Consistent Estimation of Peer Effects: Application to Korean Panel Data

Minjung Park[1]

Department of Economics, Ewha Woman's University

52 Ewhayeodae-gil Seodaemun-gu Seoul, Korea

Abstract

This paper examines intra-household peer effects in the adoption of smartphones using unique South Korean panel data. Consistent estimation of peer effects in this setting needs to address two key challenges: homophily and endogenous sample attrition. I show that first-differencing the individual-level panel data and then using deeper lags of the independent variables as instruments address both challenges, enabling consistent estimation of peer effects. I find that an individual becomes much more likely to adopt a smartphone if other household members have previously adopted one. The analysis also reveals that failure to account for endogenous attrition of individuals after product adoption would lead to a significant under-estimation of peer effects.

Keywords: Peer effects, Smartphone adoption, Endogenous sample attrition

JEL Classifications: D12, C23

[1]Correspondence: Minjung Park, Department of Economics, Ewha Woman's University, 52 Ewhayeodae-gil Seodaemun-gu Seoul, Korea. Email: minjung_park@ewha.ac.kr. The work was partly conducted when the author was an assistant professor at the Haas School of Business, University of California, Berkeley. The usual disclaimer applies.

1 Introduction

Whether an individual's product adoption is influenced by the behavior among the individual's reference group is a very important topic, and many researchers in marketing have addressed this question in various settings (Manchanda et al., 2008; Tucker, 2008; Nair et al., 2010; Iyengar et al., 2011; Bollinger and Gillingham, 2012; Narayanan and Nair, 2013; Ma et al., 2015). The economics literature on peer effects has long noted key challenges to correct inference of peer effects, a major challenge being the difficulty to separate peer effects from other confounding factors such as homophily (Manski, 1993; Moffitt, 2001; Brock and Durlauf, 2001). Park (2019) also discusses another challenge that arises when the researcher examines peer effects in product adoption, where the fact that individuals drop out of the estimation sample after adopting the product leads to endogenous sample attrition, generating selection bias in the estimates of peer effects.

In addition to these challenges to inference, another practical challenge that arises in most studies on peer effects is the difficulty of defining the reference group, or social network, over which peer effects should be measured. To define the reference group for product adoption, researchers have relied on geography (e.g., in physicians' drug adoption in Manchanda et al. (2008) and adoption of Prius in Narayanan and Nair (2013)), surveys (e.g., in physicians' drug adoption in Nair et al. (2010) and Iyengar et al. (2011)) or observed communication patterns (e.g., in adoption of a video conferencing technology in Tucker (2008) and adoption of caller ring-back tones in Ma et al. (2015)), among others. While these approaches are in general sensible, there could be some degree of arbitrariness in choosing the boundary for social network (in case of geography) or a possible concern that the reported social ties are formed by the agents as a result of product adoptions (in case of surveys or observed communication patterns), which would invalidate the interpretation of the estimates as causal impact of social ties on product adoption.

In this paper, I examine within-household peer effects in smartphone adoption using unique panel data from Korea Media Panel Survey. The data contain detailed information on smartphone ownership by individual members of households, for more than 2000 households over the span of 5 years. Because there is no ambiguity regarding the boundary of a household and also because it is extremely unlikely that a family formation is affected by smartphone purchase decisions, the dataset enables me to define the reference group in a very clean way. Furthermore, the individual-level panel data enable me to address the key challenges of homophily and endogenous sample attrition in a reasonable way. Exploiting the individual-level panel data, I use first-differencing to difference out group-level fixed effects, thereby accounting for homophily. In the widely used specification where adoption choices of other members are used to measure the installed base, first-differencing within an individual does not lead to a violation of the strict exogeneity condition required for consistent estimation of panel data models, as long as idiosyncratic shocks are uncorrelated across individuals. Therefore, in typical settings with the installed base specification, one would be able to estimate peer effects consistently using first-differencing of individual-level panel data.

However, when a researcher examines peer effects in the context of product adoption, a new challenge arises after first-differencing due to endogenous attrition of individuals over time. Those that adopt the product due to favorable idiosyncratic shocks will be no longer used in estimation after adopting the product. Therefore, the fact that someone is still observed in period t imposes restrictions on the level of her idiosyncratic shocks in previous periods, violating the assumption that the conditional expectation of the first-differenced error term should be zero. Due to the nature of endogeneity the first-differenced error term is correlated with all first-differenced regressors, and finding instrumental variables (IVs) for all of the first-differenced regressors is in general very difficult. In the considered setting, however, I show that using deeper lags of the regressors as IV addresses the problem, borrowing the insights from Anderson and Hsiao (1981) and Arellano and Bond (1991). Unlike the original setting considered by Anderson and Hsiao (1981) and Arellano and Bond (1991), the deeper lags are here used to address endogeneity owing to sample selection, rather than as a solution to endogeneity owing to the inclusion of lagged dependent variable as a regressor (i.e., dynamic panel model).

IV estimates from the first-difference model show that individuals' decision on whether to adopt a smartphone is highly influenced by other household members' adoption behavior. Specifically, the estimates indicate that an individual becomes 11.5 percentage points more likely to adopt a 3G smartphone when one more family member has previously adopted one. Given that 45% of the individuals in the sample adopt a 3G smartphone during the sample period, this amounts to a 25% increase in the probability of adoption. Empirical analysis also reveals that failure to account for endogenous sample attrition (i.e., using first-differencing without IV) would significantly under-estimate peer effects.

The literature on peer effects in product adoption is large and growing both in marketing and economics. While many papers have examined social influence coming from friends and professional networks, analysis on peer effects within a household, which I call intra-household peer effects, has been very scarce. This is somewhat surprising, given that frequent and direct interactions among family members may lead to strong imitation desire and social learning, which are commonly discussed mechanisms for peer effects (Bursztyn et al., 2014). The lack of work on intra-household peer effects might be due to the difficulty of finding household-level panel data that contain information on product adoption or ownership by each member of the household. The lack of work on intra-household peer effects might be also partly due to the fact that for some products adoption is at the household level rather than individual level. For instance, adoption of solar panels or Prius would be best considered as adoption by the household since multiple members of the household can co-consume the product. In case of smartphones, it makes sense to consider adoption as an individual choice, because rarely two family members jointly consume a smartphone.

This paper fills a gap in the literature by studying the under-explored topic of peer effects within a household, which is likely to be a highly important social network for many people. To my knowledge, the only paper that examines intra-household peer effects in product adoption is Huang (2013), who studies intra-household correlation in demand for cellular phone service. While the focus and product of interest are similar between this paper and Huang (2013), the approach is very different. Unlike this paper, Huang (2013) uses cross-sectional data on cellular phone subscription within households, and as a result is severely limited in the ability to distinguish causal intra-household peer effects from confounding effects. Framing the subscription decisions as a simultaneous game within a household, Huang (2013) focuses on simultaneity of decision making and semiparametric maximum likelihood estimator for the game. This paper instead relies on a simpler single-agent estimating equation by exploiting the panel structure of the data and also by abstracting away from possible simultaneity of decision making through the use of installed base specification.

