Necessary Luxuries: A New Social Interactions Model, Applied to Keeping Up With the Joneses in India

Arthur Lewbel, Samuel Norris, Krishna Pendakur, and Xi Qu

Boston College, Northwestern U., Simon Fraser U., and Shanghai Jiao Tong U.

original July 2014, revised Dec 2015 - PRELIMINARY AND INCOMPLETE

Abstract

We propose a new model that introduces peer effects into existing utility models of perceived needs. This combination introduces obstacles to obtaining identification that differ both from standard consumer demand and from standard models of peer effects. These obstacles arise from required nonlinearities in utility, from features of standard consumption survey data, and from heterogeneity that requires group level fixed or random effects. We first provide identification and an associated estimator for a new generic peer effects model that allows for nonlinearities, fixed effects, and the data feature that only a small number of the members of each peer group are observed. We then extend this model to our consumer demand application. We obtain estimates of the dollar costs of what is spent on keeping up with others in one's group. These estimates have important tax policy implications, since the larger these peer effects are, the smaller are the welfare gains associated with tax cuts. We find that, in our data from India, peer effects are important for luxuries and not necessities, and about 15% of income growth in India is spent on keeping up with one's peers.

1 Introduction

We propose a new model of social interactions (peer effects) in consumption, including a new method for identification and related estimation of a general class of social interactions models. Novel features of our general model include the inclusion of fixed effects or random effects, the use of nonlinearity to overcome reflection and other obstacles to identification, and applicability to data sets where only a small number of members of each peer group is observed, even asymptotically.

In our application, the nonlinear peer effects model we estimate is derived from utility maximization. Consumers perceived needs for goods such as luxuries are affected by the mean expenditures on these goods of others in one's peer group. By identifying these perceived needs, we provide measures of the welfare (utility) costs to society of "keeping up with the Joneses." These costs are associated with the utility that is lost from feeling relatively poorer when our peers get richer. As a result, I may need to spend more just to get back to the same level of utility I had before my peers got richer. These costs have important implications for tax policy. In particular, we find that cutting income taxes, and thereby increasing incomes, produces fewer welfare gains than standard estimates suggest. This is because part of what I gain from my increased income is spent on keeping up with my peers, who's income is also increased by the tax cut.

Our model starts with existing structural utility function theory on identifying individual's perceived consumption needs from Samuelson (1947), Gorman (1976), Blackorby and Donaldson (1994), and Pendakur (2005). We then incorporate into this framework peer effects as in Manski (1993, 2000) and Brock and Durlauf (2001). However, our model differs from existing models both of consumer demand and of peer effects in a number of important ways. In particular, unlike almost all of the empirical social interactions and peer effects literature, due to the nature of the required utility theory and the empirical evidence of Engel curves, we require a nonlinear model of peer effects. We also require both different methods of obtaining identification, and different asymptotics from what is usually employed in the consumer demand or the social interactions and peer effects literatures.

For consumption of goods and services (particularly conspicuous ones like luxuries), ones relevant peer group does not consist of just immediate friends and associates, but virtually anybody of comparable status to oneself. Therefore, rather than relying on typical social network data, we consider ones peers to be essentially everybody of similar geographic, educational, and job status to oneself. However, this introduces a fundamental data problem, since we only observe a very small number of each person's peers. In our empirical application we employ household survey data from India. Dividing the data into sensibly defined peer groups results in only a few dozen people being observed in each group, with the vast majority of others in each group being unobserved.

Household surveys like ours are a standard source of consumer expenditure data at the individual level, but they pose many challenges for estimating peer effects. For example, we cannot exploit variation in group sizes to aid in identification, indeed, true group sizes are unknown. It is also not appropriate to treat observed average expenditures in each group as true group mean expenditures, even asymptotically (given the small number of actual observations per group), and so both identification and estimation need to account for the resulting mismeasurement in group means. Yet another issue is that many peer effects models cannot allow for either group level fixed effects or group level random effects, because such effects typically cannot be separated from the impacts of group mean variables. Note that this is distinct from Manski's (1993) reflection problem, which, along with the above mentioned nonlinearities, is yet another issue our model must address.

After reviewing the relevant literatures, the steps of our analysis are as follows. We first consider a simple, generic model containing peer effects, and show how it can be identified and estimated with our type of data. This new methodology for identifying and estimating a peer effects model should be of general interest, since it is potentially applicable to other contexts where one only observes a small number of members of each peer group. A key feature of this methodology is that it exploits both nonlinearity and structure to deal with the above listed identification issues. To account for the small number of individuals that are observed in each group, our asymptotic theory assumes that the number of groups goes to infinity, but the number of observed individuals within each group does not. Correlations between group level variables and errors, exacerbated by the same nonlinearities that otherwise help identification, introduce difficulties for estimation. Much of the novelty of our generic peer effects methodology comes from overcoming these correlations to construct valid moment conditions used for GMM type estimation.

After presenting our generic peer effect model and methodology, we then derive our specific model of consumer behavior. This begins with the existing theory of 'needs' in utility and demand function specification and associated welfare calculations. We adapt this class of utility derived demand models to our context where perceived needs can depend on the purchases of one's peers. We then arrive at a set of demand functions that, while more complicated than our generic peer effects model, can be identifed and estimated using the same techniques.

We then implement our estimator and associated welfare analyses using a few annual cross-sections of household-level expenditure data from India. Our data offer a good laboratory for this analysis because we observe relevant characteristics of each household for constructing peer groups, including education level, industry of employment, and detailed geographic area of residence, along with other household characteristics such as household size, age, religion, and caste. We find that peer effects are large and important for luxuries but not necessities. A one rupee increase in peer group luxury spending increases one's own perceived needs for luxury spending by about half a rupee. So a large fraction of luxury purchases are perceived as necessary. The estimates imply that, in terms of utility, the costs of keeping up with the Joneses accounted for about one seventh of Indian GDP growth from 1994 to 2010.

1.1 Relevant Literature - Peer Effects in Consumption, Income, and Demand

Income gains, particularly in the upper parts of the income distribution, may not increase well being much if utility depends on other people's consumption levels. See e.g., Frank (1999, 2012). The possible mechanisms for this are varied: Veblen (1899) effects make consumers value consumption of visible status goods; reference-dependent utility functions make consumption valuable only inasmuch as it exceeds what we see around us; "keeping up with the Joneses" makes our effective consumption smaller the more our peers consume; the consumption of our peers affects what we perceive as 'necessary'; etc. What these stories have in common is that they imply that each individual's consumption may have externalities on the utility functions of others around them.

