Intra-Household Effects on Demand for Telephone Service: Empirical Evidence

Huang, Ching-I
National Taiwan University

January 2008

Intra-Household Effects on Demand for Telephone Service: Empirical Evidence

Ching-I Huang*

January 21, 2008

Abstract

I present a game-theoretical model to estimate consumption demand, accounting for intra-household interaction among household members. Although multiple Nash equilibria of consumption decisions may exist in a household, model parameters are pointwise identified from household-level data for households with only two members. I propose a semiparametric maximum likelihood estimator and apply it to empirically analyze the subscription decision for cellular phone service in Taiwan. On average, a consumer's probability of subscribing to cellular service rises 35 percentage points when the other household member chooses to subscribe. This result suggests the existence of intra-household network effects on cellular phone consumption. The intra-household effect increases in household income, but decreases in the number of kids and the age difference in a household.

*Department of Economics, National Taiwan University, Taipei, Taiwan. This paper was previously circulated under the title "Intra-Household Consumption Externalities of Cellular Phone Service". I received helpful comments from Aviv Nevo, Robert Porter, William Rogerson, James Heckman, Chung-Ming Kuan, and Yi-Chi Chen. I have also benefited from comments received from 2007 North American summer meeting of the Econometric Society and 2007 Far Eastern meeting of the Econometric Society. I appreciate the assistance from the Center for Survey Research of Academia Sinica in providing data. Financial supports from the Center for the Study of Industrial Organization at Northwestern University and the National Science Council in Taiwan are gratefully acknowledged. All remaining errors are mine.

1 Introduction

Standard microeconomic theory analyzes consumer behavior based on individual preferences. When more than one person lives in a household, we need to take into account the intra-household allocation of resources and consumption externalities among household members. Consequently, a consumer's decision depends on other household member's choices. In this paper, I use the term *intra-household effect* to refer to the effect on a consumer's willingness to pay caused by the decision of other members in the same household. I propose a game-theoretical framework to estimate the intra-household effect on the consumption of telephone service.

There is a rich literature on estimating household demand for telecommunication service. Generally, these studies use household-level survey data. Nonetheless, each household is treated as a single decision-maker in the estimation. Only household heads' individual characteristics are included in the demand estimation. This approach implicitly assumes the demand to be solely determined by household heads. Other members can influence the decision only indirectly through household-level variables.[1] This assumption is unlikely to be true in reality. Besides, most of these empirical studies on telephone demand focus on landline phone service. There are relatively few works on the demand for cellular phone service. Iyengar (2004) and Grzybowski and Pereira (2007) estimate the cellular phone service demand by using data from billing records. No demographic characteristic is observed in Iyengar (2004)'s data while only very few demographic characteristics are available in Grzybowski and Pereira (2007)'s work. In contrast, I use a household survey which provides information on many household demographic variables.[2]

[1]Train, McFadden, and Ben-Akiva (1987) and Train, Ben-Akiva, and Atherton (1989) only consider aggregate household income. In the estimation of demand for local telephone service under optional rate plans, Miravete (2002) includes several household-level characteristics. His empirical analysis only accounts for household head's individual characteristics, but not other members' characteristics. Many previous researches (Rappoport and Taylor, 1997; Solvason, 1997; Madden and Simpson, 1997; Duffy-Deno, 2001; Rodini, Ward, and Woroch, 2003; Economides, Seim, and Viard, 2006) use similar approach in estimating telephone demand.

[2]A different approach, proposed by Bajari, Fox, and Ryan (2007), estimates demand for cellular phone service by using market share ranks. Their focus is the value of national coverage. Individual demographic variables are not included in their estimation as well.

As Browning, Bourguignon, Chaippori, and Lechene (1994) point out, household behavior depends on intra-household interactions unless we impose some restrictive hypotheses such as transferable utilities. They propose a collective household model: Household members bargain with each other to allocate their overall resources. Individual consumption depends on the allocation. The bargaining power depends on individual characteristics. The resource allocation must achieve Pareto efficient in the bargaining process. Using data on couples with no kids, they find that the allocation of expenditure depends on the relative incomes and relative ages of the couples, rejecting the hypotheses of a single decision-maker in a household. See Vermeulen (2002) for further discussions on the collective household model.

I consider a model of binary subscription choices. When there is only a single person in a household, this model reduces to a standard discrete choice model. When more than one person lives in a household, consumption externalities among the members may affect their demand. For example, if the husband has a cellular phone, the wife may have a stronger desire to own a cellular phone as well. There are several possible reasons for *positive* intra-household effect. The first one is the *direct network effect.* Because the husband can be contacted by phone more frequently, the wife's demand for cellular phone service increases. The second reason is the *indirect network effect.* For instance, since the husband's knowledge of cellular phone service from his own consumption reduces the wife's search cost on her subscription decision, she is likely to have higher demand. A third explanation is the *price effect.* When the price of a cellular-to-cellular phone call is lower than that of a landline-to-cellular phone call[3], the wife can pays less for a call from a cellular phone than one from a landline phone. Consequently, she may have higher demand for cellular phone service. Similarly, if carriers offer family plans which lowers the subscription fee for a second cellular phone, there exist positive intra-household effect on the consumption. On the other hand, intra-household effect may be *negative* if a cellular phone is a *public good* in a household. Then, each household member wants to be a free-rider and shares the usage of the other person's cellular phone. Because intra-household effect can be either positive or negative in theory, finding out its

[3]This is the case when cellular carriers offers in-network discounts.

sign for a particular good is an empirical issue. The objective of the paper is to estimate the sign and the magnitude of intra-household effects for cellular phone service.

In the current paper, I restrict my attention to households with only two members. Each of the two members makes a binary choice on the subscription of telephone service. My model is similar to entry models in the industrial organization literature, such as Bresnahan and Reiss (1990). However, the spillover effects between two firms in an entry model are always negative. The entrance of one firm reduces the profit of the other firm. Their entry decisions must be strategic substitutes. For household consumption behavior, the effects may be either positive or negative. When they are positive, the decisions are strategic complements. I do not restrict the sign of intra-household effects in the estimation. In addition, the sign can vary across households. I will investigate how intra-household effects vary across households. Different from Browning et al. (1994)'s collective household model, the equilibrium allocation is not necessarily Pareto optimal in my game-theoretical model.

The primary difficulty in the estimation is to deal with multiple Nash equilibria. When intra-household effect is negative, we can estimate the model by using the equilibrium number of subscribers in a household, which always has a unique equilibrium, as in Bresnahan and Reiss (1990)'s entry model. On the contrary, when intra-household effect is positive, this approach does not work because the equilibrium number of subscribers may have multiple Nash equilibria as well. Because there is no one-to-one mapping between model primitives and outcomes in the presence of multiple equilibria, model parameters can only be partially identified in general. To explicitly deal with multiple Nash equilibria, one approach is to consider the selection rule among these equilibria. For example, Jia (2007) imposes an ad hoc selection rule to choose among multiple market equilibria. Bajari, Hong, and Ryan (2007) propose a simulation-based method to estimate the selection rule. Another approach uses bounds estimation based on inequality constraints derived from necessary conditions for pure strategy Nash equilibria (Chernozhukov, Hong, and Tamer, 2007; Ciliberto and Tamer, 2007; Pakes, Porter, Ho, and Ishii, 2006). Nonetheless, for a two-by-two game, Tamer (2003) shows that point identification can be achieved under a suitable exclusion condition. He proposes a two-step estimation procedure. In the first step, use nonparametric estimation to determine the selection among multiple equilibria. Model parameters are obtained by a maximum likelihood estimation in the second step. In Tamer's paper, however, the interaction effect between the two agents is assumed to be constant, independent of any observed characteristic. In this paper, I generalize Tamer's model and allow the interaction effect to be heterogeneous, depending on observed household characteristics. I show that the parameters in this generalized demand model can still be pointwise identified.