This paper also contributes to the large literature that studies various approaches for consistent estimation of peer effects (Bell and Song, 2007; Bollinger and Gillingham, 2012; Narayanan and Nair, 2013; Matos et al., 2014; Ma et al., 2015). By cleverly combining first-differencing of individual-level panel data with the use of deeper lags of regressors as IV, I propose a simple and attractive approach that enables consistent estimation of peer effects for researchers who want to estimate peer effects for product adoption.

The rest of the paper proceeds as follows. I describe the data in Section 2 and the empirical model in Section 3. In Section 4, I present model estimates and provide discussion. I conclude the paper in Section 5.

2 Data

A smartphone runs on a mobile operating system which enables many of the complicated functions of a personal computer, including the ability to run third-party applications. Smartphones are a significant improvement over feature phones, whose functions are limited to basic multimedia and internet capabilities in addition to voice calls and texts.

The first smartphone that achieved a widespread commercial success was Apple's iPhone. iPhone was introduced in the US in 2007. Since the introduction of iPhone, the smartphone penetration rate continued to climb. For example, 79.3% of the US population had a mobile phone in 2014 and among those, 67.6% were smartphone users with the rest being non-smartphone users such as feature phones (eMarketer, 2015).

In South Korea, which is the focus of this study, iPhone was released in late 2009, much later compared to the US. Despite the late introduction of iPhone, the smartphone penetration rate climbed very quickly in South Korea. As of 2014, 82.6% of the population had a mobile phone and 79.5% of them were smartphone users (eMarketer, 2015).

The dataset used in the study is annual household surveys called Korea Media Panel Survey. The data are collected by Korea Information Society Development Institute, a government-sponsored research institute. In June of every year, survey questionnaires are filled out by participants. To ensure a high participation rate, each household and each individual receive a small monetary reward for participation, and participants are also entered into a lottery for prizes such as laptops and gift cards (Detailed information on the dataset can be found in Lee (2019)).

Each survey consists of two sets of questionnaires – a household survey questionnaire completed by a head of the household and an individual survey questionnaire completed by each household member aged above six. In the household survey, I observe a range of household demographic information, such as household income and metropolitan area the household resides in. In the individual survey, I observe individual demographic information including age, gender, income, employment status, education level and occupation.

As the surveys are designed to gather information on ownership and usage of various technology devices, detailed information is available on mobile phone ownership and usage, such as the type of mobile phones owned by each household member, year of purchase for the currently owned phone, monthly charges, amount spent on mobile app downloads, etc.

The sample covers 5 years of data from 2011 to 2015. Since panel structure is crucial for the ability to estimate peer effects as cleanly as possible, I focus on individuals that appear in all 5 surveys, which is 73% of the individuals included in the 2015 survey, with the discrepancy explained by the addition of new households and their members in the 2015 survey as well as temporary non-response of the 2015 survey participants in some of the earlier years. I also restrict the sample to individuals who belong to a household with a minimum of 2 members in each of the 5 years, since the presence of other members in the same household is crucial to measure intra-household peer effects. The final sample includes 5672 individuals from 2191 households over 5 years. Tables 1 and 2 report summary statistics for the sample.

[Table 1 about here]

In Table 1, we see that there are three different 'generations' of smartphone technologies observed in the sample: 3G, LTE (Long-Term Evolution), and LTE-A (LTE Advanced). Simply put, the LTE-A technology offers faster internet than the LTE technology, which in turn runs faster than the 3G technology. The faster speeds mean that websites will load quicker and video streaming will suffer less buffering.[2] While smartphones based on the 3G technology were introduced before the beginning of the sample period, LTE smartphones were first introduced in South Korea in late 2011, and LTE-A smartphones were first introduced in late 2013. From Table 1, we see that LTE smartphones and LTE-A smartphones first appear in the dataset in 2012 and 2014, respectively, and it is due to the fact that the surveys are conducted in June of every year.

Generally speaking, South Koreans are keenly interested in new smartphone technologies, and major wireless service providers as well as handset makers such as Samsung and LG run very heavy advertising campaigns when a new generation of smartphone technology, and accordingly a new generation of smartphones, become available. Due to this combination of factors, smartphones of an older generation technology tend to disappear quickly as people adopt smartphones of a newer technology rapidly, and this can be seen in Table 1. In less than two years after the introduction of LTE smartphones, about twice as many people newly adopt LTE smartphones than 3G smartphones (the number of new adopters in 2013 is 1250 vs. 633 for LTE smartphones vs. 3G smartphones). In less than two years after the introduction of LTE-A smartphones, more people newly adopted LTE-A smartphones than LTE smartphones (the number of new adopters in 2015 is 725 vs. 691 for LTE-A smartphones vs. LTE smartphones). Overall, the fraction of the individuals who use a smartphone of any kind (3G, LTE or LTE-A) continually rises throughout the sample period: 15.7%, 41.6%, 59.6%, 68.2%, and 73.5% of the sample own a smartphone

[2] In South Korea, downloading speed on LTE smartphones is typically considered to be about 5 times faster than that on 3G smartphones. Downloading speed on LTE-A smartphones is typically twice as fast as that on LTE smartphones.

in year 2011, 2012, 2013, 2014 and 2015, respectively.

In Table 2, I report the characteristics of smartphone users and non-users separately. The table shows that adoption of a new technology first occurs among young people and spreads to older people over time. The average age for the set of people who have not adopted any smartphone ('Non-Users' in the table) continually increases, from 43.7 in 2011 to 62.6 in 2015. The fraction of females is higher among smartphone non-users than among smartphone users.[3] People with higher personal income tend to adopt a smartphone earlier than those with lower personal income. Furthermore, people with higher income are more likely to adopt a smartphone of the latest technology. In 2014 and 2015, the average personal income of LTE-A smartphone users is greater than the average personal income of LTE smartphone users, which is in turn greater than that of 3G smartphone users. Similar patterns hold with respect to household income.

