\leq \max_{1 \leq v \leq \bar{v}} \max_{j \in \{1,2,\ldots,M(\epsilon_2)\}} \sup_{(\boldsymbol{c}, \gamma_v) \in \bar{\Upsilon}^j(\bar{v})} \left|\frac{1}{N_v}\sum_{v=1}^{N_v} \left[\mathcal{L}_{vh}(\boldsymbol{c}, \gamma_v) - \mathcal{L}_{vh}(\boldsymbol{c}^{(j)}, \gamma_v^{(j)})\right]\right| \\
&+ \max_{1 \leq v \leq \bar{v}} \max_{j \in \{1,2,\ldots,M(\epsilon_2)\}} \sup_{(\boldsymbol{c}, \gamma_v) \in \bar{\Upsilon}^j(\bar{v})} \left|\frac{1}{N_v}\sum_{v=1}^{N_v} E_{\boldsymbol{\omega}_v}\left[\mathcal{L}_{vh}(\boldsymbol{c}, \gamma_v) - \mathcal{L}_{vh}(\boldsymbol{c}^{(j)}, \gamma_v^{(j)})\right]\right| \\
&+ \max_{1 \leq v \leq \bar{v}} \max_{j \in \{1,2,\ldots,M(\epsilon_2)\}} \left|\hat{Q}_v(\boldsymbol{c}^{(j)}, \gamma_v^{(j)}) - Q_v(\boldsymbol{c}^{(j)}, \gamma_v^{(j)})\right|.
\end{aligned} \tag{79}$$

Each of the first two terms on the majorant side is bounded by

$$\|(\boldsymbol{c}, \gamma_v) - (\tilde{\boldsymbol{c}}, \tilde{\gamma}_v)\| \times C_2\sqrt{\log[\bar{v}(\log \bar{v})^t]} \leq \epsilon_2 \times C_2\sqrt{\log[\bar{v}(\log \bar{v})^t]}, \tag{80}$$

which follows from

$$|\mathcal{L}_{vh}(\boldsymbol{c},\gamma_v) - \mathcal{L}_{vh}(\tilde{\boldsymbol{c}},\tilde{\gamma}_v)| \leq \|(\boldsymbol{c},\gamma_v) - (\tilde{\boldsymbol{c}},\tilde{\gamma}_v)\| \times C_2\sqrt{\log[\bar{v}(\log\bar{v})^t]}, \tag{81}$$

where $C_2 > 0$ is some constant that is independent of (v,h), $(\boldsymbol{c},\gamma_v)$, and $(\tilde{\boldsymbol{c}},\tilde{\gamma}_v)$; the proof of (81) is provided below. Here, by setting

$$\epsilon_2 = \sqrt{\log[\bar{v}(\log\bar{v})^t](\log N_0)/N_0}, \tag{82}$$

we have the first two terms on the RHS of (79) be $O(\sqrt{(\log[\bar{v}(\log\bar{v})^t])^2(\log N_0)/N_0})$.

We next derive the probability bound of the third term on the RHS of (79). To this end, we use the following bounds:

$$|\mathcal{L}_{vh}(\boldsymbol{c},\gamma_v)| \leq C_{\mathcal{L}}\log[\bar{v}(\log\bar{v})^t], \tag{83}$$

uniformly over (v,h) and $(\boldsymbol{c},\gamma_v)$, whose proof is provided below, as well as

$$\text{Var}[\textstyle\sum_{v=1}^{N_v}\mathcal{L}_{vh}(\boldsymbol{c},\gamma_v)] \leq N_v\left|C_{\mathcal{L}}\log[\bar{v}(\log\bar{v})^t]\right|^2,$$

which follows from $\text{Var}[\mathcal{L}_{vh}(\boldsymbol{c},\gamma_v)] \leq E[|\mathcal{L}_{vh}(\boldsymbol{c},\gamma_v)|^2] \leq \left|C_{\mathcal{L}}\log[\bar{v}(\log\bar{v})^t]\right|^2$ almost surely as $\bar{v}\to\infty$. Thus, by Bernstein's inequality for independent variables (p. 102, van der Vaart and Wellner, 1996) with these two bounds for $|\mathcal{L}_{vh}(\boldsymbol{c},\gamma_v)|$ and $\text{Var}[\sum_{v=1}^{N_v}\mathcal{L}_{vh}(\boldsymbol{c},\gamma_v)]$, we have for $s>0$,