We bring this intuition to data using a model of consumption externalities in which the welfare cost of such externalities is easily expressed. We exploit the intuition that the consumption of a person's peers affects their own perception of their needs. In the context of utility and cost functions, 'needs' are fixed costs, representing the minimum quantity vector one requires to start getting any utility. The idea that preferences have fixed costs that need to be met before expenditures start increasing utility is an old one, going back at least to Samuelson's (1947) note on the implications of linearity. Samuelson called the minimum quantity vector corresponding to needs the "necessary set" of goods, and defined "supernumerary income" as one's remaining income, after subtracting off the 'fixed cost' of these needs. Utility is then obtained by spending supernumerary income. The classical Stone (1954) and Geary (1949) Linear Expenditure System incorporates this construction. More generally, Gorman (1976) showed that these kind of fixed costs (which he calls "overheads") can be introduced into any utility function and will generally vary across consumers. This type of model has implications for the specification of demand functions that take the form of shape invariance in quantity demands (see Pendakur 2005; Pollak and Wales 1992 refer to this as ‘demographic translation’). This class of specifications is analogous to the more well known shape invariance in budget shares popularized by, e.g., Pendakur (1999) and Blundell, Chen and Kristensen (2007). See Pendakur (1999) and Lewbel (2010) for details.

Blackorby and Donaldson (1994) show how valid social welfare functions can be constructed based on differences between budgets and needs, using what they call "Absolute Equivalence Scale Exactness" or AESE. Blackorby and Donaldson (1994) also show that needs themselves are not identified from observable consumer behaviour, but differences in needs across consumers with different values of exogenously varying conditioning variables are identifiable from behaviour, and that is all that we require for our welfare analyses. Identification and welfare calculations are more complicated in our model, because in our context needs depend on peer group expenditures which are endogenous rather than exogeneous. Nevertheless we find we can adapt their framework to our setting.

Our model starts with Gorman’s specification of overheads or needs as a quantity vector. One then derives utility from the vector of differences between the vector of goods purchased and this minimum needs quantity vector. Our innovation is that we specify this vector of needs as a function of the consumption of one’s peers. Thus, consumption externalities arise because one’s perceived needs depend on the consumption of one’s peers. If my peers consume more, my perceived needs and hence my fixed costs go up. This in turn makes my supernumerary income go down, resulting in both an observable change in expenditure patterns and a loss in utility. The model therefore has testable implications, and it implies welfare losses that are quantifiable in both an individual and a social sense. Specifically, we can simply add up the estimated increases in needs across people, which then provides the dollar social cost of keeping up with the Joneses.

Note that these costs of keeping up with peers are not necessarily all wasted resources. For example, there may be a value to society if everyone has internet access, even if that makes internet access become a perceived need. We do not take a stand on what fraction of the costs of peer effects is not a waste from the standpoint of society, though the peer effects of luxuries are unlikely to be useful for the most part. Regardless, it is clearly useful to quantify these costs of peer effects, and separate those from the direct impacts of income increases on consumer’s utility.

Our model fits into the large literature on income-evaluation and income reference points, where one’s valuation of income depends on one’s reference group. Surveys of this literature include Kahneman (1992) and Clark, Frijters, and Shields (2008). A smaller literature focuses on the difference between consumption and income, allowing the valuation of one’s consumption to depend one’s reference group. These are are mostly intertemporal models that are intended to address macroeconomic puzzles. See, e.g., Gali (1994) or Maurer and

Meier (2008). At the other extreme, some papers in psychology and marketing focus on how the valuation of particular individual goods or brands depend on one's peers. See, e.g., Rabin (1998) and Kalyanaram and Winer (1998). Chao and Schor (1998) find that these effects are particularly important for goods that are visible (in their case, cosmetics), linking their findings to Veblen effects.

Many analyses of peer effects of these types are essentially nonstructural, including the above Chao and Schor (1998) paper. Boneva (2013) regresses household quantity demand vectors on household budgets (total expenditures) and on the average budgets of reference groups, using PROGRESA/Opportunidades related variables to instrument group averages. Ravina (2008) and Clark and Senik (2010) regress self-reported utility on own budgets and reference group average budgets. Virtually all such studies find these group average variables to be significant, though Ravaillon and Lokshin (2010) say that it is not important for the poor. But the magnitude or statistical significance of coefficients on group variables do not measure welfare effects. This requires a model of utility with sufficient structure to permit social welfare analysis.

1.2 Relevant Literature - Identification

Our model where each individual's outcome depends on the mean of the outcomes of one's peer group is a form of social interactions model. It can also be seen as a spatial model, where all individuals within a group are equidistant from each other.

A well known obstacle to identification of this kind of model is the reflection problem described by Manski (1993, 2000). See also Brock and Durlauf (2001), and Blume, Brock, Durlauf, and Ioannides (2010). Our model will have specific behaviorally derived structure (in particular some nonlinearity) that overcomes reflection problems. In some peer effects model, network information is available and can help identification. For example, Bramoullé, Djebbari, and Fortin (2009) show identification of peer effects in social networks that are sufficiently interconnected, and where for each member of group g, the peer effect is linear in a mean taken over all other group members. These types of models exploit variation in group sizes to aid identification, and requires that the number of observed members of each group increases with sample size (See, e.g., Devezies et. al, 2006). Using a somewhat related approach, Graham (2008) estimates peer effects in a linear model by comparing the covariance of test scores in large versus small classrooms. In our context, variation in group size does not provide any identifying power, both because we only see a small number of members of each group, and because we do not know actual group sizes.

The interactions of peer group members may be modeled as a game. Suppose there is private information that cannot observed by econometricians. We assume that group members have utility functions that depend on peers only through the true mean of the peer group's outcomes. If group members also all observe each other's private information and make decisions simultaneously (corresponding to a complete information game), then each individual's actual behavior will only depend on others through the group mean. Complete games are generally plausible only when the size of each group is small, and are typically estimated assuming the econometrician's data includes all members of each observed group. An example is Lee (2007). However, in our case the true group sizes are large, but we only observe a small number of members of each group. An alternative model of group behaviour is a Bayes equilibrium derived from a game of incomplete information, in which each individual has private information and makes decisions based on rational expectations regarding others. This type of incomplete game of group interactions can result in the reflection problem again, where endogenous effects, exogeneous effects, and the correlated effects cannot in general be separately identified. In either type of game there is also the potential problem of no equilibrium or multiple equilibria existing, resulting in the problems of incompleteness or incoherence and the associated difficulties they introduce for identification as discussed by Tamer (2003).