[Table 2 about here]

The literature has discussed two main mechanisms that can generate peer effects: the desire to imitate and learning about product benefits as well as costs (such as difficulty of using the product) from others' adoption (Banerjee, 1992; Bikhchandani et al., 1992; Bernheim, 1994; Duflo and Saez, 2003; Kremer and Miguel, 2007; Oster and Thornton, 2012; Bursztyn et al., 2014). While the empirical literature on peer effects has mostly focused on social networks outside family ties, it seems reasonable to expect that those mechanisms likely operate within households as well. Family ties can be classified as

[3] There are more females than males in the sample, which explains why the fraction of females is greater than 0.5 for all 4 categories in 2015.

strong ties, according to the terminology of Granovetter (1973), and it seems plausible to posit that the desire to imitate and the possibility of learning among family ties might be as high, if not higher, as among other strong ties such as close friends. Family members spend a significant amount of time together, and information flows and idea exchanges naturally occur on a frequent basis for most family members.

Specifically in case of smartphone adoption, the high upfront costs of smartphones would give people a strong incentive to rely on the opinions and experiences of trusted peers to gather sufficient information about the benefits of the product. The chance of going over the data allowance adds to the uncertainty of actual costs of using a smartphone, which would create an additional incentive to see how owning a smartphone works out for other people before making one's own purchase decision. Also, seeing other family members use a new smartphone with more versatile features might prompt one to consider upgrading a phone herself by activating the desire to imitate or simply by activating willingness to search. These discussions suggest a priori reasons why we might expect one's smartphone adoption choice to be influenced by other household members' decisions.

Note that in this paper the term 'adoption' is used to refer to the *first-time* adoption of a smartphone with a given generation of technology. Although rare, it is possible that an individual adopts a 3G smartphone in one year, and then switches to another 3G smartphone a few years later. In this paper, the first-time purchase of a 3G smartphone will be considered as the individual's adoption of 3G smartphones, while the second-time purchase will not be considered as adoption of 3G smartphones. This definition of adoption is reasonable for this paper's research question, because the influence of peers' choices would be most relevant for the first-time adoption of the product while the subsequent purchases of a smartphone with the same generation of technology are likely to be mostly determined by the individual's own prior experiences.

In Table 3, I examine whether there is evidence of clustered adoption behavior, which is a necessary but not sufficient condition for the presence of intra-household peer effects. For this table, I focus on two-person households, and compute the fraction of individuals who ever adopted a smartphone during the sample period, for each generation of smartphones (3G, LTE and LTE-A). 45.3%, 49.6% and 20.7% of the individuals in that restricted sample ever adopted a 3G smartphone, LTE smartphone and LTE-A smartphone, respectively. The sum of the three numbers exceeds 100% since some people adopt more than one generation of smartphones at different points in time during the sample period. In Table 3, I report the fraction of two-person households that are expected to have zero adopter, one adopter and two adopters, respectively, if each household member's decision on whether to adopt a smartphone is independent from other members' decisions. I also report the fraction of two-person households that have zero adopter, one adopter and two adopters, respectively, that are actually observed in the data. If household members tend to make similar choices, the observed fractions for zero-adopter households and two-adopter households should be greater than the respective fractions predicted under the assumption of independent choices, while the observed fraction for one-adopter households should be less than the predicted fraction.

[Table 3 about here]

From the table, we see that the observed fraction exceeds the predicted fraction for zero-adopter households and two-adopter households for each generation of smartphones, while the observed fraction of one-adopter households is far less than the predicted fraction. The χ^2 statistic for the test of independence has an essentially zero p-value for all three generations of smartphones, indicating that smartphone adoption choices are correlated across family members.

Obviously, clustered behavior does not necessarily imply the presence of intra-household peer effects, since there could be common traits among household members that lead to similar choices (homophily). Table 3 simply shows that the data are consistent with the presence of intra-household peer effects, and we will try to control for other sources of correlated behavior through regressions in the next section. In the empirical analysis, I will focus on 3G and LTE smartphones only, because LTE-A smartphones were introduced very late in the sample period. Instead of pooling data on adoption of 3G and LTE smartphones, I will examine adoption decision for 3G smartphones and adoption decision for LTE smartphones separately.

3 Empirical Model

Each individual in the model decides whether to adopt a smartphone or not in each period, if she has not yet adopted. This decision is modeled as a binary choice. The latent utility from adoption for individual i in period t is given by

$$U_{it} = \beta_0 + \beta_1 b_{h(i)t} + \beta_2 x_{it} + \beta_3 z_{h(i)t} + \omega_i + \omega_{h(i)} + \omega_t + \epsilon_{it}. \tag{1}$$

$h(i)$ denotes the household individual i belongs to, and $b_{h(i)t}$ represents the installed base of adopters in household $h(i)$ in period t, e.g., the number or fraction of other household members who adopted the product before period t. x_{it} represents characteristics of i in period t, such as age, employment status, and personal income, and $z_{h(i)t}$ represents characteristics of i's household h in period t, such as household income and region of residence. $b_{h(i)t}$, x_{it} and $z_{h(i)t}$ are observed by the researcher, and those that are time-varying, such as income and employment status, will still appear in the equation after first-differencing, while time-invariant characteristics such as gender will drop from the equation after first-differencing.

ω_i represents i's valuation of the product, assumed to be iid across i and independent of x_{it} and $z_{h(i)t}$ in the underlying population (i.e., random effects), while $\omega_{h(i)}$ measures product valuation of i's household h and is allowed to be correlated with x_{it} and $z_{h(i)t}$ in any flexible fashion (i.e., fixed effects). Inclusion of $\omega_{h(i)}$ allows for homophily, a phenomenon that people in the same group share similar preferences and thus tend to behave similarly. I assume random effects for ω_i because individual fixed effects (i.e., ω_i which can be correlated with x_{it} and $z_{h(i)t}$ in any flexible fashion) are not possible in the considered setting, given that each individual experiences no more than one adoption event. Since ω_i and $\omega_{h(i)}$ are time-invariant, they will disappear from the equation after first-differencing. ω_t represents year effects common to all individuals in the sample, and first-differencing will transform year effects into time trend effects.

ϵ_{it} represents an idiosyncratic shock that influences i's adoption decision in period t, which is neither observed by the researcher nor captured by the included regressors. ϵ_{it} is assumed iid across i and t. As will become clear later, the assumption of no correlation in ϵ_{it} across i and t is crucial for the validity of the estimation procedure proposed in the paper. There are N individuals, $i = 1, ..., N$, and T time periods, $t = 1, ..., T$. Each individual belongs to only one household, and belongs to the same household in all periods.

Individual i will adopt a smartphone in period t if and only if the utility from adoption is greater than the utility from no adoption, which is normalized to zero. Thus, the adoption choice can be written as

$$y_{it} = 1[\beta_0 + \beta_1 b_{h(i)t} + \beta_2 x_{it} + \beta_3 z_{h(i)t} + \omega_i + \omega_{h(i)} + \omega_t + \epsilon_{it} > 0], \tag{2}$$

where y_{it} is a binary variable that takes a value of one if individual i adopts a smartphone in period t and zero otherwise. Since those who already adopted the product no longer make an adoption decision, only those who have not yet adopted before period t will comprise the estimation sample for period t. As a result, the estimation sample will get increasingly smaller, as more people adopt the product and are dropped from the estimation sample over time.