$$\begin{aligned}
&P_{\boldsymbol{\omega}_v}\left[\max_{1\leq v\leq\bar{v}}\max_{j\in\{1,2,\ldots,M(\epsilon_2)\}}\left|\hat{Q}_v(\boldsymbol{c}^{(j)},\gamma_v^{(j)}) - Q_v(\boldsymbol{c}^{(j)},\gamma_v^{(j)})\right| \geq s\right]\\
&\leq \bar{v}\max_{1\leq v\leq\bar{v}}M(\epsilon_2)\,P_{\boldsymbol{\omega}_v}\left[\sum_{v=1}^{N_v}\left\{\mathcal{L}_{vh}(\boldsymbol{c}^{(j)},\gamma_v^{(j)}) - E[\mathcal{L}_{vh}(\boldsymbol{c}^{(j)},\gamma_v^{(j)})]\right\} \geq N_v s\right]\\
&\leq \bar{v}\max_{1\leq v\leq\bar{v}}M(\epsilon_2)\,2\exp\left\{-\frac{1}{2}\frac{(N_v s)^2}{N_v\left|C_{\mathcal{L}}\log[\bar{v}(\log\bar{v})^t]\right|^2 + \frac{1}{3}C_{\mathcal{L}}\log[\bar{v}(\log\bar{v})^t](N_v s)}\right\}\\
&= \bar{v}\max_{1\leq v\leq\bar{v}}M(\epsilon_2)\,2\exp\left\{-\frac{1}{2}\frac{(N_v/N_0)\,s_0^2(\log N_0)}{|C_{\mathcal{L}}|^2 + \frac{C_{\mathcal{L}}}{3}s_0\sqrt{(\log N_0)/N_0}}\right\}\\
&\leq \bar{v}\times O\left(\epsilon_2^{-(d_W+1)}\sqrt{\log[\bar{v}(\log\bar{v})^t]}\right)\times N_0^{-\underline{r}s_0^2/2(|C_{\mathcal{L}}|^2+s_0)},
\end{aligned} \tag{84}$$

where the equality holds with

$$s = s_0\sqrt{\left|\log[\bar{v}(\log\bar{v})^t]\right|^2(\log N_0)/N_0}\quad (s_0>0 \text{ is a constant}), \tag{85}$$

and the last inequality holds since $(N_v/N_0)\geq\underline{r}$ (for any v) and $\frac{C_U}{3}\sqrt{(\log N_0)/N_0}\leq 1$ for sufficiently large N_0. Since we have defined $\epsilon_2 = \sqrt{\log[\bar{v}(\log\bar{v})^t](\log N_0)/N_0}$ in (82) and $\log\bar{v}$ is at most of polynomial order of N_0, for some sufficiently large $s_0>0$, the majorant side of (84) tends to zero as $N_0\to\infty$. That is, given (85) and $N_v = r_v N_0$ with $\underline{r}\leq r_v\leq\bar{r}$, we have

$$\max_{1\leq v\leq\bar{v}}\max_{j\in\{1,2,\ldots,M(\epsilon_2)\}}\left|\hat{Q}_v(\boldsymbol{c}^{(j)},\gamma_v^{(j)}) - Q_v(\boldsymbol{c}^{(j)},\gamma_v^{(j)})\right| = O_p(\sqrt{(\log[\bar{v}(\log\bar{v})^t])^2(\log N_0)/N_0}).$$

Noting that $\left(\log[\bar{v}\,(\log\bar{v})^t]\right)^2 < (\log\bar{v})^3$, this bound, together with (79) and (80), implies the conclusion of this Lemma 3.

To complete the proof of the lemma, it remains to show inequalities (81) and (83).