I apply the econometric approach to study the demand for cellular phone service in Taiwan. The estimated marginal intra-household effect of cellular phone service increases a consumer's probability of subscription by 35.24 percentage points on average. The finding suggests the existence of network effect within a household. Heterogeneity of the effects across households can be explained by the observed characteristics. Intra-household effects increase in household income, but decrease in the number of kids and the age difference in a household.

Another important contribution of this paper is to estimate the direct effect of both household-level and individual-level characteristics on telephone demand after controlling for intra-household effects. Previous researches only include household head's individual characteristics, but not other member's individual characteristics, in the estimation. Therefore, it is difficult to identify the effect of some individual demographic variables such as gender since most household heads are male. In this paper, I find that males have higher demand for cellular phone service than females. Moreover, household income by itself has a *negative* effect on telephone consumption, but individual income has a *positive* effect. Because the latter effect is substantially larger than the former one, the total effect for an increase in a consumer's income is positive.

In the next section, I introduce the econometric model and propose my estimation approach. Section 3 describes the data to be used in the estimation. I then present my empirical results on the consumption of cellular phone service in Taiwan and demonstrate the estimated intra-household effects in Section 4. Concludes are given in the final section.

2 Econometric Model

The presence of intra-household effect means that consumption depends on the decision of other household members. In this section, I present a static discrete choice model which is an extension of the probit model. I restrict my attention to households with two members and show that the model parameters are fully identified despite the existence of multiple Nash equilibria in a noncooperative game between the two household members. I then present a semiparametric maximum likelihood estimator and conduct a Monte Carlo experiment to demonstrate the performance of the estimator.

2.1 Discrete Choice Model

For household i, there are two members $j \in \{1, 2\}$. All characteristics of each member are observed by both members. *Household-level characteristics*, such as household income, residence location, ..., etc., are common to both member, while *individual-level characteristics*, such as gender, age, education ...etc., are not. Furthermore, some characteristics, such as taste on new technology, are observed only by the two household members, but not by the econometrician.

A consumer's subscription decision depends on the *direct effect of consumption* and the *intra-household effect of consumption.* The former effect is determined by the consumer's own individual-level characteristics as well as the household-level characteristics in his household. The latter effect depends on the choice of the other household member. Its magnitude is normalized to zero when the other member does not subscribe. I assume the intra-household effect is reciprocal between the two members and its magnitude is determined by household-level characteristics.

Let the binary variable $y_{ij} \in \{0, 1\}$ denote *the subscription decision* of individual j in household i. Let $y_{ij} = 1$ if and only if the individual subscribes to the telephone service. The demand is characterized by

$$y_{ij} = 1 \quad \Leftrightarrow \quad [\mathbf{x}'_{ij}\boldsymbol{\beta} + \varepsilon_{ij}] + y_{i(3-j)}[\mathbf{z}'_i\boldsymbol{\gamma}] > 0, \tag{1}$$

where $(3 - j)$ is the index for the other member in the household. The terms in the first bracket of equation (1) represents the direct effect of consumption. The term in the second bracket, $\mathbf{z}'_i\boldsymbol{\gamma}$, captures the magnitude of the intra-household effect. The vector $\mathbf{x}_{ij}$ is member j's *observed characteristics* (including both household-level and individual-level characteristics) and the scalar ε_{ij} represents his *unobserved characteristics*. The vector $\mathbf{z}_i$ includes all *household-level characteristics* which affect the intra-household effect. To identify the model parameters, at least one of the elements in the vector $\mathbf{x}_{ij}$ (such as member j's age) is not a household-level characteristic. Furthermore, both $\mathbf{x}_{ij}$ and $\mathbf{z}_i$ contain a constant term. My model reduces to the standard probit model if the intra-household effect vanishes ($\boldsymbol{\gamma} = \mathbf{0}$). If the intra-household effect is restricted to be constant across households, as in Tamer (2003)'s model, the vector $\mathbf{z}_i$ only contains the constant term ($\mathbf{z}'_i\boldsymbol{\gamma} = \gamma_0$).

The unobserved characteristics $(\varepsilon_{i1}, \varepsilon_{i2})$ are assumed to be jointly normally distributed, independently across households.

$$\begin{bmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right). \tag{2}$$

The variance of ε_{ij} is normalized to one. The correlation coefficient ρ in (2) is to be estimated.

Finally, let $Y_i = y_{i1} + y_{i2}$ denote the total number of subscribers in the household.

2.2 Nash Equilibria

Consider a simultaneous-move non-cooperative game.[4] This is similar to the incomplete model discussed in Tamer (2003). Figure 1 shows the set of equilibria for positive intra-

[4]The set of Nash equilibria under a cooperative game is a subset of Nash equilibria under a non-cooperative game. Consequently, the results under a cooperative game can be viewed as imposing an equilibrium selection rule on the results under a non-cooperative game. I will discuss more on this issue in Subsection 2.4.

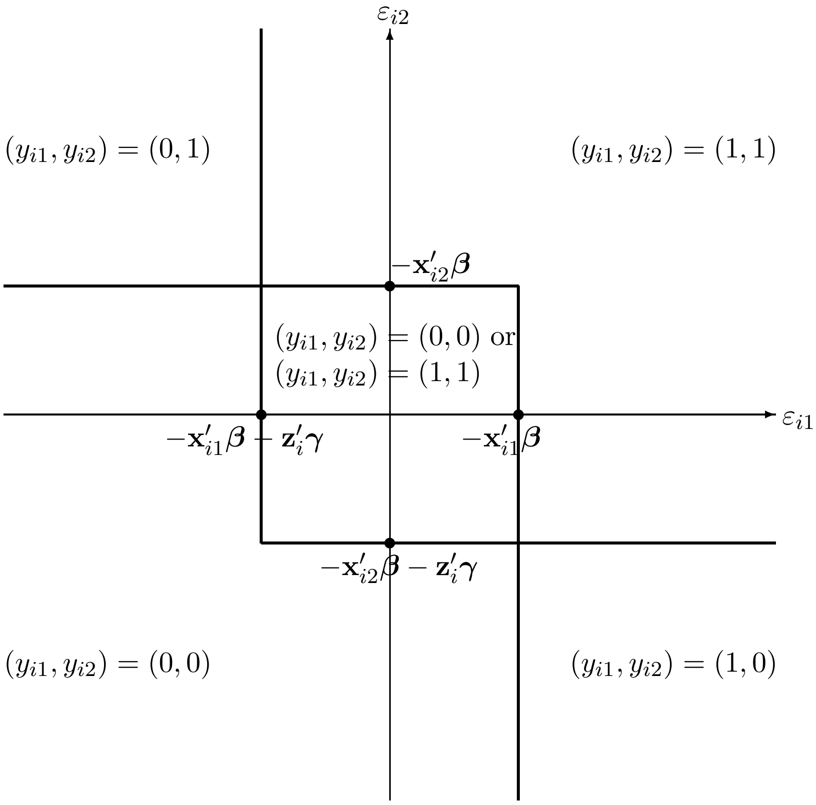

Figure 1: Nash equilibria for positive intra-household effects

household effect ($\mathbf{z}_i'\boldsymbol{\gamma} > 0$) conditional on observed characteristics ($\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i$) and unobserved characteristics ($\varepsilon_{i1}, \varepsilon_{i2}$). There are multiple Nash equilibria when $(\varepsilon_{i1}, \varepsilon_{i2}) \in (-\mathbf{x}_{i1}'\boldsymbol{\beta} - \mathbf{z}_i'\boldsymbol{\gamma}, -\mathbf{x}_{i1}'\boldsymbol{\beta}) \times (-\mathbf{x}_{i2}'\boldsymbol{\beta} - \mathbf{z}_i'\boldsymbol{\gamma}, -\mathbf{x}_{i2}'\boldsymbol{\beta})$. Both $(y_{i1}, y_{i2}) = (0,0)$ and $(y_{i1}, y_{i2}) = (1,1)$ are equilibria in this region. Nonetheless, the model predicts the exact probability for $(y_{i1}, y_{i2}) = (0,1)$ and $(y_{i1}, y_{i2}) = (1,0)$. Given the observed characteristics $(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i)$, the probability of the event $(y_{i1}, y_{i2}) = (0,0)$ is bounded by

$$\Pr\left(\{\varepsilon_{i1} < -\mathbf{x}_{i1}'\boldsymbol{\beta} - \mathbf{z}_i'\boldsymbol{\gamma},\ \varepsilon_{i2} < -\mathbf{x}_{i2}'\boldsymbol{\beta}\} \cup \{\varepsilon_{i1} < -\mathbf{x}_{i1}'\boldsymbol{\beta},\ \varepsilon_{i2} < -\mathbf{x}_{i2}'\boldsymbol{\beta} - \mathbf{z}_i'\boldsymbol{\gamma}\} | \mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i\right)$$

and

$$\Pr(\varepsilon_{i1} < -\mathbf{x}_{i1}'\boldsymbol{\beta},\ \varepsilon_{i2} < -\mathbf{x}_{i2}'\boldsymbol{\beta} | \mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i).$$