Peer effects in (2) are specified as the influence of past adoption decisions of other household members on i's current adoption choice. The specification with the installed base abstracts away from possible simultaneity of decision making, and is extensively used in models of network contagion, product diffusion, and peer effects, e.g., Bass (1969), Manchanda et al. (2008), Iyengar et al. (2011), Bollinger and Gillingham (2012), and Narayanan and Nair (2013).

I also chose a specification that ignores the possibility of exogenous peer effects (what Manski (1993) called contextual effects) in the analysis, i.e., characteristics of other household members do not directly enter Eq. (2), because it seems highly unlikely that individuals' smartphone adoptions are directly influenced by the exogenous characteristics (such as age and sex) of their family members. In other contexts, exogenous peer effects might be relevant. For instance, if the product under consideration is a video conferencing technology as in Tucker (2006) and Ryan and Tucker (2012), the utility of adopting the technology could depend on characteristics of other members in the same group, e.g., my desire to adopt the video conferencing technology could be greater if many members in my network work in regional offices that are far away.

The key parameter of interest is the parameter on the installed base $b_{h(i)t}$, β_1, which captures the degree of intra-household peer influence on i's product adoption. There are various threats to correctly identifying β_1. One commonly discussed confounding factor that challenges identification of peer effects is homophily. Family ties are not randomly assigned, and people with similar preferences tend to form a family. Thus, family members might have similar preferences for the latest technology products, in which case i's propensity to adopt a smartphone would be positively correlated with other family members' adoption behavior even in the absence of (causal) peer effects. To address this concern, group fixed effects $\omega_{h(i)}$ are included in the above equation.

Another challenge to identification of causal peer effects is the possible presence of correlated time-varying unobservables. If members of a group experience similar shocks to their propensity to adopt in ways that are unobservable to researchers, and such shocks are serially correlated (i.e., a shock to another member in period $t-1$ is correlated with a shock to i in period t), it will generate bias in the estimates of peer effects. The likely source of such a shock is changes in family income and employment status of the head of the household. Since data on family income and employment status of the head of the household are available, I can directly control for them in the regression, and I will assume that correlated time-varying unobservables are no longer a concern once such pieces of information are controlled for.

For estimation, I take first-differencing of Eq. (2) between two consecutive years for a given individual. To facilitate first-differencing as well as use of instrumental variables below, a linear probability model is assumed, which is equivalent to assuming a uniform distribution on ϵ_{it}. First-differencing of a linear probability specification for Eq. (2) will result in the following:

$$\Delta y_{it} = \beta_1 \Delta b_{h(i)t} + \beta_2 \Delta x_{it} + \beta_3 \Delta z_{h(i)t} + \Delta\omega_t + \Delta\epsilon_{it}, \tag{3}$$

where $\Delta y_{it} = y_{it} - y_{it-1}$, $\Delta b_{h(i)t} = b_{h(i)t} - b_{h(i)t-1}$, $\Delta x_{it} = x_{it} - x_{it-1}$, $\Delta z_{h(i)t} = z_{h(i)t} - z_{h(i)t-1}$, $\Delta\omega_t = \omega_t - \omega_{t-1}$ and $\Delta\epsilon_{it} = \epsilon_{it} - \epsilon_{it-1}$.

First-differencing removes both group fixed effects ($\omega_{h(i)}$) and individual random effects (ω_i), thereby enabling estimation of peer effects that is not contaminated by time-invariant unobserved heterogeneity either at the group level or individual level. Because the model in (1) uses adoption choices of *other* members to measure the installed base, first-differencing within an individual does not lead to a violation of the strict exogeneity condition required for consistent estimation of panel data models, as long as idiosyncratic shocks are uncorrelated across individuals.

However, a new challenge arises due to the endogenous attrition of individuals. Park (2019) shows that endogenous sample attrition could lead to a downward bias in the estimates of peer effects, borrowing key insights from Heckman (1979) that sample selection bias can be viewed as an omitted variable bias, and a similar issue arises in this paper as well. Specifically, in the first-difference equation (3), those that adopted the product due to favorable idiosyncratic shocks ϵ_{i1}, ϵ_{i2}, ..., ϵ_{it-1} are no longer used in estimation for period t adoption decision. Therefore, the fact that someone is still observed in period t imposes restrictions on the level of her idiosyncratic shocks in previous periods, violating the assumption that the conditional expectation of the first-differenced error term should be zero. Furthermore, if individual i is still observed in period t (i.e., has not yet adopted the product by the end of period $t-1$) although the values of her regressors in period $t-1$ were favorable for adoption, it allows us to infer that she had a particularly unfavorable realization of ϵ_{it-1}, resulting in a negative correlation between ϵ_{it-1} and the values of her regressors in period $t-1$. As a result, $\Delta\epsilon_{it} = \epsilon_{it} - \epsilon_{it-1}$ is correlated with all first-differenced regressors, conditional on individual i being observed in period t, i.e., $E[\Delta\epsilon_{it} \mid \Delta b_{h(i)t}, \Delta x_{it}, \Delta z_{h(i)t}, \text{observed in } t] \neq 0$, which would lead to inconsistent estimation of peer effects unless further actions, such as use of instrumental variables, are taken. In particular, we can expect that the endogenous sample attrition would lead to a downward bias in the estimates of peer effects given the negative correlation between the residual and the regressors in the first-difference equation.

Since the first-differenced error term $\Delta\epsilon_{it}$ is correlated with all of the first-differenced regressors in (3), we would need to find instrumental variables for all of the first-differenced regressors, which is in general very difficult. In the considered setting, however, finding instruments for all the first-differenced regressors is feasible because of the nature of endogeneity. As discussed, endogeneity arises due to correlation between ϵ_{it-1} and ($b_{h(i)t-1}$, x_{it-1}, $z_{h(i)t-1}$) among those that survive until period t. Borrowing the insights from Anderson and Hsiao (1981) and Arellano and Bond (1991), I propose to use deeper lags of the regressors to construct instruments.

In the original setting considered by Anderson and Hsiao (1981) and Arellano and Bond (1991), they consider a panel data model where one of the regressors is a lagged dependent variable y_{it-1} and first-differencing is performed to remove fixed effects (unobserved heterogeneity). The presence of the lagged dependent variable among the regressors of the first-difference model creates endogeneity due to correlation between y_{it-1} and ϵ_{it-1}, and they propose to use deeper lags of regressors, including deeper lags of the dependent variable (y_{i1}, y_{i2}, ..., y_{it-2}), to generate instruments for $\Delta y_{it-1} = y_{it-1} - y_{it-2}$. In other words, deeper lags of the regressors are used to address endogeneity owing to the inclusion of lagged dependent variable as a regressor (i.e., dynamic panel model).