We do not take a stand on whether the true game in our case is one of complete or incomplete information. We just assume that players are basing their behavior on the true group means. This assumption is most easily rationalized by assuming that consumers either have complete information, or can observe a sufficiently large number of members in each group that their errors in calculating group means are negligible. A more difficult problem would be allowing for the possibility that each group member also only observes group means with error. We do not attempt to tackle that issue in this paper. In that case we would need to model how individuals estimate group means, how they incorporate uncertainty regarding group means into their purchasing decisions, and how all of that could be identified in the presence of all of the other obstacles to identification that we face. These obstacles include only observing a small number of members of each group, the reflection problem, group level fixed effects, nonlinearities resulting from utility maximization, and a multiple equation system where each equation depends on the vector of peer means from all of the equations.

Identification depends on what we assume is observable from data. Standard models of within group interactions with large groups assume that there are no interactions between groups, and that both the number of groups G and the number of observed members n_g within each group goes to infinity. However, for reasonable definitions of peer groups, standard consumer expenditure surveys only sample a relatively small number of individuals within each group (even in our relatively large Indian data set, n_g is less than one or two dozen for many groups). So while it is reasonable to assume that G goes to infinity, we take a completely new approach to identifying and estimating peer effects, by assuming that n_g is small and fixed. This means that observed within group sample averages are mismeasured estimates of true within group means, and that these measurement errors do not disappear asymptotically as the sample size grows with G. Moreover, these measurement errors are by construction correlated with individual specific covariates, further exacerbating the difficulties listed earlier in obtaining identification and constructing consistent estimators of model parameters.

Measurement error more broadly has long been recognized as potentially important in social interactions models (e.g., Moffitt 2001 and Angrist, 2014), though this work focuses on standard issues of mismeasurement in regressors, recognizing that, unlike in ordinary models, outcomes are also regressors and hence measurement error in outcomes matters. This is quite different from our situation, which recognizes that only observing a limited number of individuals in each group results in measurement errors in group means. This can also be interpreted as a missing data problem where what is missing is the outcomes of most group members. Others have looked at different missing data problems in peer models. For example Sojourner (2009) considers peer effects in Project STAR classrooms, where the missing data consists of pre-intervention information on student achievement. In his model, the difficulties of missing data are addressed in part by assuming a linear model where student are randomly assigned to their peer groups, defined as classrooms.

As is standard in models with measurement errors, we will assume we have valid instruments that are correlated with true group means. However, even with instruments, the obvious two stage least squares or GMM estimator that assumes model errors are uncorrelated with instruments (after replacing true group means with their sample analogs) will not be consistent in our context. This is because: a) in a linear model such instruments will not overcome the reflection problem; and b) in a nonlinear model we will have interaction terms between the measurement errors and the true regressors. An analogous problem arises in the polynomial model with measurement errors considered by Hausman, Newey, Ichimura, and Powell (1991). We show that overcoming these issues requires some novel transformations that ultimately lead to a valid GMM estimator.

Finally, even given complete identification of model parameters, the Blackorby and Donaldson (1994) result discussed earlier still applies, namely, that only relative needs across consumers are identified, not the absolute level of needs. However, this will suffice for all of our welfare analyses.

2 Generic Model Identification

Before introducing our general model of peer effects in consumer demand, in this section we consider a simple generic model where individual outcomes depend on group means. We use this model to illustrate the difficulties associated with identification in our general context, and to show how we overcome these difficulties, and how we construct a corresponding estimator. This generic model should be useful in other applications where peer effects are nonlinear, and the models require fixed effects or random effects.

Here we summarize the main structure of our generic social interactions model, and the associated logic of its identification and estimation. In the Appendix we provide detailed assumptions regarding the model and a formal proof of its identification. Let i index individuals. Each individual i is in a peer group $g \in \{1, ...G\}$. The number of peer groups G is large, so we assume $G \to \infty$. In our data we will only observe a small number n_g of the individuals in each peer group g. So asymptotics assuming $n_g \to \infty$ would be a poor approximation for our data. We therefore assume n_g is fixed and so does not grow with the sample size.

Let y_i be an outcome which is affected by an observed scalar regressor x_i (We later generalize the model to allow y and x to be vectors of outcomes and of regressors). Denote the group mean outcome $\overline{y}_g = E(y_i \mid i \in g)$, and similarly define $\overline{x}_g$. The general form of our model is

$$y_i = h\left(\theta \mid \overline{y}_g, x_i\right) + v_g + u_i \tag{1}$$

where v_g for $g \in \{1, ...G\}$ are group level random or fixed effects, u_i are mean zero errors, independent of $x_{i'}$ for all individuals i', and θ is a vector of parameters to be identified and estimated. The dependence of h on $\overline{y}_g$ are peer effects we want to identify. Note that $\overline{x}_g$ does not appear explicitly in this model, however, we have allowed for a fixed effect v_g, which could be an unknown function of both $\overline{x}_g$ and of any other group level covariates. Although excluding $\overline{x}_g$ would solve the reflection problem in a model without v_g, the problem is not avoided by excluding $\overline{x}_g$ in our model.

Suppose h were linear, i.e., suppose $h\left(\theta \mid \overline{y}_g, x_i\right)$ equalled $\overline{y}_g a + x_i b$. A constant term is omitted here because it would trivially be included in v_g. Then the peer effect, given by the parameter a, could not be identified because we could not separate $\overline{y}_g$ from v_g. To overcome this linear model nonidentification (and because there is substantial empirical evidence of nonlinearity in our empirical application), we propose the nonlinear model[1]

$$h\left(\theta \mid \widehat{y}_g, x_i\right) = \left(\overline{y}_g a + x_i b\right)^2 d + \left(\overline{y}_g a + x_i b\right) \tag{2}$$

where $\theta = (a, b, d)$.

Now $\overline{y}_g$ cannot actually be observed (even asymptotically, because we have assumed n_g is fixed), so we will need to replace it with some estimator. Let $\widehat{y}_g$ be an estimator of $\overline{y}_g$. This introduces an additional error term ε_{gi} defined by $\varepsilon_{gi} = h\left(\theta \mid \overline{y}_g, x_i\right) - h\left(\theta \mid \widehat{y}_g, x_i\right)$, and the model becomes

$$y_i = \left(\widehat{y}_g a + x_i b\right)^2 d + \left(\widehat{y}_g a + x_i b\right) + v_g + u_i + \varepsilon_{gi}$$

where

$$\varepsilon_{gi} = \left(\overline{y}_g - \widehat{y}_g\right) a + \left(\overline{y}_g^2 - \widehat{y}_g^2\right) a^2 d + 2\left(\overline{y}_g - \widehat{y}_g\right) x_i bad$$

Inspection of this equation shows a number of obstacles to identifying and estimating θ. First, v_g will in general be correlated with $\overline{y}_g$ and hence with $\widehat{y}_g$ (this was the main cause of nonidentification in the linear model). Second, since n_g does not go to infinity, if $\widehat{y}_g$ contains y_i, then $\widehat{y}_g$ will correlate with u_i. Third, again because n_g is fixed, ε_{gi} doesn't vanish asymptotically, and is by construction correlated with some functions of $\widehat{y}_g$ and x_i. Equivalently, we should think of $\left(\overline{y}_g - \widehat{y}_g\right)$ and $\left(\overline{y}_g^2 - \widehat{y}_g^2\right)$ as measurement errors in $\overline{y}_g$ and $\overline{y}_g^2$, leading to the standard measurement error problem that mismeasured regressors are correlated with errors in the model.