On the other hand, when the effect is negative ($\mathbf{z}_i'\boldsymbol{\gamma} < 0$), there are multiple equilibria of $(0,1)$ and $(1,0)$ if $(\varepsilon_{i1}, \varepsilon_{i2}) \in (-\mathbf{x}_{i1}'\boldsymbol{\beta}, -\mathbf{x}_{i1}'\boldsymbol{\beta} - \mathbf{z}_i'\gamma) \times (-\mathbf{x}_{i2}'\boldsymbol{\beta}, -\mathbf{x}_{i2}'\boldsymbol{\beta} - \mathbf{z}_i'\gamma)$. (See Figure 2.) The model gives the exact probabilities of $(y_{i1}, y_{i2}) = (0,0)$ and $(y_{i1}, y_{i2}) = (1,1)$, but not $(y_{i1}, y_{i2}) = (0,1)$ and $(y_{i1}, y_{i2}) = (1,0)$.

Regardless the sign of intra-household effect, the exact probability of observing one subscriber in a household ($Y_i = y_{i1} + y_{i2} = 1$) for given observed characteristics $(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i)$ can summarized as

$$\begin{aligned}
& P_1\left(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i; \boldsymbol{\beta}, \boldsymbol{\gamma}, \rho\right) \\
\equiv\, & \Pr\left(Y_i = 1 | \mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i; \boldsymbol{\beta}, \boldsymbol{\gamma}\right) \\
=\, & \Pr\left(\varepsilon_{i1} < -\mathbf{x}_{i1}'\boldsymbol{\beta} - \mathbf{z}_i'\boldsymbol{\gamma}, \varepsilon_{i2} > -\mathbf{x}_{i2}'\boldsymbol{\beta} | \mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i\right) + \Pr\left(\varepsilon_{i1} > -\mathbf{x}_{i1}'\boldsymbol{\beta}, \varepsilon_{i2} < -\mathbf{x}_{i2}'\boldsymbol{\beta} - \mathbf{z}_i'\boldsymbol{\gamma} | \mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i\right) \\
& - \mathbf{1}\{\mathbf{z}_i'\boldsymbol{\gamma} < 0\} \Pr\left(-\mathbf{x}_{i1}'\boldsymbol{\beta} < \varepsilon_{i1} < -\mathbf{x}_{i1}'\boldsymbol{\beta} - \mathbf{z}_i'\boldsymbol{\gamma}, -\mathbf{x}_{i2}'\boldsymbol{\beta} < \varepsilon_{i2} < -\mathbf{x}_{i2}'\boldsymbol{\beta} - \mathbf{z}_i'\boldsymbol{\gamma} | \mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i\right), \qquad (3)
\end{aligned}$$

where $\mathbf{1}\{\cdot\}$ denotes the indicator function. However, the exact probabilities of no subscriber ($Y_i = 0$) and two subscribers ($Y_i = 2$) in a household are unknown when intra-household effect is positive because we do not know how individuals choose among multiple Nash equilibria.[5]

[5] Contrary to my model, in Bresnahan and Reiss (1990)'s entry model, the effect must be negative. As a

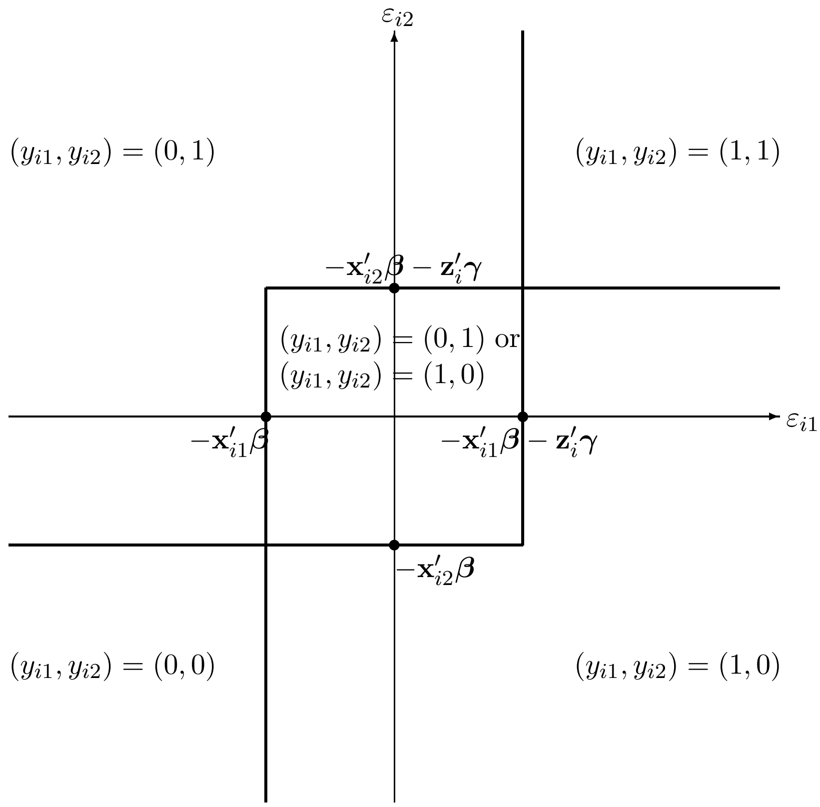

Figure 2: Nash equilibria for negative intra-household effects

Without loss of generality, we only need to focus on the probability $\Pr(Y_i = 0|\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i)$ because $\Pr(Y_i = 2|\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i)$ can be obtained from $1 - \Pr(Y_i = 0|\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i) - \Pr(Y_i = 1|\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i)$. Without imposing any equilibrium selection rule, the probability of no subscriber in a household is bounded in an interval. The upper bound occurs when individuals always fail to coordinate their decisions in the event of multiple Nash equilibria.

$$P_0^U(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i; \boldsymbol{\beta}, \boldsymbol{\gamma}, \rho) \equiv \Pr(\varepsilon_{i1} < -\mathbf{x}_{i1}'\boldsymbol{\beta},\ \varepsilon_{i2} < -\mathbf{x}_{i2}'\boldsymbol{\beta}|\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i). \tag{4}$$

The lower bound is achieved when individuals can perfectly coordinate.

$$\begin{aligned} P_0^L(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i; \boldsymbol{\beta}, \boldsymbol{\gamma}, \rho) &\equiv \Pr(\varepsilon_{i1} < -\mathbf{x}_{i1}'\boldsymbol{\beta},\ \varepsilon_{i2} < -\mathbf{x}_{i2}'\boldsymbol{\beta}|\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i) \\ &- \mathbf{1}\{\mathbf{z}_i'\boldsymbol{\gamma} > 0\} \Pr(-\mathbf{x}_{i1}'\boldsymbol{\beta} - \mathbf{z}_i'\boldsymbol{\gamma} < \varepsilon_{i1} < -\mathbf{x}_{i1}'\boldsymbol{\beta}, -\mathbf{x}_{i2}'\boldsymbol{\beta} - \mathbf{z}_i'\boldsymbol{\gamma} < \varepsilon_{i2} < -\mathbf{x}_{i2}'\boldsymbol{\beta}|\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i). \end{aligned} \tag{5}$$

2.3 Identification

Although multiple Nash equilibria are possible, the parameters in the econometric model are pointwise identified. My model is similar to but more complicated than Tamer (2003). Tamer's model is identified when we have data on the individual decisions (y_{i1}, y_{i2}). However, the data set that I use only reports the aggregate decision in a household ($Y_i = y_{i1} + y_{i2}$), not individual choices. Nonetheless, the following theorem shows the parameters are still identified.