Our setting differs because the installed base is a function of others' past adoption choices only and does not depend on the focal individual's past choice. In other words, our model is not a dynamic panel model. Rather, endogeneity in our setting arises due to endogenous sample attrition, where attrition depends on the lagged dependent variable y_{it-1}. However, I argue that we can borrow the same intuition from Anderson and Hsiao (1981) and Arellano and Bond (1991) and use deeper lags of the regressors ($b_{h(i)t-2}$, x_{it-2}, $z_{h(i)t-2}$) to construct instruments. Under the assumption of no serial correlation in ϵ_{it}, the variables $\Delta_2 b_{h(i)t} = b_{h(i)t} - b_{h(i)t-2}$, $\Delta_2 x_{it} = x_{it} - x_{it-2}$, and $\Delta_2 z_{h(i)t} = z_{h(i)t} - z_{h(i)t-2}$ are uncorrelated with $\Delta\epsilon_{it} = \epsilon_{it} - \epsilon_{it-1}$, even if we just look at those that survive until period t, while they are likely to be highly correlated with $\Delta b_{h(i)t}$, Δx_{it}, and $\Delta z_{h(i)t}$,. Thus, deeper lags of the regressors can be used to address endogeneity owing to sample selection in our setting.

It is clear that the validity of the proposed instruments crucially relies on the assumption of no serial correlation in ϵ_{it}. If ϵ_{it} is serially correlated, $b_{h(i)t-2}$, x_{it-2}, and $z_{h(i)t-2}$, which are correlated with ϵ_{it-2} among those that survive until period $t-1$, will be also correlated with ϵ_{it-1}, rendering the deeper lags of the regressors invalid as instruments. If time series is sufficiently long, it is also possible to create additional instruments by using further lags. For instance, one can add $\Delta_3 b_{h(i)t} = b_{h(i)t} - b_{h(i)t-3}$, $\Delta_3 x_{it} = x_{it} - x_{it-3}$, and $\Delta_3 z_{h(i)t} = z_{h(i)t} - z_{h(i)t-3}$ to the set of instruments, in which case the model would be over-identified instead of being exactly identified.

A few remarks are in order regarding the first-difference equation (3). First of all, it is obvious that prices and features of smartphones are important determinants of one's smartphone adoption, so one might wonder why they do not appear in the estimating equation. The reason is that smartphone prices and features are mostly common to all individuals and thus are absorbed by the year dummies. As a result, an increase in the probability of smartphone adoption due to lower prices or better features over time would be captured by the time trend effects in Eq. (3).

Second, an individual's smartphone adoption could be influenced by many people in her various social networks, and family is only one of the many social networks a person belongs to. Thus, one might wonder whether omitting other social ties from the analysis might create a bias in the estimates of intra-household peer effects. Essentially, the omitted social ties become part of the error term, so the question boils down to whether the *change* in the level of the installed base within the person's other social networks is correlated with the *change* in the level of the installed base within her family. The shocks that determine the change in the level of the installed base within the household are things like changes in other family members' income. It is hard to argue why these shocks would be correlated with the shocks that determine the change in the installed base among her other social ties such as friends or co-workers. Therefore, the omission of other social ties from the analysis is unlikely to cause a bias in the estimates of intra-household peer effects obtained from Eq. (3).

Finally, while a bias arises due to endogenous sample attrition in both Park (2019) and this paper, there are key differences in the exact source of endogeneity. In Park (2019), the original model was used for estimation without any first-differencing, and the focus was on endogeneity due to correlation between time-invariant, unobserved individual heterogeneity (ω_i) and the regressors resulting from endogenous sample attrition. In this paper, first-difference model in (3) is used for estimation and endogeneity arises due to correlation between time-varying idiosyncratic shocks (ϵ_{it}) and the regressors resulting from endogenous sample attrition.

4 Estimation Results

In this section, I present estimates of peer effects under various specifications. First, I estimate the first-difference equation (3) using a linear regression without any IV. The resulting estimates of peer effects will be robust to possible bias from homophily, but are likely to suffer from bias due to endogenous sample attrition. Second, I estimate Eq. (3) using ($\Delta_2 b_{h(i)t}$, $\Delta_2 x_{it}$, $\Delta_2 z_{h(i)t}$) as instruments for ($\Delta b_{h(i)t}$, Δx_{it}, $\Delta z_{h(i)t}$). Estimation is done using the Generalized Method of Moments. The instruments are valid as they are likely to be highly correlated with the instrumented regressors, while being uncorrelated with the error term $\Delta\epsilon_{it}$ under the assumption of no serial correlation in ϵ_{it}. Therefore, the resulting estimates of peer effects will represent true causal peer influence, free from both homophily bias and sample selection bias. Comparison between the results from the first and second specifications allows us to infer the direction and size of the sample selection bias.

Third, I estimate Eq. (3) using a few alternative sets of instruments: (A1) using ($\Delta_2 b_{h(i)t}$, $\Delta_2 x_{it}$, $\Delta_2 z_{h(i)t}$, $\Delta_3 x_{it}$) as instruments, (A2) using ($\Delta_2 b_{h(i)t}$, $\Delta_2 x_{it}$, $\Delta_2 z_{h(i)t}$, $\Delta_3 x_{it}$, $\Delta_3 z_{h(i)t}$) as instruments, and (A3) using ($\Delta_2 b_{h(i)t}$, $\Delta_2 x_{it}$, $\Delta_2 z_{h(i)t}$, $\Delta_3 b_{h(i)t}$, $\Delta_3 x_{it}$, $\Delta_3 z_{h(i)t}$) as instruments, respectively. By the same argument as before, the resulting estimates from (A1)-(A3) will reflect bias-free, causal peer influence, if the proposed instruments are valid. I add increasingly more instruments in order to check the sensitivity of the results to different sets of instruments. Also, using an increasingly larger set of instruments enables me to test the validity of different sets of instruments by performing an over-identifying restriction test separately for each specification. For instance, if I find that the over-identifying restriction test does not reject the null hypothesis of correct moment conditions for (A1), while rejecting the null hypothesis for (A2) and (A3), I can infer that the additional instruments in (A2) and (A3), namely $\Delta_3 b_{h(i)t}$ and $\Delta_3 z_{h(i)t}$, might not be valid instruments while the instruments in (A1), namely $\Delta_2 b_{h(i)t}$, $\Delta_2 x_{it}$, $\Delta_2 z_{h(i)t}$ and $\Delta_3 x_{it}$, are likely to be valid instruments.