Proof of (81): We look at

$$\begin{aligned}
&|\mathcal{L}_{vh}(\boldsymbol{c},\gamma_v) - \mathcal{L}_{vh}(\tilde{\boldsymbol{c}},\tilde{\gamma}_v)| \\
&= A_{vh}\left[\log F_\varepsilon\left(W_{vh}\boldsymbol{c}' + \gamma_v\right) - \log F_\varepsilon\left(W_{vh}\tilde{\boldsymbol{c}}' + \tilde{\gamma}_v\right)\right] \\
&+ (1-A_{vh})\left[\log\left(1 - F_\varepsilon\left(W_{vh}\boldsymbol{c}' + \gamma_v\right)\right) - \log\left(1 - F_\varepsilon\left(W_{vh}\tilde{\boldsymbol{c}}' + \tilde{\gamma}_v\right)\right)\right] \\
&\leq \|(\boldsymbol{c},\gamma_v) - (\tilde{\boldsymbol{c}},\tilde{\gamma}_v)\|\left\{\sup_{(\boldsymbol{c},\gamma_v)\in\bar{\Upsilon}(\bar{v})}\frac{f_\varepsilon\left(W_{vh}\boldsymbol{c}' + \gamma_v\right)}{F_\varepsilon\left(W_{vh}\boldsymbol{c}' + \gamma_v\right)} + \sup_{(\boldsymbol{c},\gamma_v)\in\bar{\Upsilon}(\bar{v})}\frac{f_\varepsilon\left(W_{vh}\boldsymbol{c}' + \gamma_v\right)}{1 - F_\varepsilon\left(W_{vh}\boldsymbol{c}' + \gamma_v\right)}\right\}.
\end{aligned}$$

Since the parameter space Υ_1 in which $\boldsymbol{c}$ lies is compact, the support of W_{vh} is bounded, and $|\gamma_v| \leq \sqrt{4(\sigma_e^*)^2\log[\bar{v}\,(\log\bar{v})^t]}$, we can bound possible minimum and maximum values of $W_{vh}\boldsymbol{c}' + \gamma_v$ uniformly over (v,h). That is, by letting

$$I_{\bar{v}} := C^0 + C^1\sqrt{\log[\bar{v}\,(\log\bar{v})^t]} \tag{86}$$

with sufficiently large constants, $C^0, C^1 > 0$, we may suppose that $-I_{\bar{v}} \leq W_{vh}\boldsymbol{c}' + \gamma_v \leq I_{\bar{v}}$. By the inequalities in (87), we can find

$$\begin{aligned}
\sup_{(\boldsymbol{c},\gamma_v)\in\bar{\Upsilon}(\bar{v})}\frac{f_\varepsilon\left(W_{vh}\boldsymbol{c}' + \gamma_v\right)}{F_\varepsilon\left(W_{vh}\boldsymbol{c}' + \gamma_v\right)} &\leq \sup_{-I_{\bar{v}}\leq x\leq I_{\bar{v}}}\frac{f_\varepsilon(x)}{F_\varepsilon(x)} \\
&\leq \sup_{-I_{\bar{v}}\leq x<-1}\frac{\phi(x)}{\frac{|x|}{1+x^2}\phi(x)} + \sup_{x\geq -1}\frac{f_\varepsilon(x)}{F_\varepsilon(x)} \leq \sup_{-I_{\bar{v}}\leq x<-1}(1+|x|) + \frac{\phi(0)}{\Phi(-1)}
\end{aligned}$$

and

$$\begin{aligned}
\sup_{(\boldsymbol{c},\gamma_v)\in\Upsilon_{\boldsymbol{c}}\times\Upsilon_\gamma^{\bar{v}}}\frac{f_\varepsilon\left(W_{vh}\boldsymbol{c}' + \gamma_v\right)}{1 - F_\varepsilon\left(W_{vh}\boldsymbol{c}' + \gamma_v\right)} &\leq \sup_{-I_{\bar{v}}\leq x\leq I_{\bar{v}}}\frac{f_\varepsilon(x)}{1 - F_\varepsilon(x)} \\
&\leq \sup_{x\leq 1}\frac{f_\varepsilon(x)}{1 - F_\varepsilon(x)} + \sup_{1<x\leq I_{\bar{v}}}\frac{\phi(x)}{\frac{x}{1+x^2}\phi(x)} \\
&\leq \frac{\phi(0)}{1 - \Phi(1)} + \sup_{1<x\leq I_{\bar{v}}}(1+x).
\end{aligned}$$