Theorem 1. *Suppose that there exists a regressor of individual characteristics* (x_{i1k}, x_{i2k}) *with* $x_{i1k}, x_{i2k} \notin \mathbf{z}_i$ *and* $\beta_k \neq 0$ *and such that the conditional distribution of* $x_{i1k}|x_{-i1k}$ *has an everywhere positive Lebesgue density where* $\mathbf{x}_{-i1k} = (x_{i11}, \ldots, x_{i1,k-1}, x_{i1,k+1}, \ldots, x_{i1K})'$. *Then the parameters,* $(\boldsymbol{\beta}, \boldsymbol{\gamma}, \rho)$, *are identified if the matrices* $X_1 \equiv [\mathbf{x}_{11}\ \mathbf{x}_{21} \cdots \mathbf{x}_{N1}]$, $X_2 \equiv [\mathbf{x}_{12}\ \mathbf{x}_{22} \cdots \mathbf{x}_{N2}]$, *and* $Z \equiv [\mathbf{z}_1\ \mathbf{z}_2 \cdots \mathbf{z}_N]$ *have full rank.*

Proof. In equation (3), I have shown that the exact probabilities of $Y_i = 1$, which is denoted by $P_1(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i; \boldsymbol{\beta}, \boldsymbol{\gamma}, \rho)$, can be obtained for any given observed characteristics $(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i)$.

result, the value of Y_i is unique in equilibrium.

Without loss of generality, assume $\beta_k > 0$. Let $(\widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}}, \widetilde{\rho})$ be different from $(\boldsymbol{\beta}, \boldsymbol{\gamma}, \rho)$. I will consider four possible cases.

Case 1: $\widetilde{\boldsymbol{\beta}} \neq \boldsymbol{\beta}$ and $\widetilde{\beta}_k > 0$: As x_{i1k} goes to minus infinity for given x_{-i1k}, both $x_{i1k}\beta_k$ and $x_{i1k}\widetilde{\beta}_k$ go to minus infinity. Because X_2 has full rank, there exists $\mathbf{x}_{i2}^*$ such that $\mathbf{x}_{i2}^{*\prime}\boldsymbol{\beta} \neq \mathbf{x}_{i2}^{*\prime}\widetilde{\boldsymbol{\beta}}$. Consequently, as $x_{i1k} \to -\infty$,

$$\begin{aligned} P_1\left(\mathbf{x}_{i1}, \mathbf{x}_{i2}^*, \mathbf{z}_i; \boldsymbol{\beta}, \boldsymbol{\gamma}, \rho\right) &\simeq \Pr\left(\varepsilon_{i2} > -\mathbf{x}_{i2}^{*\prime}\boldsymbol{\beta}\right) \\ &\neq \Pr\left(\varepsilon_{i2} > -\mathbf{x}_{i2}^{*\prime}\widetilde{\boldsymbol{\beta}}\right) \simeq P_1\left(\mathbf{x}_{i1}, \mathbf{x}_{i2}^*, \mathbf{z}_i; \widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}}, \widetilde{\rho}\right). \end{aligned}$$

This implies the parameters $(\boldsymbol{\beta}, \boldsymbol{\gamma}, \rho)$ are identified.

Case 2: $\widetilde{\boldsymbol{\beta}} \neq \boldsymbol{\beta}$ and $\widetilde{\beta}_k < 0$: Since Z has full rank, there exists $\mathbf{z}_i^*$ such that $\mathbf{z}_i^{*\prime}\boldsymbol{\gamma} \neq \mathbf{z}_i^{*\prime}\widetilde{\boldsymbol{\gamma}}$ when $\widetilde{\boldsymbol{\gamma}} \neq \boldsymbol{\gamma}$. If the parameters are not identified, then

$$\begin{aligned} P_1\left(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i^*; \boldsymbol{\beta}, \boldsymbol{\gamma}, \rho\right) &\simeq \Pr(\varepsilon_{i2} > -\mathbf{x}_{i2}^{\prime}\boldsymbol{\beta}) \\ &= \Pr\left(\varepsilon_{i2} < -\mathbf{x}_{i2}^{\prime}\widetilde{\boldsymbol{\beta}} - \mathbf{z}_i^{*\prime}\widetilde{\boldsymbol{\gamma}}\right) \simeq P_1\left(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i^*; \widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}}, \widetilde{\rho}\right). \end{aligned} \tag{6}$$

for any $\mathbf{x}_{i2}$ as $x_{i1k} \to -\infty$ for given x_{-i1k}, and

$$\begin{aligned} P_1\left(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i^*; \boldsymbol{\beta}, \boldsymbol{\gamma}, \rho\right) &\simeq \Pr\left(\varepsilon_{i2} < -\mathbf{x}_{i2}^{\prime}\boldsymbol{\beta} - \mathbf{z}_i^{*\prime}\boldsymbol{\gamma}\right) \\ &= \Pr\left(\varepsilon_{i2} > -\mathbf{x}_{i2}^{\prime}\widetilde{\boldsymbol{\beta}}\right) \simeq P_1\left(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i^*; \widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}}, \widetilde{\rho}\right). \end{aligned} \tag{7}$$

for any $\mathbf{x}_{i2}$ as $x_{i1k} \to +\infty$ for given x_{-i1k}. Since ε_{i2} is a symmetric distribution with zero mean, Equations (6) and (7) together imply

$$\mathbf{x}_{i2}^{\prime}\boldsymbol{\beta} = -\mathbf{x}_{i2}^{\prime}\widetilde{\boldsymbol{\beta}} - \mathbf{z}_i^{*\prime}\widetilde{\boldsymbol{\gamma}} = \mathbf{x}_{i2}^{\prime}\boldsymbol{\beta} + \mathbf{z}_i^{*\prime}\boldsymbol{\gamma} - \mathbf{z}_i^{*\prime}\widetilde{\boldsymbol{\gamma}} \neq \mathbf{x}_{i2}^{\prime}\boldsymbol{\beta}.$$

This is a contradiction. Therefore, equations (6) and (7) cannot hold together, implying the parameters $(\boldsymbol{\beta}, \boldsymbol{\gamma}, \rho)$ are identified when $\widetilde{\boldsymbol{\gamma}} \neq \boldsymbol{\gamma}$.

If $\widetilde{\boldsymbol{\gamma}} = \boldsymbol{\gamma}$, either equation (6) or (7) implies that $\mathbf{x}_{i2}^{\prime}(\boldsymbol{\beta} + \widetilde{\boldsymbol{\beta}}) + \mathbf{z}_i^{*\prime}\boldsymbol{\gamma} = 0$ holds for any

$(\mathbf{x}_{i2}, \mathbf{z}_i)$. This contradicts with the fact that X_2 and Z both have full rank.

Case 3: $\widetilde{\boldsymbol{\beta}} = \boldsymbol{\beta}$ but $\widetilde{\boldsymbol{\gamma}} \neq \boldsymbol{\gamma}$: Because $\beta_k > 0$, I know $\widetilde{\beta}_k > 0$. Let x_{i1k} go to positive infinity. Both $x_{i1k}\beta_k$ and $x_{i1k}\widetilde{\beta}_k$ go to positive infinity. Because Z has full rank, there exists $\mathbf{z}_i^{**}$ such that $\mathbf{z}_i^{**\prime}\boldsymbol{\gamma} \neq \mathbf{z}_i^{**\prime}\widetilde{\boldsymbol{\gamma}}$. As $x_{-i1k} \to +\infty$, for any $\mathbf{x}_{i2}$, I have

$$\begin{aligned} P_1\left(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i^{**}; \boldsymbol{\beta}, \boldsymbol{\gamma}, \rho\right) &\simeq \Pr\left(\varepsilon_{i2} < -\mathbf{x}_{i2}'\boldsymbol{\beta} - \mathbf{z}_i^{**\prime}\boldsymbol{\gamma}\right) \\ &\neq \Pr\left(\varepsilon_{i2} < -\mathbf{x}_{i2}'\widetilde{\boldsymbol{\beta}} - \mathbf{z}_i^{**\prime}\widetilde{\boldsymbol{\gamma}}\right) \simeq P_1\left(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i^{**}; \widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}}, \rho\right). \end{aligned}$$

Therefore, I can identify the parameters.