For all the specifications, I perform estimation twice, once focusing on adoption of 3G smartphones, and then separately focusing on adoption of LTE smartphones. Because the LTE technology is considered to be significantly superior to the 3G technology, it seems sensible to examine adoption decisions for 3G smartphones and LTE smartphones separately, instead of pooling data on both. I also repeat each estimation twice using two different measures of installed base in order to examine robustness of the results. The first measure of installed base is the number of other household members who have previously adopted a smartphone of the considered technology, and the second measure is the fraction of such household members.

Table 4 reports the estimates from the analysis on adoption of 3G smartphones and Table 5 reports the estimates from the analysis on adoption of LTE smartphones. The upper panel in each table reports the estimates from the specifications where the installed base is measured using the number of other household members who have previously adopted a smartphone of the considered technology. The estimates in the lower panel of each table come from the specifications where the installed base is measured using the fraction of such household members. Within each panel, the first, second, third, fourth, and fifth columns report the results obtained when the first-difference equation (3) is estimated using a linear regression without any IV, using ($\Delta_2 b_{h(i)t}$, $\Delta_2 x_{it}$, $\Delta_2 z_{h(i)t}$) as instruments, using ($\Delta_2 b_{h(i)t}$, $\Delta_2 x_{it}$, $\Delta_2 z_{h(i)t}$, $\Delta_3 x_{it}$) as instruments, using ($\Delta_2 b_{h(i)t}$, $\Delta_2 x_{it}$, $\Delta_2 z_{h(i)t}$, $\Delta_3 x_{it}$, $\Delta_3 z_{h(i)t}$) as instruments, and using ($\Delta_2 b_{h(i)t}$, $\Delta_2 x_{it}$, $\Delta_2 z_{h(i)t}$, $\Delta_3 b_{h(i)t}$, $\Delta_3 x_{it}$, $\Delta_3 z_{h(i)t}$) as instruments, respectively. For the last three columns, the model is over-identified, and the χ^2 statistics from the over-identifying restriction tests are reported at the bottom. To facilitate comparison across different specifications within each table, I use the same estimation sample for all specifications, and this means that the first three years of data will not be used for estimation (to construct $\Delta_3 b_{h(i)t}$, $\Delta_3 x_{it}$, and $\Delta_3 z_{h(i)t}$).

[Table 4 about here]

Starting with the upper panel of Table 4, comparison between the first and second columns shows that the estimates of peer effects become more than 70% larger once instruments are used to address the potential bias due to endogenous sample attrition. In other words, failure to account for endogenous sample attrition resulted in a significant downward bias in the estimates of peer effects, a finding similar to that in Park (2019). In terms of the magnitude of peer effects, IV estimates from the second column of the upper panel indicate that an individual becomes 11.5 percentage points more likely to adopt a

3G smartphone when one more family member has previously adopted one. Given that 45.3% of the individuals in the sample adopt a 3G smartphone during the sample period, this amounts to a 25% increase in the probability of adoption.

Comparison across the second, third, fourth, and fifth columns in the upper panel of Table 4 shows that the estimates of peer effects are robust to the exact set of IVs being used. Furthermore, the χ^2 statistics for the over-identifying restriction tests in columns 3-5 fail to reject the null hypothesis of correct moment conditions. This suggests that there is no evidence against the validity of the IVs used in columns 3-5, and given that the IVs used in the second column are a subset of the IVs in columns 3-5, this also implies validity of the IVs used in the second column.

The results in Table 4 also indicate that other coefficient estimates are mostly insignificant. Since the estimates are obtained from the first-difference equation, the time-invariant portion of the variables has been differenced out, so the lack of significance for variables such as personal income or household income does not necessarily mean that those variables do not impact the adoption choice. Rather, the results indicate that changes in those variables between two consecutive periods do not lead to any meaningful change in the adoption probability of the individual (while the level of those variables could very well impact the adoption choice). Similarly, some time-invariant characteristics such as gender and education level are likely to impact adoption choices and we allow that possibility, but their effects cannot be estimated from the first-difference equation. Finally, variables that are time-varying but common to all individuals in the sample, such as prices and features of smartphones, and variables that change in a deterministic fashion that is common to all individuals, such as age, are allowed to impact the adoption choice, but we cannot estimate their effects separately as they are absorbed by the time trend term in the first-difference equation.

The results in the lower panel of Table 4, where the installed base is measured using the fraction of other household members who have previously adopted a 3G smartphone, are overall similar to the results from the upper panel. The only noticeable difference is that the χ^2 statistic for the over-identifying restriction test in column 5 of the lower panel rejects the null hypothesis of correct moment conditions. Since we fail to reject the null hypothesis in columns 3 and 4 of the lower panel, we can infer that the instrument that was additionally used in column 5, namely $\Delta_3 b_{h(i)t}$, might be the culprit that leads to a rejection of the null hypothesis, while the other instruments are OK. As a result, we expect that the IV estimates in columns 2-4 of the lower panel are likely to represent true causal peer effects within the household.

Table 5 reports the estimates from the analysis on adoption of LTE smartphones, and the results are qualitatively similar to those in Table 4. Comparison between the first and second columns in the upper panel of Table 5 indicates that failure to account for endogenous sample attrition resulted in a significant downward bias in the estimates of peer effects for adoption of LTE smartphones. In terms of the magnitude of peer effects, IV estimates from the second column of the upper panel indicate that an individual becomes 9.4 percentage points more likely to adopt an LTE smartphone when one more family member has previously adopted one. Given that 49.6% of the individuals in the sample adopt an LTE smartphone during the sample period, this amounts to a 19% increase in the probability of adoption.

[Table 5 about here]

Comparison across the second, third, fourth, and fifth columns of Table 5 shows that the estimates of peer effects are robust to the exact set of IVs being used. The estimates of peer effects are very stable across columns 2-5. While the χ^2 statistic for the over-identifying restriction test rejects the null hypothesis of correct moment conditions in column 5, the null hypothesis is not rejected for columns 3 and 4, indicating that the instrument that was additionally used in column 5, namely $\Delta_3 b_{h(i)t}$, might be correlated with the error term $\Delta\epsilon_{it}$, but the IV estimates in columns 2-4 are likely to represent true causal peer effects within a household. Other results are overall quite similar to the findings in Table 4, with the only noticeable difference being that the change in personal income ΔInc_t is positive and significant in Table 5.

5 Conclusion

In this paper, I empirically examine whether an individual's smartphone adoption is influenced by the adoption behavior of other household members. Consistent estimation of peer effects is a well-known challenge, and I show that first-differencing the individual-level panel data and then using deeper lags of the independent variables to construct instruments enable consistent estimation of peer effects. The empirical analysis indicates that the probability of smartphone adoption increases by 19-25% when one more family member has previously adopted one. Given the increasing availability of individual-level panel data, I expect the proposed method to become appealing to many researchers who are interested in estimation of peer effects for product adoption, where the issues of homophily and endogenous sample attrition are bound to arise.