So, while nonlinearity overcomes the fundamental nonidentification of the linear model, it introduces a host of other obstacles to identification that we need to overcome. We emply two somewhat different methods for identifying the model, depending on whether each v_g is assumed to be a fixed effect or a random effect. For each case, we construct a set of moment conditions that suffice to identify θ, and can be used for estimated via GMM (Generalized Method of Moments, see Hansen 1982).

2.1 Generic Model Identification - Fixed Effects

We begin by looking at the difference between the outcomes of two people i and i' in group g.

$$y_i - y_{i'} = h\left(\theta \mid \overline{y}_g, x_i\right) - h\left(\theta \mid \overline{y}_g, x_{i'}\right) + u_i - u_{i'}$$

[1]We show in the appendix that the seemingly more general model $y_i = \left(\overline{y}_g a + x_i b + c\right)^2 d + \left(\overline{y}_g a + x_i b + c\right) + v_g + u_i$ is observationally equivalent to the simpler form given above.

This differencing removes the fixed effects v_g. This also differences out the quadratic term $\overline{y}_g^2 a^2$ inside h. Define the leave-two-out group mean estimator

$$\widehat{y}_{g,-ii'} = \frac{1}{n_g - 2} \sum_{l \in g, l \neq i, i'} y_l$$

This is just the sample average of y for everyone who is observed in group g except for the individuals i and i'. Let $\widehat{y}_g$ from before be the estimator $\widehat{y}_{g,-ii'}$. Then

$$y_i - y_{i'} = h\left(\theta \mid \widehat{y}_{g,-ii'}, x_i\right) - h\left(\theta \mid \widehat{y}_{g,-ii'}, x_{i'}\right) + u_i - u_{i'} + \varepsilon_{gi} - \varepsilon_{gi'}. \tag{3}$$

We can then show (see Theorem 1 in the Appendix) that, with these definitions,

$$E\left(u_i - u_{i'} + \varepsilon_{gi} - \varepsilon_{gi'} \mid x_i, x_{i'}\right) = 0 \tag{4}$$

which we can then use to construct moments for estimation of equation (3).

The intuition for this result can be seen by reexamining the obstacles to identification listed earlier. The correlation of v_g with $\overline{y}_g$ and hence with $\widehat{y}_{g,-ii'}$ doesn't matter because v_g has been differenced out. $\widehat{y}_{g,-ii'}$ does not correlate with u_i or $u_{i'}$ because individuals i and i' are omitted from the construction of $\widehat{y}_{g,-ii'}$. Finally, we can verify that $\varepsilon_{gi} - \varepsilon_{gi'}$ is linear in $x_i - x_{i'}$, with a conditionally mean zero coefficient.

Equation (3) contains functions of $\widehat{y}_{g,-ii'}$, x_i, and $x_{i'}$ as regressors, and equation (4) shows that we can use functions of x_i and $x_{i'}$ as instruments (equivalently, x_i and $x_{i'}$ are exogenous regressors). But what can we use as an instrument for $\widehat{y}_{g,-ii'}$? An obvious candidate instrument would be some estimate $\widehat{x}_g$ of $\overline{x}_g$, the reason being that y_i depends on x_i and therefore the average within group value of y should be correlated with the average within group value of x. The problem is that, although $E\left(\varepsilon_{gi} - \varepsilon_{gi'} \mid x_i, x_{i'}\right) = 0$, the error $\varepsilon_{gi} - \varepsilon_{gi'}$ will in general be correlated with x_l for all observed individuals l in the group other than the individuals i and i'. Note that this problem is due to the assumption that n_g is fixed. If it were the case that $n_g \to \infty$, then $\varepsilon_{gi} - \varepsilon_{gi'} \to 0$, and this problem would disappear.

To overcome this final obstacle to identification in the fixed effects model (finding an instrument for $\widehat{y}_{g,-ii'}$), we require some source of group level data. For example, in our application x_i is total consumption expenditures. A valid instrument for $\widehat{y}_{g,-ii'}$ would then be something that correlates with $\overline{x}_g$ e.g., some measure of the average level of income, wealth or socioeconomic status of the group, perhaps obtained from census data.

An alternative source of group level instruments is what we actually use in our empirical application. Our data set, which is typical of consumption surveys, is repeated cross section data, where different consumers are sampled in each time period. Now $\varepsilon_{gi} - \varepsilon_{gi'}$ is correlated with x_l for individuals l in group g that we observed and used in constructing $\widehat{y}_{g,-ii'}$. But $\varepsilon_{gi} - \varepsilon_{gi'}$ will not in general be correlated with other individuals, and in particular will not be correlated with individuals that are observed in group g in other time periods (again, see the appendix for details). We can therefore construct an instrument that correlates with $\overline{x}_g$ by taking the sample average of x_l for individuals l who are observed in group g in *other* time periods. These will be useful and valid instruments as long as group level total expenditures $\overline{x}_g$ are autocorrelated over time.

Let $\mathbf{r}_g$ denote a vector of valid group level instruments, constructed as above either from other datasets or from other time periods. Combining these with equations (3) and (4) then gives conditional moments

$$E\left[y_i - y_{i'} - h\left(\theta \mid \widehat{y}_{g,-ii'}, x_i\right) + h\left(\theta \mid \widehat{y}_{g,-ii'}, x_{i'}\right) \mid x_i, x_{i'}, \mathbf{r}_g\right] = 0$$

Since it is easier to estimate models using unconditional moments, let $\mathbf{r}_{gii'}$ denote a vector of functions of $x_i, x_{i'}, \mathbf{r}_g$. Since h is quadratic, a natural choice of elements comprising $\mathbf{r}_{gii'}$ would be $x_i, x_{i'}, \mathbf{r}_g$, and squares and cross products of these variables. We then have the unconditional moments

$$E\left[\left(y_i - y_{i'} - h\left(\theta \mid \widehat{y}_{g,-ii'}, x_i\right) + h\left(\theta \mid \widehat{y}_{g,-ii'}, x_{i'}\right)\right) \mathbf{r}_{gii'}\right] = 0. \tag{5}$$

Theorem 1 in the Appendix extends this model to a vector $\mathbf{x}_i$, and proves that the parameters θ are identified from these unconditional moments.