Case 4: $(\widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}}) = (\boldsymbol{\beta}, \boldsymbol{\gamma})$ but $\widetilde{\rho} \neq \rho$: For $\mathbf{z}_i'\boldsymbol{\gamma} > 0$, I have

$$\frac{\partial P_1(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i; \boldsymbol{\beta}, \boldsymbol{\gamma}, \rho)}{\partial \rho} = -\frac{e^{-\frac{(\mathbf{x}_{i1}'\boldsymbol{\beta}+\mathbf{z}_i'\boldsymbol{\gamma})^2 + (\mathbf{x}_{i2}'\boldsymbol{\beta})^2 - 2\rho(\mathbf{x}_{i1}'\boldsymbol{\beta}+\mathbf{z}_i'\boldsymbol{\gamma})(\mathbf{x}_{i2}'\boldsymbol{\beta})}{2(1-\rho^2)}}}{\pi\sqrt{1-\rho^2}} < 0.$$

Similarly, for $\mathbf{z}_i'\boldsymbol{\gamma} < 0$, I can obtain

$$\begin{aligned} &\frac{\partial P_1(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i; \boldsymbol{\beta}, \boldsymbol{\gamma}, \rho)}{\partial \rho} = \\ &\quad -\frac{e^{-\frac{(\mathbf{x}_{i1}'\boldsymbol{\beta})^2 + (\mathbf{x}_{i2}'\boldsymbol{\beta})^2 - 2\rho(\mathbf{x}_{i1}'\boldsymbol{\beta})(\mathbf{x}_{i2}'\boldsymbol{\beta})}{2(1-\rho^2)}}}{2\pi\sqrt{1-\rho^2}} - \frac{e^{-\frac{(\mathbf{x}_{i1}'\boldsymbol{\beta}+\mathbf{z}_i'\boldsymbol{\gamma})^2 + (\mathbf{x}_{i2}'\boldsymbol{\beta}+\mathbf{z}_i'\boldsymbol{\gamma})^2 - 2\rho(\mathbf{x}_{i1}'\boldsymbol{\beta}+\mathbf{z}_i'\boldsymbol{\gamma})(\mathbf{x}_{i2}'\boldsymbol{\beta}+\mathbf{z}_i'\boldsymbol{\gamma})}{2(1-\rho^2)}}}{2\pi\sqrt{1-\rho^2}} < 0. \end{aligned}$$

Therefore, ρ can be identified from the data. □

Note that identification of the coefficients $(\boldsymbol{\beta}, \boldsymbol{\gamma})$ only depends on the marginal distribution of the idiosyncratic preferences $(\varepsilon_{i1}, \varepsilon_{i2})$, not on their joint distribution.

2.4 Semiparametric Maximum Likelihood Estimator

If intra-household effect is negative, I know the exact probability of the events $\{Y_i = 0\}$, $\{Y_i = 1\}$, and $\{Y_i = 2\}$ conditional on the observed characteristics $(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i)$. Consequently, the usual likelihood can be computed. On the contrary, the exact probabilities of $\{Y_i = 0\}$ and $\{Y_i = 2\}$ are unknown when the effect is positive. I use a semiparametric maximum likelihood estimator, extended from Tamer (2003)'s approach, to obtain the parameters in the demand model. Specifically, the estimation consists of two steps. In the first step, I use a kernel regression to obtain the empirical probability of $\{Y_i = 0\}$. In the second step, I replace the unknown probabilities in the likelihood function by the empirical probabilities obtained in the first step and maximize the likelihood to obtain the parameter estimates.

Define the conditional probability of the event $\{Y_i = 0\}$ for observed characteristics $(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i)$ as

$$H(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i) = \Pr\left(Y_i = 0 | \mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i\right).$$

When the function H is known, I can write down the likelihood, and the parameters $(\boldsymbol{\beta}, \boldsymbol{\gamma}, \rho)$ are estimated by maximizing the logarithm of the likelihood function. For a random sample with size N,[6] the logarithm of the likelihood function is

$$\begin{aligned} L(\boldsymbol{\beta}, \boldsymbol{\gamma}, \rho; H) = \frac{1}{N} \sum_i \Big\{ & \mathbf{1}[Y_i = 0] \log(H(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i)) \\ & + \mathbf{1}[Y_i = 1] \log\left(P_1\left(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i; \boldsymbol{\beta}, \boldsymbol{\gamma}, \rho\right)\right) \\ & + \mathbf{1}[Y_i = 2] \log\Big(1 - H(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i) - P_1\left(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i; \boldsymbol{\beta}, \boldsymbol{\gamma}, \rho\right)\Big)\Big\} \end{aligned} \quad (8)$$

The unknown function $H(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i)$ represents the probability of observing no subscriber in a household. From equations (4) and (5), we know that $H(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i)$ is bounded by the closed interval $[P_0^L(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i; \boldsymbol{\beta}, \boldsymbol{\gamma}, \rho), P_0^U(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i; \boldsymbol{\beta}, \boldsymbol{\gamma}, \rho)]$ when multiple Nash equilibria exist, but the model cannot predict the exact probability. I follow Tamer (2003)'s suggestion to approximate the unknown function by a kernel regression of the event $\{Y_i = 0\}$ on the observed characteristics $(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i)$.[7] Since the function $H(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i)$ is bounded by $[P_0^L(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i; \boldsymbol{\beta}, \boldsymbol{\gamma}, \rho), P_0^U(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i; \boldsymbol{\beta}, \boldsymbol{\gamma}, \rho)]$, I truncate the result of the kernel regression

[6] The survey data I use to perform estimation is not a random sample. Therefore, I need to adjust for the sampling weights in my calculation. To ease the exposition, however, I present the estimator without writing down the sampling weights.

[7] Any function which locally approximates the true probability of $\{Y_i = 0\}$ can be used in the estimation. In the current paper, I use Gaussian kernel regression to estimate $H(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i)$.

$$\hat{H}(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i) = \frac{\frac{1}{N} \sum_{i'} \mathbf{1}[Y_{i'} = 0] \phi\left(\frac{1}{B} d[(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i), (x_{i'1}, x_{i'2}, \mathbf{z}_i)]\right)}{\frac{1}{N} \sum_{i'} \phi\left(\frac{1}{B} d[(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i), (x_{i'1}, x_{i'2}, \mathbf{z}_i)]\right)},$$

by the upper and lower bounds and denote the value as $\hat{H}(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i; \boldsymbol{\beta}, \gamma, \rho)$. Replace H in the likelihood (8) by $\hat{H}$. I can obtain a consistent estimate of $(\boldsymbol{\beta}, \gamma, \rho)$. To obtain the variance of the estimator, I need to account for the standard errors resulting from the kernel regression in the first step. Instead of computing the analytic variance, I compute the variance by bootstrapping.

Estimating the model under the assumption of a simultaneous-move non-cooperative game seems restrictive, but it is actually not. When the interaction within a household is not a simultaneous-move non-cooperative game, the proposed estimator is still consistent, though it is less efficient. If all households can perfectly coordinate their consumption decisions as in a simultaneous-move cooperative game, the kernel estimator of $H(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i)$ in the first step will converge to $P_0^L(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i)$ in probability. My proposed estimator remains valid for this situation. Similarly, if individuals make decisions sequentially rather than simultaneously, then the unique subgame-prefect equilibrium is also a subset of the Nash equilibria under a simultaneous-move non-cooperative game. The kernel estimator of $H(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i)$ will also converge to the exact probability of no subscriber in the subgame-perfect equilibrium. Consequently, my estimation approach can also apply to this case.