Therefore, given these bounds and the definition of $I_{\bar{v}}$, we can write

$$|\mathcal{L}_{vh}(\boldsymbol{c},\gamma_v) - \mathcal{L}_{vh}(\tilde{\boldsymbol{c}},\tilde{\gamma}_v)| \leq \|(\boldsymbol{c},\gamma_v) - (\tilde{\boldsymbol{c}},\tilde{\gamma}_v)\| \times C_2\sqrt{\log[\bar{v}\,(\log\bar{v})^t]},$$

for any large $\bar{v}$, where $C_2 > 0$ is some constant that is independent of (v,h), $\bar{v}$, $(\boldsymbol{c},\gamma_v)$, and $(\tilde{\boldsymbol{c}},\tilde{\gamma}_v)$.

Proof of (83). Note that $F_\varepsilon\left(W_{vh}\boldsymbol{c}' + \gamma_v\right) \in (0,1)$ and thus $\mathcal{L}_{vh}(\boldsymbol{c},\gamma_v) \leq 0$. To find a lower bound of $\mathcal{L}_{vh}(\boldsymbol{c},\gamma_v)$, we use the following results:

$$\frac{x}{1+x^2}\phi(x) \leq 1 - F_\varepsilon(x) \;\text{ for } x \geq 0; \text{ and } \frac{|x|}{1+x^2}\phi(x) \leq F_\varepsilon(x) \quad \text{for } x < 0, \tag{87}$$

which follows p. 112 of Karatzas and Shreve (1991) and the symmetry of the standard normal distribution (noting that $1 - F_\varepsilon(x) = \int_x^\infty \phi(u)\,du$), and separately look at $\log(1 - F_\varepsilon(W_{vh}\boldsymbol{c}' + \gamma_v))$ and $\log F_\varepsilon(W_{vh}\boldsymbol{c}' + \gamma_v)$. This inequality (87) and the bound of $W_{vh}\boldsymbol{c}' + \gamma_v$, $I_{\bar{v}}$ (defined in(86)), we obtain

$$\begin{aligned}\log(1 - F_\varepsilon\left(W_{vh}\boldsymbol{c}' + \gamma_v\right)) &\geq \log(1 - F_\varepsilon(I_{\bar{v}})) \geq \log\left(\frac{I_{\bar{v}}}{1 + (I_{\bar{v}}^*)^2}\phi(I_{\bar{v}})\right) \\ &= \log I^* - \log(1 + |I_{\bar{v}}|^2) + \log\phi(I_{\bar{v}})\end{aligned}$$

and analogously,

$$\begin{aligned}\log F_\varepsilon\left(W_{vh}\boldsymbol{c}' + \gamma_v\right) &\geq \log F_\varepsilon(-I_{\bar{v}}) \\ &\geq \log I_{\bar{v}} - \log(1 + |I_{\bar{v}}|^2) + \log\phi(-I_{\bar{v}}).\end{aligned}$$

Therefore, since $\phi(I_{\bar{v}}) = \phi(-I_{\bar{v}})$ (by the symmetry of ϕ), we have

$$0 \geq \mathcal{L}_{vh}(\boldsymbol{c}, \gamma_v) \geq \log I_{\bar{v}} - \log(1 + |I_{\bar{v}}|^2) + \log\phi(I_{\bar{v}}).$$