After plugging equation (2) for the function h into equation (5), we obtain an expression that can immediately be used for estimation by GMM. For estimation, observations are defined as every pair of individuals i and i' in each group. By construction, the errors in this model are correlated across observations within each group. It is therefore necessary to estimate the model using clustered standard errors, where each group is a cluster (again, details are provided in the Appendix).

2.2 Generic Model - Random Effects

A drawback of the fixed effects model is that differencing across individuals, which was needed to remove the fixed effects, results in a substantial loss of information. So in this section we instead assume that v_g is independent of x_i (a random effects assumption) and provide addtional moments that do not entail differencing. The moments obtained under fixed effects remain valid under the additional random effects asumptions. So the proof of identification under fixed effects (Theorem 1) also shows identification of the random effects model. The goal here is to show how additional moments (that do not require differencing) can be obtained exploiting the random effects independence of v_g from x_i. We may then do GMM using both the fixed effects moments from before with the additional random effects moments obtained here.

For random effects it will be convenient to rewrite the quadratic model, equations (1) and (2), as

$$y_i = \overline{y}_g^2 a^2 d + (a + 2x_i abd)\, \overline{y}_g + \left(x_i b + x_i^2 b^2 d\right) + v_g + u_i \tag{6}$$

As before, we will need to replace the unobserved $\overline{y}_g$ with some estimate, and this replacement will add an additional epsilon term to the errors. However, in the fixed effects case, when we pairwise differenced this model, the quadratic term $\overline{y}_g^2$ also dropped out. Now, since we will not be differencing, we will need to cope not just with estimation error in $\overline{y}_g$, but also in $\overline{y}_g^2$ (recall also that since n_g is fixed, this estimation error is equivalent to measurement error, which does not disappear asymptotically). To obtain valid moment conditions, we employ a variant of the trick we used before. Again let i' denote an individual other than i in group g, and $\widehat{y}_{g,-ii'}$. Suppose we replaced $\overline{y}_g$ with $\widehat{y}_{g,-ii'}$ as before. The problem now is that the error $\widehat{y}_{g,-ii'}^2 - \overline{y}_g^2$ would in general be correlated with x_l for every individual l in the group, including i and i'.

To circumvent this problem, we replace the linear term $\overline{y}_g$ with the estimate $\widehat{y}_{g,-ii'}$ as before, but we replace the squared term $\widehat{y}_{g,-ii'}^2$ with $\widehat{y}_{g,-ii'} y_{i'}$. This latter replacement might seem problematic, since a single individual's $y_{i'}$ provides a very crude estimate of $\overline{y}_g$. However, we repeat this construction for every individual i' (other than i) in the group, and essentially average the resulting moments over all individuals i' in g. With this replacement, equation (6) becomes

$$y_i = \widehat{y}_{g,-ii'} y_{i'} a^2 d + (a + 2x_i abd)\, \widehat{y}_{g,-ii'} + \left(x_i b + x_i^2 b^2 d\right) + v_g + u_i + \widetilde{\varepsilon}_{gii'}$$

where

$$\widetilde{\varepsilon}_{gii'} = \left(\overline{y}_g^2 - \widehat{y}_{g,-ii'} y_{i'}\right) a^2 d + (a + 2x_i abd)\left(\overline{y}_g - \widehat{y}_{g,-ii'}\right)$$

We can then show (see the Appendix for details), that $E(\widetilde{\varepsilon}_{gii'} | x_i, \mathbf{r}_g) = -da^2 Var\,(v_g)$ and hence,

$$E\left[y_i - \widehat{y}_{g,-ii'} y_{i'} a^2 d - (a + 2x_i abd)\, \widehat{y}_{g,-ii'} - \left(x_i b + x_i^2 b^2 d\right) - v_0 \mid x_i, \mathbf{r}_g\right] = 0 \tag{7}$$

where $v_0 = E\,(v_g) - da^2 Var\,(v_g)$ is a constant to be estimated along with the other parameters, and $\mathbf{r}_g$ are the same group level instruments we defined earlier. Letting $\mathbf{r}_{gi}$ be functions of x_i and $\mathbf{r}_g$ (such as x_i, $\mathbf{r}_g$, x_i^2, and $x_i\mathbf{r}_g$), we immediately obtain unconditional moments

$$E\left[\left(y_i - \widehat{y}_{g,-ii'} y_{i'} a^2 d - (a + 2x_i abd)\,\widehat{y}_{g,-ii'} - \left(x_i b + x_i^2 b^2 d\right) - v_0\right)\mathbf{r}_{gi}\right] = 0 \tag{8}$$

which we can estimate using GMM exactly as before. The moments from the fixed effects model, equation (5), remain valid under random effects, so both equations (5) and (8) could be combined in a single GMM estimator to increase asymptotic efficiency.

3 Utility, Welfare, and Demands With Needs Containing Peer Effects

Here we first summarize some existing demand theory regarding utility and welfare calculations in the presence of perceived needs. These needs take the form of fixed costs or overheads in the utility function. We then introduce peer effects into these perceived needs, and obtain an associated demand system that we later identify and estimate.

3.1 Utility and Welfare With Needs

Let i index consumers and let $\mathbf{q}_i = (q_{1i}, .., q_{Ji})$ be a J–vector of commodity quantities chosen by consumer or household i. Let $\mathbf{p}$ be the corresponding J-vector of prices of each commodity, and let x_i be the total budget for commodities of consumer i. Commodities here are aggregates of goods or services that are assumed to be purchased and consumed in continuous quantities. Each consumer i is assumed to choose $\mathbf{q}_i$ to maximize a direct utility function, subject to the budget constraint that $\mathbf{p}'\mathbf{q}_i \leq x_i$. Preferences and the utility function of a consumer i can be represented by an indirect utility function $V_i(\mathbf{p}, x_i)$ which gives the utility level attained by consumer i when facing the budget constraint defined by $(\mathbf{p}, x_i)$. The purchase choice $\mathbf{q}_i$ of consumer i in all different price and budget regimes describes their demand functions $\mathbf{q}_i = \mathbf{g}_i(\mathbf{p}, x_i)$. The demand functions $\mathbf{g}_i$ are related to V_i by Roy's identity. If consumer's preferences are stable, then demand functions $\mathbf{q}_i = \mathbf{g}_i(\mathbf{p}, x_i)$ may be observed from demand data, that is, by seeing what the consumer chooses to buy in every possible price and budget regime.