2.5 Monte Carlo

To demonstrate the performance of my estimator, I conduct a Monte Carlo experiment using the discrete choice model described in (1). For each individual, there are three observed characteristics (x_{ij1}, x_{ij2}, z_i). The first two variables have direct effect on the subscription decision, while the final one affects the intra-household effect. In addition, assume $x_{i11} = x_{i21}$, that is, this variable is common to both household members. The choice model of (1) can be

where ϕ is the density function of a standard normal distribution, and the metric d is defined as

$$d[(\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{z}_i), (x_{i'1}, x_{i'2}, \mathbf{z}_{i'})] \equiv \sqrt{\frac{1}{2K+L}\left[\sum_{j=1}^{2}\sum_{k=1}^{K}\frac{(x_{ijk}-x_{i'jk})^2}{Var(x_{\cdot jk})} + \sum_{l=1}^{L}\frac{(z_{il}-z_{i'l})^2}{Var(z_{\cdot l})}\right]}.$$

A bandwidth $B = 0.3$ is used for the following results. The parameter estimates are robust to changes in the bandwidth B.

Table 1: Monte Carlo Results

Parameter	True Value	Mean	Median	Standard Deviation	MSE
β_0	-1.0	-1.006	-0.999	0.073	0.005
β_1	2.5	2.605	2.580	0.207	0.054
β_2	1.5	1.551	1.535	0.116	0.016
γ_0	1.0	0.955	0.962	0.151	0.024
γ_1	1.5	1.500	1.510	0.120	0.014
ρ	-0.5	-0.535	-0.543	0.177	0.074

Notes: The Monte Carlo simulation is conducted for 100 samples. Each sample consists of 1000 households.

expresses as

$$y_{ij} = 1 \quad \Leftrightarrow \quad [\beta_0 + \beta_1 x_{ij1} + \beta_2 x_{ij2} + \varepsilon_{ij}] + y_{i(3-j)}[\gamma_0 + \gamma_1 z_i] > 0. \tag{9}$$

All the observed characteristics are generated from the standard normal distribution $N(0, 1)$ independently. The unobserved characteristics $(\varepsilon_{i1}, \varepsilon_{i2})$ are drawn from a joint normal distribution with unit variance and correlation coefficient ρ.

I generate 100 samples of size 1000 to assess the property of my proposed estimator. True parameters are $\beta_0 = -1.0$, $\beta_1 = 2.5$, $\beta_2 = 1.5$, $\gamma_0 = 1.0$, $\gamma_1 = 1.5$, and $\rho = -0.5$. When multiple Nash equilibria (i.e. both $Y_i = 0$ and $Y_i = 2$ are equilibria.) exist, I assume that the event $Y_i = 2$ occurs randomly with probability 0.7. The results are shown in Table 1. All the parameters can be estimated reasonably well. The estimator for the correlation of unobserved characteristics (ρ) has the largest mean squared error. In addition, this value is identified only through the functional form assumption, not directly from the data. Therefore, in my empirical study, I will have more confidence on the estimator of the coefficients $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, and not make any inference about the estimated correlation ρ.

3 Data

The data come form the 2003 Survey of Family Income and Expenditure in Taiwan. This survey was conducted by the Directorate-General of Budget, Accounting and Statistics in early 2004. It adopts a stratified two-stage sampling method with counties and cities as subpopulations. The universal sampling rate is 0.20%, which is 13,681 households. Because young kids are unlikely to make their own decisions and they are unlikely to use telephones, young kids are not counted as household members in my empirical work. I define young kids as people who are less than 6 years old. The estimation results do not change much for different definition of young kids. Based on this age criterion, there are 3,489 households with two members.

Descriptive statistics are presented in Table 2. The first two columns are means and standard deviations for the subsample with two household members. The final two columns are for the entire sample in the survey. The upper panel shows the household-level variables, while the lower panel presents individual-level variables. I construct three household-level variables (*age difference*, *education difference*, and *income difference*) from the original data by computing the difference of individual-level variables in each two-member household. Incomes are measured in Taiwan dollars (TWD). The average exchange rate between US dollars and Taiwan dollars in 2003 is 1 USD = 34.42 TWD. Note that household income is more than twice of individual income in the subsample because part of the household income cannot be attributed to either member. The mean of age in the subsample is considerably older than that in the entire population. This is reasonable because families with one or more teenagers living together with their parents are excluded in the subsample. Households in the subsample also tends to have lower total income because their sizes are smaller on average. Besides, households in the subsample are modestly more likely to live in the South region.[8]

I only observe the total numbers of cellular phones in a household. Table 3 summarizes

[8]As defined by the Directorate General of Telecommunications, the counties and cities included in each of the three regions are the following. (1) North: Keelung, Taipei, Taoyuan, Hsinchu, Yilan, Hualien, and Lienchiang; (2) Central: Miaoli, Taichung, Changhua, Nantou, and Yunlin; (3) South: Chiayi , Tainan, Kaohsiung, Pingtung, Taitung, Penghu, and Kinmen.

Table 2: Descriptive Statistics

Variable	Subsample Mean	Subsample Std. Dev.	Entire Sample Mean	Entire Sample Std. Dev.	Description
Household-level Variables					
Cellular Phone	1.071	0.876	1.859	1.324	Number of cell phones
Household Income	0.789	0.603	1.065	0.740	Annual income (10^6 TWD)
City	0.793	0.405	0.807	0.395	Living in a city
Town	0.168	0.374	0.163	0.369	Living in a town
Rural	0.039	0.193	0.030	0.171	Living in rural area
North	0.444	0.497	0.472	0.499	Living in North region
Central	0.223	0.416	0.228	0.419	Living in Cental region
South	0.333	0.471	0.300	0.458	Living in South region
Number of Kids	0.270	0.597	0.217	0.530	Number of young kids
Household Size	2.000	0.000	3.310	1.490	Number of HH members
Age Difference	0.107	0.122			Age difference
Education Differ.	0.305	0.330			Education difference (10 yr)
Income Difference	0.502	0.483			Individual income difference
Individual-level Variables					
Gender	0.511	0.500	0.501	0.500	Female = 1
Age	0.523	0.185	0.387	0.202	Age (100 years)
Education	0.884	0.478	0.963	0.424	Years of Education (10 years)
Employment	0.478	0.500	0.466	0.499	Employed = 1
Individual Income	0.368	0.463	0.301	0.449	Annual Income (10^6 TWD)
sample size	3489		13681		

Notes: The sampling weights are used to compute means and standard deviations.

Table 3: Distribution of the Number of Telephones in a Household

Number of Cellular Phone	Percentage
0	30.57
1	35.32
2	31.32
3	2.02
4	0.73
5	0.05

Notes: Percentages are computed according to the sampling weights.

the distributions of the number of telephones among households with two members. When the total is zero, obviously neither member subscribes to the phone service. When it is one, only one member in the household choose to subscribe, and the other member does not. When there is more than one phone, I assume that both individuals choose to have one. In my data, 3% of two-member households own more than two cellular phones.

4 Empirical Results

In this section, I apply the estimation method introduced in Section 2 to analyze the demand for cellular phone service in Taiwan. The parameter estimates for the coefficients in (1) are presented in Table 4.

First, in Column (A), intra-household effects are restricted to be zero. The model is identical to the standard probit model except that unobserved characteristics ε_{ij} are allowed to be correlated within a household. The estimated coefficients for household income and individual income are both significantly positive, suggesting cellular phone service is a normal good. Moreover, age has a negative effect on demand while education and employment have positive effects.

In Column (B), I include intra-household effect in the choice model but assume the effect to be a constant γ_0 across households. The marginal intra-household effect is 28.4 percentage points, which is significantly positive. Therefore, I can reject the hypothesis $\gamma_0 = 0$.[9] This result suggests the subscription decisions are strategically complements within a household. There exists within-household networks effect on the consumption of cellular phone service. The estimate for the coefficient of household income remains positive, but its magnitude is much smaller than that in Column (A) once I control for the intra-household effect.