References

[1] Anderson, T. and Hsiao, C. 1981. “Estimation of Dynamic Models with Error Components.” *Journal of the American Statistical Association*, 76: 598-606.

[2] Arellano, M. and Bond, S. 1991. “Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations.” *Review of Economic Studies*, 58: 277-298.

[3] Banerjee, A. 1992. “A Simple Model of Herd Behavior.” *Quarterly Journal of Economics*, 107(3): 797-817.

[4] Bass, F. 1969. “A New Product Growth for Model Consumer Durables.” *Management Science,* 15(5): 215-227.

[5] Bell, D. R. and Song, S. 2007. “Neighborhood Effects and Trial on the Internet: Evidence from Online Grocery Retailing.” *Quantitative Marketing and Economics,* 5: 361-400.

[6] Bernheim, B. D. 1994. “A Theory of Conformity.” *Journal of Political Economy,* 102(5): 841-877.

[7] Bikhchandani, S., Hirshleifer, D. and Welch, I. 1992. “A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades.” *Journal of Political Economy*, 100(5): 992-1026.

[8] Bollinger, B. and Gillingham, K. 2012. “Peer Effects in the Diffusion of Solar Photovoltaic Panels.” *Marketing Science,* 31(6): 900-912.

[9] Brock, W. and Durlauf, S. 2001. “Discrete Choice with Social Interactions.” *Review of Economic Studies*, 235-260.

[10] Bursztyn, L., Ederer, F., Ferman, B. and Yuchtman, N. 2014. “Understanding Mechanisms Underlying Peer Effects: Evidence From a Field Experiment on Financial Decisions” *Econometrica*, 82(4): 1273-1301.

[11] Duflo, E. and Saez, E. 2003. “The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment.” *Quarterly Journal of Economics*, 118(3): 815-842.

[12] Granovetter, M. 1973. “The Strength of Weak Ties.” *American Journal of Sociology*, 78(6): 1360-1380.

[13] Heckman, J. 1979. “Sample Selection Bias as a Specification Error.” *Econometrica*, 47(1): 153-161.

[14] Huang, CI. 2013. “Intra-household Effects on Demand for Telephone Service: Empirical Evidence.” *Quantitative Marketing and Economics*, 11(2): 231-261.

[15] Iyengar, R., Van den Bulte, C., and Valente, T. 2011. "Opinion Leadership and Social Contagion in New Product Diffusion." *Marketing Science,* 30(2): 195-212.

[16] Kremer, M. and Miguel, E. 2007. "The Illusion of Sustainability." *Quarterly Journal of Economics*, 122(3): 1007-1065.

[17] Lee, S. 2019. "Quantifying the Benefits of Smartphone Adoption: Digital Device Substitution and Digital Consumption Expansion." Working Paper.

[18] Ma, L., Krishnan, R. and Montgomery, A. 2015. "Latent Homophily or Social Influence? An Empirical Analysis of Purchase Within a Social Network." *Management Science*, 61(2): 454-473.

[19] Manchanda, P., Xie, Y. and Youn, R. 2008. "The Role of Targeted Communication and Contagion in Product Adoption." *Marketing Science,* 27(6): 961-976.

[20] Manski, C. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *Review of Economic Studies*, 60(3): 531-542.

[21] Matos, M., Ferreira, P. and Krackhardt, D. 2014. "Peer Inflence in the Diffusion of Iphone 3G over a Large Social Network." *MIS Quarterly*, 38(4): 1103-1133.

[22] Moffitt, R. 2001. "Policy Interventions, Low-Level Equilibria, and Social Interactions." In: Durlauf S., Young P. (eds.), *Social Dynamics,* 45-82.

[23] Nair, H., Manchanda, P. and Bhatia, T. 2010. "Asymmetric Social Interactions in Physician Prescription Behavior: The Role of Opinion Leaders." *Journal of Marketing Research*, 47(5): 883-895.

[24] Narayanan, S. and Nair, H. 2013. "Estimating Causal Installed-Base Effects: A Bias-Correction Approach." *Journal of Marketing Research*, 50(1): 70-94.

[25] Oster, E. and Thornton, R. 2012. "Determinants of Technology Adoption: Peer Effects in Menstrual Cup Take-up." *Journal of the European Economic Association*, 10: 1263-1293.

[26] Park, M. 2019. "Selection Bias in Estimation of Peer Effects in Product Adoption." *Journal of Choice Modelling*, 30: 17-27.

[27] Ryan, S and Tucker, C. 2012. "Heterogeneity and the Dynamics of Technology Adoption." *Quantitative Marketing and Economics*, 10(1): 63-109.

[28] Tucker, C. 2008. "Identifying Formal and Informal Influence in Technology Adoption with Network Externalities." Marketing Science, 54(12): 2024–2038.

Table 1: Number of Smartphone Users and Adopters in the Sample						
Year		2011	2012	2013	2014	2015
Number of Households		2191	2191	2191	2191	2191
Number of Individuals		5672	5672	5672	5672	5672
3G Smartphone:	Users	889	1834	1700	1226	664
	- New Adopters	499	1141	633	299	167
	- Continued Users	390	693	1067	927	497
LTE Smartphone:	Users	0	526	1681	2082	2374
	- New Adopters	0	526	1250	732	691
	- Continued Users	0	0	431	1350	1683
LTE-A Smartphone:	Users	0	0	0	559	1132
	- New Adopters	0	0	0	559	725
	- Continued Users	0	0	0	0	407

Table 2: Characteristics of Smartphone Users vs. Non-users

Year		2011	2012	2013	2014	2015
Age	3G Users	33.5	34.3	37.3	39.9	41.5
	LTE Users	.	34.1	35.5	37.6	40.4
	LTE-A Users	.	.	.	37	38.5
	Non-Users	43.7	49.4	55.4	59.7	62.6
Female Fraction	3G Users	0.44	0.50	0.52	0.53	0.52
	LTE Users	.	0.50	0.49	0.51	0.52
	LTE-A Users	.	.	.	0.49	0.50
	Non-Users	0.54	0.54	0.55	0.54	0.54
Personal Income	3G Users	164.5	140.7	129.0	123.9	121.3
10,000 won / m	LTE Users	.	158.2	149.6	143.4	138.8
	LTE-A Users	.	.	.	156.6	166.9
	Non-Users	92.3	82.5	71.4	67.1	60.9
Household Income	3G Users	373.6	381.8	369.9	378.3	370.4
10,000 won / m	LTE Users	.	394.5	384.7	388.5	394.5
	LTE-A Users	.	.	.	396.8	415.4
	Non-Users	296.7	274.6	247.7	238.6	234.1