For welfare comparisons, we need to be able to compare well being across consumers. Define the *equivalent-income function* $X_i(\mathbf{p}, x)$ as the income (budget) needed by consumer i to get the same level of utility as that of some reference consumer $i = 0$ having a budget x. Blackorby and Donaldson (1994) define Absolute Equivalence Scale Exactness (AESE) as a property of utility functions that holds if and only if, for all consumers i, $X_i(\mathbf{p}, x) = x - \widetilde{F}_i(\mathbf{p})$

for some function $\widetilde{F}_i(\mathbf{p})$. They show that AESE holds if and only if $\widetilde{F}_i(\mathbf{p}) = F_i(\mathbf{p}) - F_0(\mathbf{p})$ where

$$V_i(\mathbf{p}, x_i) = V\left(\mathbf{p}, x_i - F_i(\mathbf{p})\right)$$

Here V is some indirect utility function (not the utility function V_0 of the reference household) and the function $F_i(\mathbf{p})$ for each consumer i can be interpreted as the cost of satisfying the perceived needs of consumer i. This model implies that each consumer derives utility (through a common utility function V^0) from the extent to which their available budget x_i exceeds their perceived cost of needs $F_i(\mathbf{p})$.

One property of AESE is that it involves both testable and untestable restrictions on utility. If follows from Roy's identity that quantity demand functions are given by:

$$\mathbf{g}_i(\mathbf{p}, x_i) = \mathbf{g}_0(\mathbf{p}, x_i - \widetilde{F}_i(\mathbf{p})) + \frac{\partial \widetilde{F}_i(\mathbf{p})}{\partial \mathbf{p}}.$$

This is shape-invariance in quantity demands as in Pendakur (2005). Here, demand functions are identical across any two consumers i and j, except that they are translated in x by $\widetilde{F}_i(\mathbf{p}) - \widetilde{F}_j(\mathbf{p})$ and translated in quantities by the vector $\frac{\partial \widetilde{F}_i(\mathbf{p})}{\partial \mathbf{p}} - \frac{\partial \widetilde{F}_j(\mathbf{p})}{\partial \mathbf{p}}$. Figure 1 shows a quantity demand equation for 2 consumers where this shape-invariance property is satisfied. Here, the quantity demand equations over expenditure of both (all) consumers have the same shape, but may be shifted vertically (in the quantity demand) or horizontally (in the budget x). This restriction must hold for all quantity demand equations.

Identification of the differences in needs is easy to see. Given exogenous variation in a variable that shifts $\widetilde{F}_i$, we can recover the response of $\widetilde{F}_i$ to that variation by estimating the horizontal shift in quantity demand curves for consumer i. The heavy lifting is in finding moment conditions that give us that exogenous variation.

Shape invariance in quantity demands is testable. Testing involves comparing the demand functions $\mathbf{g}_i(\mathbf{p}, x)$ of each consumer i with the demand functions $\mathbf{g}_0(\mathbf{p}, x)$ of a reference consumer, and seeing if a function $\widetilde{F}_i(\mathbf{p})$ exists for each consumer i that makes the above equation hold.

However, AESE is not fully testable. In particular, since quantity demands depend only on the ordinal properties of utility functions, we cannot test the cardinal restrictions imposed by AESE. That is, the exact same observable restrictions on demand (that is, shape-invariance) will hold if, for each consumer i, $V_i(\mathbf{p}, x_i) = H_i\left[V^0\left(\mathbf{p}, x_i - F_i(\mathbf{p})\right)\right]$ for some strictly monotonically increasing function H_i. The untestable restriction of AESE is that H_i equals the identity function.

Another property of AESE documented by Blackorby and Donaldson (1994) is that the cost of needs functions $F_i(\mathbf{p})$ cannot themselves be identified from demand data (except possibly by arbitrary functional form restrictions). All that can be identified is the differences in needs functions across consumers, that is, $\widetilde{F}_i(\mathbf{p}) = F_i(\mathbf{p}) - F_0(\mathbf{p})$ is identified but not $F_i(\mathbf{p})$ itself. Consequenty, we normalize $F_0(\mathbf{p}) = 0$.

Identification of $\widetilde{F}_i(\mathbf{p})$ is in the following sense. Suppose utility functions satisfy AESE. Then, preferences satisfy shape-invariance, and we can recover $\widetilde{F}_i(\mathbf{p})$ from demand behaviour. However, by applications of monotonic transformations H_i, we could generate an infinite number of other equivalent-income functions X_i, and all of these would be consistent with observed behaviour. Blackorby and Donaldson show that only one of these satisfies the additive property of the AESE equivalent-income function.

The usefulness of AESE for our purpose is that, since $X_i(\mathbf{p}, x) = x - \widetilde{F}_i(\mathbf{p})$, we can use AESE to directly measure, in dollar terms, the welfare impacts of changes in costs, which in our application will come from peer effects. It follows from the equivalent income function that, under AESE, we can define anonymous social welfare functions over $X_i(\mathbf{p}, x)$ instead over over utilities V_i. One such valid social welfare function is the simple sum $\sum_i x_i - \widetilde{F}_i(\mathbf{p})$ (it is not inequality-averse). Notice that this welfare function is easy to understand, and can be evaluated just from identification of the relative costs of needs functions $\widetilde{F}_i(\mathbf{p})$. If needs change, social welfare changes by the sum of those needs changes.

In our work we will assume that $F_i(\mathbf{p}) = \mathbf{p}'\mathbf{f}_i$. This is the most natural form of AESE model, where $\mathbf{f}_i$ is a quantity vector. Each element f_{ji} of $\mathbf{f}_i$ equals the quantity of commodity j that consumer i perceives as needs, that is, f_{ji} is the minimum amount household i feels must be consumed, before they can start to generate any utility. This implies the model

$$V_i(\mathbf{p}, x_i) = V\left(\mathbf{p}, x_i - \mathbf{p}'\mathbf{f}_i\right)$$

which by Roys identity has associated demand functions

$$\mathbf{g}_i(\mathbf{p}, x_i) = \mathbf{g}(\mathbf{p}, x_i - \mathbf{p}'\mathbf{f}_i) + \mathbf{f}_i \tag{9}$$

where $\mathbf{g}$ is the vector of demand functions that correspond to the utility function V. For linear $\mathbf{g}$, Samuelson (1947) calls the vector of needs $\mathbf{f}_i$ the necessary set, and calls $x_i - \mathbf{p}'\mathbf{f}_i$ supernumerary income. The Stone (1954) and Geary (1949) linear expenditure system is a special case of this model. Gorman (1976) and Pollak and Wales (1981) assume that $\mathbf{f}_i$ is a function of observable household characteristics (preference shifters) and call $\mathbf{f}_i$ overheads, interpreting $V\left(\mathbf{p}, x_i - \mathbf{p}'\mathbf{f}_i\right)$ as a consumer's production function where the ultimate product being produced is utility. They consider linear and quadratic specifications for $\mathbf{g}$. Pendakur (2005) showed the semiparametric characterisation of this model, and showed how identification works with unspecified form for $\mathbf{g}$.