By definition of $I_{\bar{v}}$, $\log\phi(I_{\bar{v}}) = -\log\sqrt{2\pi} - \frac{\left|C^0 + C^1\sqrt{\log[\bar{v}(\log\bar{v})^t]}\right|^2}{2}$ and thus we can find some constant $C_{\mathcal{L}} \in (0, \infty)$ that is independent of (v, h) and $\bar{v}$ satisfying

$$|\mathcal{L}_{vh}(\boldsymbol{c}, \gamma_v)| \leq C_{\mathcal{L}} \log[\bar{v}(\log\bar{v})^t],$$

which is the desired result. The proof of Lemma 3 is complete. ■

A.5 Simulation Exercise

To see how our two-step probit performs in finite samples, we set up the following simulation exercise. We generate data according to the model specified in (37), (38) (39) and (40) above. We vary $\bar{v}$ while holding $N_v = 250$ in each case, to resemble our application. We choose $\dim(W_{vh}) = 2$, and pick for $v = 1, ..., \bar{v}$,

$$d_v = [v \times U_1 + 0.1 \times U_2,\ \ v \times 0.02 \times N(0,1)],$$

where U_1 and U_2 are independent uniform [0,1], I.I.D. across villages. Then generate $P_{vh} \simeq 2 \times U[0,1] + v \times 0.1 \times U[0,1]$, $e_v \simeq N\left(0, 0.2^2\right)$, $\delta = (1,1)$ and $\tau_{vh}, \varepsilon_{vh} \simeq N(0,1)$, I.I.D. across v and h. The multiplication by v in the generation of d_v, P_{vh} lead to variation in the distribution of observables across villages, which produces variation in π_v necessary to point-identify α. Finally, as in (37), (38) above, we generate

$$W_{vh} = d_v + \tau_{vh} \ \text{ and } \ \xi_v = d_v\boldsymbol{\delta}' + e_v = \bar{W}_v\boldsymbol{\delta}' + e_v - \bar{\tau}_v\boldsymbol{\delta}',$$

and the π_v's by solving the fixed point problem

$$\arg\min_{\pi_v}\left\{\pi_v-\frac{1}{N_v}\sum_{h=}^{N_v}1\left(W'_{vh}\begin{pmatrix}1\\2\end{pmatrix}-P_{vh}+2\pi_v+\xi_v+\varepsilon_{vh}>0\right)\right\}^2$$

Using the π_vs obtained from the previous step. we generate the outcome as

$$A_{vh}=1\left\{1.75+0.5\times W^1_{vh}+W^2_{vh}-2\times P_{vh}+2\times\pi_v+\xi_v+\varepsilon_{vh}>0\right\},$$

and compute σ_e, α and the price coefficient by the double probit exercise described in equations (36) and (41)/(42) above. For each choice of $\bar{v}$, we repeat the process over 100 replications. The results of the simulation exercise are reported in Table 6, where the true values of the parameters used to generate the data are displayed at the top.

Table 6: Simulation Results

True values	**sigma**	**alfa**	**price_coeff**			
	0.200	2.000	-2.000			
	Mean	S.D.	Median	10th %ile	90th %ile	RMSE
		vbar=5				
sigma	0.201	0.207	0.168	0.070	0.335	0.21
alfa	2.180	2.780	2.430	-2.470	5.610	2.79
price_coeff	-2.060	0.140	-2.040	-2.240	-1.870	0.15
		vbar=10				
sigma	0.285	0.050	0.273	0.140	0.780	0.10
alfa	2.220	0.740	2.171	0.841	3.630	0.77
price_coeff	-2.030	0.120	-2.040	-2.200	-1.880	0.12
		vbar=25				
sigma	0.41	0.54	0.36	0.113	0.317	0.58
alfa	2.02	3.21	2.38	1.233	3.25	3.21
price_coeff	-2.05	0.1	-2.07	-2.21	-1.95	0.11

Based on the root mean square error and quantile values of the estimates, it is clear that for $\bar{v}=10$, the estimates are more precise than when we have $\bar{v}=5,25$. As $\bar{v}$ rises from 5 to 10, the standard deviation of the estimated α and the root-mean square error decrease due to larger variation in π_v which helps pin down α. On the other hand, the deterioration from $\bar{v}=10$ to $\bar{v}=25$ results from the need to estimate many more village fixed effects γ_v in (36). This represents the fundamental trade-off discussed in the asymptotic results in the paper; a larger $\bar{v}$ helps average out the e_vs, but also increases the number of nuisance parameters γ_v to be estimated in the first stage probit. Given that $\bar{v}=11$ in our application, the good performance under $\bar{v}=10$ in our simulation exercise is reassuring.