In the final two columns of Table 4, I allow for the heterogeneity of intra-household effect. The effect is captured by $\mathbf{z}_i'\boldsymbol{\gamma}$. In Column (C), I include all observed household-level characteristics into the vector $\mathbf{z}_i$. In addition to the household-level variables shown in Table

[9] Alternatively, the hypothesis can also be rejected by a likelihood-ratio test. The likelihood in Column (B) is significantly larger than the likelihood in Column (A).

Table 4: Estimation Results

	(A)	(B)		(C)		(D)	
Characteristics	$\boldsymbol{\beta}$	$\boldsymbol{\beta}$	$\boldsymbol{\gamma}$	$\boldsymbol{\beta}$	$\boldsymbol{\gamma}$	$\boldsymbol{\beta}$	$\boldsymbol{\gamma}$
constant	0.240	-0.485***	0.952***	-0.425***	1.140***	-0.465***	1.188***
	(0.160)	(0.112)	(0.056)	(0.151)	(0.200)	(0.124)	(0.179)
			[0.289]				
Household Income	0.304***	0.087*		-0.546***	0.589***	-0.544***	0.627***
	(0.083)	(0.047)		(0.134)	(0.114)	(0.119)	(0.124)
	[0.077]	[0.025]		[-0.156]	[0.163]	[-0.156]	[0.173]
Town	-0.129***	-0.077*		-0.007	-0.135		-0.139*
	(0.047)	(0.046)		(0.047)	(0.089)		(0.076)
	[-0.033]	[-0.022]		[-0.002]	[-0.038]		[-0.039]
Rural	-0.139	0.076		0.069	0.047		0.104
	(0.103)	(0.081)		(0.105)	(0.156)		(0.126)
	[-0.035]	[0.022]		[0.020]	[0.013]		[0.028]
Central	-0.003	0.078		0.099	-0.050	0.058	
	(0.045)	(0.056)		(0.068)	(0.073)	(0.055)	
	[-0.001]	[0.023]		[0.029]	[-0.014]	[0.017]	
South	-0.083*	-0.018		-0.022	0.002	-0.019	
	(0.046)	(0.046)		(0.049)	(0.064)	(0.043)	
	[-0.021]	[-0.005]		[-0.006]	[0.001]	[-0.006]	
Number of Kids	-0.029	-0.021		0.047	-0.173***		-0.123***
	(0.041)	(0.032)		(0.057)	(0.062)		(0.041)
	[-0.007]	[-0.006]		[0.013]	[-0.048]		[-0.034]
Average Age					-0.428		-0.472*
					(0.267)		(0.257)
					[-0.118]		[-0.130]
Average Education					0.044		
					(0.097)		
					[0.012]		
Average Employment					-0.128		-0.188**
					(0.101)		(0.092)
					[-0.035]		[-0.052]
Age Difference					-1.013***		-1.054***
					(0.271)		(0.288)
					[-0.280]		[-0.290]
Education Difference					0.081		0.112
					(0.075)		(0.069)
					[0.022]		[0.031]
Income Difference					0.036		
					(0.150)		
					[0.010]		
Gender	0.007	-0.124*		-0.120*		-0.128*	
	(0.081)	(0.064)		(0.068)		(0.067)	
	[0.002]	[-0.036]		[-0.035]		[-0.037]	
Age	-2.677***	-1.588***		-1.364***		-1.312***	
	(0.156)	(0.148)		(0.153)		(0.126)	
	[-0.677]	[-0.461]		[-0.391]		[-0.378]	
Education	0.742***	0.598***		0.549***		0.601***	
	(0.064)	(0.084)		(0.090)		(0.082)	
	[0.188]	[0.174]		[0.157]		[0.173]	
Employment	0.419***	0.284***		0.246***		0.284***	
	(0.068)	(0.063)		(0.068)		(0.060)	
	[0.114]	[0.086]		[0.073]		[0.085]	
Individual Income	0.617***	0.617***		1.134***		1.080***	
	(0.179)	(0.161)		(0.223)		(0.178)	
	[0.156]	[0.179]		[0.325]		[0.311]	
ρ	0.014	-0.841***		-0.908***		-0.899***	
	(0.039)	(0.030)		(0.019)		(0.021)	
Likelihood	-2548.813	-2520.171		-2483.426		-2488.481	

Notes: Standard errors, computed from 50 bootstrap draws, are in parentheses. Marginal effects, computed as average derivatives of the subscription probability except for for dummy variables whose effects are evaluated for a move from 0 to 1, are in square brackets. Superscripts ***, **, and *represent significance at 1%, 5%, and 10%, respectively. The sample size is 3489 households.

2, I construct three additional household-level characteristics by averaging individual-level variables within a household: *average age*, *average education years*, and *average employment status*. The last specification, Column (D), removes the characteristics which are not significantly different from zero in (C). According to likelihood-ratio tests, there are significant improvements in the likelihood from Column (B) to (C), but no significant difference between (C) and (D) at the 5% significance level. Consequently, I select Column (D) as the preferred specification.

The magnitude of intra-household effect in a household can be expressed as the marginal effect of one member's subscription decision on the other member. Specifically, the *marginal intra-household effect* is the change in the subscription probability for given observed characteristics when the other member change his decision. For member j in household i, it can be expressed as

$$\Pr\left([\mathbf{x}_{ij}'\boldsymbol{\beta} + \varepsilon_{ij}] + [\mathbf{z}_i'\boldsymbol{\gamma}] > 0 | \mathbf{x}_{ij}, \mathbf{z}_i\right) \; - \; \Pr\left([\mathbf{x}_{ij}'\boldsymbol{\beta} + \varepsilon_{ij}] > 0 | \mathbf{x}_{ij}, \mathbf{z}_i\right).$$

Based on the parameters $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\gamma}}$ estimated in Column (D), I compute the marginal intra-household effect for each individual. Figure 3 shows the estimated distribution. All of the estimated effects are positive. On average, the effect increases subscription by 35.24 percentage points, with a standard deviation of 10.50 percentage points. This is a substantial effect when comparing with the average subscription rate 51.77%. When one household member chooses to subscribe, its average effect on the other member is equivalent to the effect caused by increasing the other member's own individual annual income by 2.283 million TWD (equal to 66,326 USD) and holding the first household member's income fixed. The existence of positive intra-household effects suggests the existence of network effects of cellular phone service within a household.

As Figure 3 illustrates, intra-household effects vary a lot across households. Estimate of the vector-valued parameter $\boldsymbol{\gamma}$ differs significantly from zero at 5% level for several variables, providing explanations for the heterogeneity. The impact of household income on

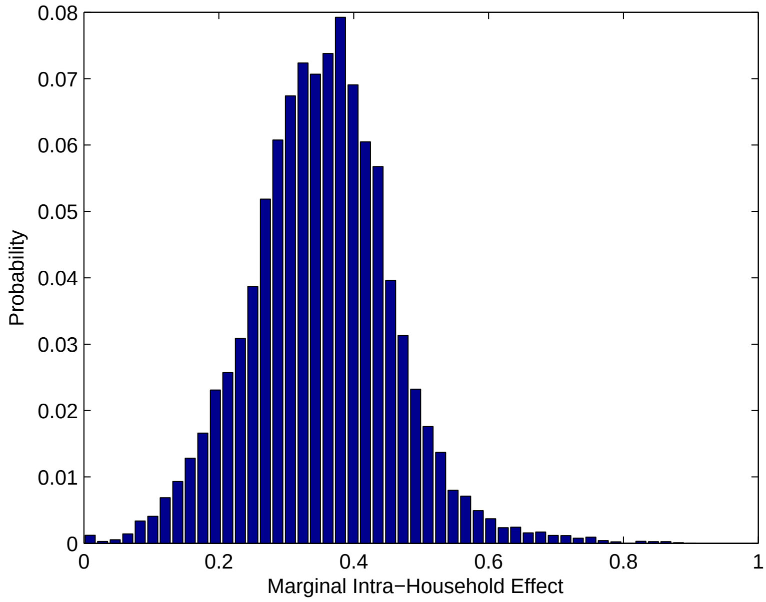

Figure 3: Histogram of the estimated marginal intra-household effects

intra-household effects is both statistically and economically significant. Increasing household income by one standard deviation (i.e. 603 thousand TWD) raises the marginal intra-household effect by 10.4 percentage points. This is probably because households with higher income tend to work longer and spend less time together. Unfortunately, I do not have data on working hours to verify this conjecture. Households in cities have larger marginal intra-household effect than those in towns by 3.9 percentage points. The number of kids has a negative effect, probably because families with more kids tend to spend more time together and hence reduce the network effect of cellular phone service. Each additional kid reduces the marginal intra-household effect by 3.4 percentage points. Interestingly, similar to the findings in Browning et al. (1994), intra-household effects are affected by the difference of individual characteristics within a household. Age difference reduces the intra-household effect. For instance, at the mean age difference (i.e. 10.7 years), marginal intra-household effect is smaller than a household with identical-age members by 3.0 percentage points. Within-household network effect of cellular phone consumption decreases in the relative age. On the contrary,