3.2 Needs With Peer Effects

We modify the existing literature on needs, as summarized in the previous subsection, by specifying utility functions in which the needs vector $\mathbf{f}_i$ of a consumer i depends on quantities purchased by that consumer's peers.

Let $i \in g$ denote that consumer i belongs to group g. Let $\overline{\mathbf{q}}_g = E(\mathbf{q}_i \mid i \in g)$, so $\overline{\mathbf{q}}_g$ is the mean level of quantities consumed by consumers in group g. We specify $\mathbf{f}_i$ as a function of peer demands, specifically, we let

$$\mathbf{f}_i = \mathbf{A}\overline{\mathbf{q}}_g + \mathbf{C}\mathbf{z}_i$$

for some J by J matrix $\mathbf{A}$ and some J by K matrix $\mathbf{C}$, where g is the peer group of consumer i and $\mathbf{z}_i$ is a K vector of observed characteristics of consumer i. The larger the elements of $\mathbf{A}$ are, the greater are the peer effects. If $\mathbf{A}$ is a diagonal matrix, then the perceived needs for any commodity depend only on the group mean purchases of that commodity. We more generally allow for nonzero off diagonal elements as well. So, e.g., my peer's expenditures on luxuries could affect not only my perceived needs for luxuries, but also my perceived needs for necessities. In general, we elements of $\mathbf{A}$, particularly diagonal elements, to be nonnegative. However, they cannot be too large (and in particular diagonal elements cannot exceed one), since otherwise stable equilibria may not exist (analogous to the Assumption A2 inequality being violated in the generic model). See the Appendix for details.

Having specified $\mathbf{f}_i$ and hence the functions defining needs, now consider the indirect utility function V. A long empirical literature on commodity demands finds that observed demand functions are close to polynomial, and have a property known as rank equal to three. See, e.g. Lewbel (1991) and Banks, Blundell, and Lewbel (1997), and references therein. Gorman (1981) showed that any polynomial demand system will have a maximum rank of three, and Lewbel (1989) shows that the simplest tractible class of indirect utility functions that yields rank three polynomials in x is $V(\mathbf{p}, x) = (x - F(\mathbf{p}))^{1-\lambda} B(\mathbf{p}) / (1-\lambda) - D(\mathbf{p})$ for some constant λ and some differentiable functions F, B and D. Combining the shape invariant AESE model of the previous section with this specification of utility gives the model

$$V_i(\mathbf{p}, x) = \left(x_i - \mathbf{p}'\mathbf{A}\overline{\mathbf{q}}_g - \mathbf{p}'\mathbf{C}\mathbf{z}_i\right)^{1-\lambda} B(\mathbf{p}) / (1-\lambda) - D(\mathbf{p})$$

Preserving homogeneity (i.e., the absence of money illusion, which is a necessary condition for rationality of preferences), requires $B(\mathbf{p})^{1/(\lambda-1)}$ to be homogeneous of degree one in $\mathbf{p}$ and $D(\mathbf{p})$ to be homogeneous of degree zero in $\mathbf{p}$.

Applying Roys identity to this indirect utility function then yields the demand functions

$$\mathbf{q}_i = \left(x_i - \mathbf{p}'(\mathbf{A}\overline{\mathbf{q}}_g + \mathbf{C}\mathbf{z}_i)\right)^{\lambda} \frac{\nabla D(\mathbf{p})}{B(\mathbf{p})} + \left(x_i - \mathbf{p}'(\mathbf{A}\overline{\mathbf{q}}_g + \mathbf{C}\mathbf{z}_i)\right) \frac{\nabla B(\mathbf{p})}{(\lambda - 1) B(\mathbf{p})} + \mathbf{A}\overline{\mathbf{q}}_g + \mathbf{C}\mathbf{z}_i.$$

The most commonly assumed rank three models are quadratic (see the above references, and Pollak and Wales 1980), which corresponds to $\lambda = 2$. Convenient specifications of the price functions are $\ln B(\mathbf{p}) = \mathbf{b}' \ln \mathbf{p}$ with $\mathbf{b}'\mathbf{1} = \lambda - 1$ and $D(\mathbf{p}) = \mathbf{d}' \ln \mathbf{p}$ with $\mathbf{d}'\mathbf{1} = 0$, which yields the model

$$\mathbf{q}_i = \left(x_i - \mathbf{p}'(\mathbf{A}\overline{\mathbf{q}}_g + \mathbf{C}\mathbf{z}_i)\right)^2 \left(e^{-\mathbf{b}' \ln \mathbf{p}}\right) \mathbf{d}/\mathbf{p} + \left(x_i - \mathbf{p}'(\mathbf{A}\overline{\mathbf{q}}_g + \mathbf{C}\mathbf{z}_i)\right) \mathbf{b}/\mathbf{p} + \mathbf{A}\overline{\mathbf{q}}_g + \mathbf{C}\mathbf{z}_i, \quad (10)$$

where $\mathbf{d}/\mathbf{p}$ and $\mathbf{b}/\mathbf{p}$ are vectors with entries d_j/p_j and b_j/p_j for $j = 1, ..., J$. The goal will be estimation of the parameters $\mathbf{A}$, $\mathbf{C}$, $\mathbf{d}$, and $\mathbf{b}$, and our welfare analyses will be based on estimates of the system of equations (10).