Table 3: Clustered Behavior			
	Zero-Adopter HHs	One-Adopter HHs	Two-Adopter HHs
3G Smartphones: Predicted	29.9%	49.6%	20.5%
3G Smartphones: Observed	41.7%	26%	32.3%
χ^2 statistic for independence		309.8 (p-value = 0)	
LTE Smartphones: Predicted	25.4%	50%	24.6%
LTE Smartphones: Observed	39.2%	22.4%	38.4%
χ^2 statistic for independence		416.3 (p-value = 0)	
LTE-A Smartphones: Predicted	62.9%	32.9%	4.3%
LTE-A Smartphones: Observed	71.7%	15.2%	13.1%
χ^2 statistic for independence		392.9 (p-value = 0)	

Table 4: Adoption of 3G Smartphones

	OLS	IV	Alt IV1 (A1)	Alt IV2 (A2)	Alt IV3 (A3)
ΔIB_t	0.035 (0.006)***	0.061 (0.015)***	0.061 (0.015)***	0.058 (0.015)***	0.065 (0.015)***
ΔInc_t	-0.008 (0.005)*	-0.013 (0.010)	-0.012 (0.009)	-0.013 (0.009)	-0.013 (0.009)
ΔEmp_t	0.028 (0.012)**	0.031 (0.021)	0.030 (0.021)	0.029 (0.021)	0.027 (0.021)
ΔHInc_t	-0.002 (0.003)	-0.005 (0.007)	-0.005 (0.007)	-0.004 (0.007)	-0.003 (0.007)
ΔHEmp_t	-0.004 (0.013)	-0.028 (0.025)	-0.027 (0.025)	-0.028 (0.025)	-0.030 (0.025)
χ^2 stat	.	.	0.08	1.88	8.43
Obs	5,867	5,867	5,867	5,867	5,867
	OLS	IV	Alt IV1 (A1)	Alt IV2 (A2)	Alt IV3 (A3)
ΔIB_t	0.051 (0.014)***	0.115 (0.032)***	0.115 (0.032)***	0.113 (0.032)***	0.125 (0.031)***
ΔInc_t	-0.009 (0.005)*	-0.014 (0.010)	-0.013 (0.009)	-0.014 (0.009)	-0.014 (0.009)
ΔEmp_t	0.028 (0.012)**	0.030 (0.021)	0.029 (0.021)	0.028 (0.021)	0.029 (0.021)
ΔHInc_t	-0.002 (0.003)	-0.002 (0.007)	-0.003 (0.007)	-0.001 (0.007)	-0.001 (0.007)
ΔHEmp_t	-0.003 (0.013)	-0.023 (0.025)	-0.022 (0.025)	-0.024 (0.025)	-0.028 (0.025)
χ^2 stat	.	.	0.08	1.91	11.08**
Obs	5,867	5,867	5,867	5,867	5,867

IB is the installed base. Inc is the monthly income of the individual.
Emp is 1 if the individual is employed and 0 otherwise. HInc is the monthly income of the household.
HEmp is 1 if the head of the household is employed and 0 otherwise.
IV in the second column: $\Delta_2\text{IB}_t$, $\Delta_2\text{Inc}_t$, $\Delta_2\text{Emp}_t$, $\Delta_2\text{HInc}_t$, $\Delta_2\text{HEmp}_t$
IV in the third column: add $\Delta_3\text{Inc}_t$ and $\Delta_3\text{Emp}_t$ to the IV set in the second column
IV in the fourth column: add $\Delta_3\text{HInc}_t$ and $\Delta_3\text{HEmp}_t$ to the IV set in the third column
IV in the fifth column: add $\Delta_3\text{IB}_t$ to the IV set in the fourth column.
Time trend dummies are included in all specifications.

Table 5: Adoption of LTE Smartphones					
	OLS	IV	Alt IV1 (A1)	Alt IV2 (A2)	Alt IV3 (A3)
ΔIB_t	0.061 (0.007)***	0.094 (0.011)***	0.095 (0.011)***	0.095 (0.011)***	0.095 (0.011)***
ΔInc_t	0.023 (0.009)**	0.047 (0.020)**	0.056 (0.018)***	0.058 (0.018)***	0.058 (0.018)***
ΔEmp_t	0.019 (0.021)	0.043 (0.039)	0.045 (0.038)	0.044 (0.038)	0.045 (0.038)
$\Delta HInc_t$	0.003 (0.006)	0.017 (0.011)	0.017 (0.011)	0.015 (0.010)	0.016 (0.010)
$\Delta HEmp_t$	0.009 (0.023)	0.021 (0.042)	0.019 (0.042)	0.020 (0.041)	0.025 (0.041)
χ^2 stat	.	.	3.87	4.25	9.64*
Obs	7,060	7,060	7,060	7,060	7,060
	OLS	IV	Alt IV1 (A1)	Alt IV2 (A2)	Alt IV3 (A3)
ΔIB_t	0.154 (0.018)***	0.250 (0.027)***	0.252 (0.027)***	0.252 (0.027)***	0.255 (0.027)***
ΔInc_t	0.022 (0.009)**	0.042 (0.020)**	0.052 (0.018)***	0.053 (0.018)***	0.054 (0.018)***
ΔEmp_t	0.019 (0.021)	0.046 (0.039)	0.048 (0.038)	0.046 (0.038)	0.049 (0.038)
$\Delta HInc_t$	0.004 (0.006)	0.019 (0.011)*	0.019 (0.011)*	0.0181 (0.010)*	0.018 (0.010)*
$\Delta HEmp_t$	0.008 (0.023)	0.022 (0.042)	0.020 (0.042)	0.022 (0.042)	0.026 (0.042)
χ^2 stat	.	.	4.20	4.40	12.63**
Obs	7,060	7,060	7,060	7,060	7,060

IB is the installed base. Inc is the monthly income of the individual.
Emp is 1 if the individual is employed and 0 otherwise. HInc is the monthly income of the household.
HEmp is 1 if the head of the household is employed and 0 otherwise.
IV in the second column: $\Delta_2 IB_t$, $\Delta_2 Inc_t$, $\Delta_2 Emp_t$, $\Delta_2 HInc_t$, $\Delta_2 HEmp_t$
IV in the third column: add $\Delta_3 Inc_t$ and $\Delta_3 Emp_t$ to the IV set in the second column
IV in the fourth column: add $\Delta_3 HInc_t$ and $\Delta_3 HEmp_t$ to the IV set in the third column
IV in the fifth column: add $\Delta_3 IB_t$ to the IV set in the fourth column.
Time trend dummies are included in all specifications.