Table 5: Cellular Phone Ownership by Region

	North	Central	South
Cellular Phone per Household	2.0422	1.7941	1.6212
Cellular Phone per Person	0.5739	0.4892	0.4806

Table 6: Cellular Phone Ownership by Urbanization Level

	City	Town	Rural Area
Cellular Phone per Household	1.9522	1.5314	1.1450
Cellular Phone per Person	0.5557	0.4198	0.3489

there is no significant difference across the three regions: North, Central, and South. Income difference within a household only has a small and insignificant effect.

Contrary to most previous researches on telecommunication demand, I can estimate the direct effect ($\boldsymbol{\beta}$) of both household-level and individual-level characteristics. Different from the findings in Column (A), where intra-household effect is abstracted away, the demand increases in individual income but decreases in household income when I account for the intra-household effects in Column (D). The magnitude of the former effect (31.1%) is almost twice the magnitude of the latter one (15.6%). Consequently, the overall effect of raising a consumer's income is positive. For instance, raising a consumer's income by one standard deviation (i.e. 463 thousand TWD) increases his own demand by 7.16 percentage points. Nonetheless, the finding is in contrast with several previous studies on the demand for *landline* phone service.[10]

As for the geographic variables, there is no significant difference in the direct effect across regions and across urbanization levels. Table 5 and Table 6 show the penetration rate of cellular service across regions and across urbanization levels. Although the penetration rates are higher in the North region and in cities, the demand for cellular phone service does

[10] For example, Miravete (2002) finds household income has negative effects on landline phone service in two cities in Kentucky in 1986. Economides et al. (2006) also find a negative effect of income on the demand in New York State in the period 1999 – 2003.

not have a systematic relationship with the penetration rates. Consequently, there is no evidence showing the existence of network effects resulting from *geographic neighborhoods*. Furthermore, while there are four cellular phone carriers operating in the North region, there are five carriers operating in the Central and the South regions. More carriers in the Central and South regions provide more varieties to consumers. However, varieties of cellular phone service do not have significant effect on the demand.[11]

Lastly, the direct effects on demand resulting from individual characteristics are consistent with intuition. The demand is stronger for young, better-educated, and employed people. Females have weaker demand than males though the difference is significant only at the 10% level. The estimation result is probably caused by the fact that young and better-educated people are more familiar with new technology. Employed people usually spend more time away from home, so they are likely to have higher demand.

5 Conclusion

I empirically analyze intra-household effects on the demand for cellular phone service under a game-theoretical framework. Because of the interaction between household members, it is possible to have multiple Nash equilibria in a non-cooperative simultaneous-move game. Nonetheless, the model parameters are fully identified from the household-level data. I use a semiparametric maximum likelihood estimator to analyze the demand for cellular phone service in Taiwan. The intra-household effect of cellular phone service is positive on average, supporting the existence of network effect on cellular phone consumption within a household. I also analyze how the intra-household effect varies with household characteristics. This effect increases in household income but decreases in the number of kids and the age difference in a household. Furthermore, my estimator provide an estimation of the direct effect on telephone service consumption resulting from observed characteristics, after accounting for the intra-household effects.

[11]It is possible that the effects caused by higher penetration rates and fewer varieties in the North region cancel out. Without more information, I cannot distinguish between these two effects.

In the current paper, I restrict my attention to household with only two members. An important extension is to include households with more than two individuals. Contrary to the two-member case, the exact probability of any observed event is unknown due to multiple equilibria. The parameters are only partially identified by inequality conditions.

References

Bajari, P., J. T. Fox, and S. Ryan (2007). Evaluating wireless carrier consolidation using semi-parametric demand estimation. Mimeo. University of Minnesota, University of Chicago, and Massachusetts Institute of Technology.

Bajari, P., H. Hong, and S. Ryan (2007). Identification and estimation of a discrete game of complete information. Mimeo. University of Minnesota, Stanford University, and Massachusetts Institute of Technology.

Bresnahan, T. F. and P. C. Reiss (1990). Entry in monopoly markets. *Review of Economic Studies 57*, 531–553.

Browning, M., F. Bourguignon, P.-A. Chaippori, and V. Lechene (1994). Income and outcomes: A structural model of intrahousehold allocation. *Journal of Political Economics 102*, 1067–1096.

Chernozhukov, V., H. Hong, and E. Tamer (2007). Estimation and confidence regions for parameter sets in econometric models. *Econometrica 75*, 1243–1284.

Ciliberto, F. and E. T. Tamer (2007). Market structure and multiple equilibria in the airline industry. Mimeo. Available at SSRN: `http://ssrn.com/abstract=605723`.

Duffy-Deno, K. T. (2001). Demand for additional telephone lines: An empirical note. *Information Economics and Policy 13*, 283–299.

Economides, N., K. Seim, and B. V. Viard (2006). Quantifying the benefits of entry into local phone service. Mimeo. New York University and Stanford University.

Grzybowski, L. and P. Pereira (2007). Subscription choices and switching costs in mobile telephony. Mimeo. NET Institute Working Paper #07-12.

Iyengar, R. (2004). A structural demand analysis for wireless services under nonlinear pricing schemes. Mimeo. University of Pennsylvania.

Jia, P. (2007). What happens when Wal-Mart comes to town: An empirical analysis of the discount retailing industry. Mimeo. MIT. Available at `http://econ-www.mit.edu/faculty/pjia/`.

Madden, G. and M. Simpson (1997). Residential broadband subscription demand: an econometric analysis of australian choice experiment. *Applied Economics 29*, 1073–1078.

Miravete, E. J. (2002). Estimating demand for local telephone service with asymmetric information and optional calling plans. *Review of Economic Studies 69*, 943–971.

Pakes, A., J. Porter, K. Ho, and J. Ishii (2006). Wmoment inequalities and their application. Mimeo. Harvard University.

Rappoport, P. N. and L. D. Taylor (1997). Toll price elasticities estimated from a sample of U.S. residential telephone bills. *Information Economics and Policy 9*, 51–70.

Rodini, M., M. R. Ward, and G. A. Woroch (2003). Going mobile: Substitutability between fixed and mobile access. *Bell Journal of Economics and Management Science 27*, 457–476.

Solvason, D. L. (1997). Cross-sectional analysis of residential telephone subscription in canada using 1994 data. *Information Economics and Policy 9*, 241–264.

Tamer, E. (2003). Incomplete simultaneous discrete response model with multiple equilibria. *Review of Economic Studies 70*, 147–165.

Train, K. E., M. Ben-Akiva, and T. Atherton (1989). Consumption patterns and self-selecting tariffs. *Review of Economics and Statistics 71*, 62–73.

Train, K. E., D. L. McFadden, and M. Ben-Akiva (1987). The demand for local telephone service: A fully discrete model of residential calling patterns and service choices. *RAND Journal of Economics 18*, 109–123.

Vermeulen, F. (2002). Collective household models: Principles and main results. *Journal of Economic Surveys 16*, 533–564.