To allow for unobserved heterogeneity in behavior, we append the error term $\mathbf{v}_g + \mathbf{u}_i$ to the above set of demand functions, where $\mathbf{v}_g$ is a $J-$vector of group level fixed effects and $\mathbf{u}_i$ is a $J-$vector of individual specific error terms that are assumed to have zero means conditional on all x_l, $\mathbf{z}_l$, and $\mathbf{p}$ with $l \in g$. The group level fixed effect $\mathbf{v}_g$ is not assumed to be independent of peer effects $\overline{\mathbf{q}}_g$. These error terms and fixed effects can be interpreted either as departures from utility maximization by individuals, or as unobserved preference heterogeneity. Assuming that the price weighted sum $\mathbf{p}'(\mathbf{v}_g + \mathbf{u}_i)$ is zero suffices to keep each individual on their budget constraint. Under this restriction, if desired one could replace $\mathbf{C}\mathbf{z}_i$ with $(\mathbf{C}\mathbf{z}_i + \mathbf{v}_g + \mathbf{u}_i)$ in the indirect utility function above, and treat error terms as unobserved preference heterogeneity. With this substitution into equation (10), the system of demand functions we have to identify and estimate are

$$\mathbf{q}_i = \left(x_i - \mathbf{p}'(\mathbf{A}\overline{\mathbf{q}}_g + \mathbf{C}\mathbf{z}_i)\right)^2 (e^{-\mathbf{b}' \ln \mathbf{p}}) \mathbf{d}/\mathbf{p} + \left(x_i - \mathbf{p}'(\mathbf{A}\overline{\mathbf{q}}_g + \mathbf{C}\mathbf{z}_i)\right) \mathbf{b}/\mathbf{p} + \mathbf{A}\overline{\mathbf{q}}_g + \mathbf{C}\mathbf{z}_i + \mathbf{v}_g + \mathbf{u}_i. \quad (11)$$

Proof that the parameters $\mathbf{A}$, $\mathbf{C}$, $\mathbf{d}$, and $\mathbf{b}$ are identified from the system of equations (11), analogous to the generic model, is shown as Theorem 2 in the Appendix. For ease of exposition, the specific assumptions needed to prove identification, analogous to Assumptions A1 to A5 in the generic model, are also deferred to the Appendix.

4 Implementing the Demand System

Here we first outline how the parameters in the system of demand equations (11) can be identified and estimated. As in the generic model, we show identification with fixed effects that uses pairwise differencing, we construct a corresponding GMM estimator, and we then show how additional moments to increase efficiency can be constructed by making additional random effects assumptions regarding $\mathbf{v}_g$.

4.1 Demand System Identification With Fixed Effects

As with the generic model, there are two main obstacles to identifying equation (11). First is that the fixed effects $\mathbf{v}_g$ correlate with all the regressors. Second is that $\overline{\mathbf{q}}_g$ is not observed. Let n_g denote the number of consumers we observe in group g. Assume $n_g \geq 3$. The actual number of consumers in each group may be large, but we assume only a small, fixed number of them are observed. Our asymptotics assume that the number of observed groups goes to infinity as the sample size grows, but for each group g, the number of observed consumers n_g is fixed. We may estimate $\overline{\mathbf{q}}_g$ by a sample average of $\mathbf{q}_i$ across observed consumers in group i, but the error in any such average is like measurement error, that does not shrink as our sample size grows.

Here we summarize how the parameters of the demand system (11) are identified. Formal assumptions and the proof of identification are provided in the Appendix. Identification of our demand model is proven in two steps. First, just consider data from a single time period, so there is no price variation and $\mathbf{p}$ can be treated as a vector of constants. Let $\boldsymbol{\alpha} = \mathbf{A}'\mathbf{p}$, $\boldsymbol{\gamma} = \mathbf{C}'\mathbf{p}$, $\boldsymbol{\delta} = \mathbf{b}/\mathbf{p}$, and $\mathbf{m} = \left(e^{-\mathbf{b}'\ln \mathbf{p}}\right)\mathbf{d}/\mathbf{p}$ with constraints of $\mathbf{b}'\mathbf{1} = 1$ and $\mathbf{d}'\mathbf{1} = 0$. Then equation (11) reduces to the system of Engel curves

$$\mathbf{q}_i = \left(x_i - \boldsymbol{\alpha}'\overline{\mathbf{q}}_g - \boldsymbol{\gamma}'\mathbf{z}_i\right)^2 \mathbf{m} + \left(x_i - \boldsymbol{\alpha}'\overline{\mathbf{q}}_g - \boldsymbol{\gamma}'\mathbf{z}_i\right)\boldsymbol{\delta} + \mathbf{A}\overline{\mathbf{q}}_\mathbf{g} + \mathbf{C}\mathbf{z}_i + \mathbf{v}_g + \mathbf{u}_i, \qquad (12)$$

This has a very similar structure to the generic multiple equation system of equations (28), and by a similar derivation we first show that $\boldsymbol{\alpha}$, $\mathbf{C}$, $\boldsymbol{\delta}$, and $\mathbf{m}$ are identified, as follows.

Define $\widetilde{\mathbf{v}}_g = \left(\boldsymbol{\alpha}'\overline{\mathbf{q}}_g\right)^2 \mathbf{m} - \boldsymbol{\alpha}'\overline{\mathbf{q}}_g\boldsymbol{\delta} + \mathbf{A}\overline{\mathbf{q}}_\mathbf{g} + \mathbf{v}_g$. Then equation (12) can be rewritten more simply as

$$\mathbf{q}_i = \left(x_i - \boldsymbol{\gamma}'\mathbf{z}_i\right)^2 \mathbf{m} - 2\left(x_i - \boldsymbol{\gamma}'\mathbf{z}_i\right)\left(\boldsymbol{\alpha}'\overline{\mathbf{q}}_g\right)\mathbf{m} + \left(x_i - \boldsymbol{\gamma}'\mathbf{z}_i\right)\boldsymbol{\delta} + \mathbf{C}\mathbf{z}_i + \widetilde{\mathbf{v}}_g + \mathbf{u}_i, \qquad (13)$$

Here the fixed effect $\mathbf{v}_g$ has been replaced by a new fixed effect $\widetilde{\mathbf{v}}_g$. As in the generic fixed effects model, we begin by taking the difference $q_{ji} - q_{ji'}$ for each good $j \in \{1, ..., J\}$ and each pair of individuals i and i' in group g. This pairwise differencing of equation (13) gives, for each good j,

$$\begin{aligned} q_{ji} - q_{ji'} &= \left((x_i - \boldsymbol{\gamma}'\mathbf{z}_i)^2 - (x_{i'} - \boldsymbol{\gamma}'\mathbf{z}_{i'})^2 \right) m_j + \left(\delta_j - 2m_j \boldsymbol{\alpha}'\overline{\mathbf{q}}_g\right) \left((x_i - \boldsymbol{\gamma}'\mathbf{z}_i) - (x_{i'} - \boldsymbol{\gamma}'\mathbf{z}_{i'}) \right) \\ &\quad + \mathbf{c}_j'(\mathbf{z}_i - \mathbf{z}_{i'}) + (u_{ji} - u_{ji'}) \end{aligned}$$

where $\mathbf{c}_j'$ equals the j'th row of $\mathbf{C}$. Then, again as in the generic model, we replace the unobservable true group mean $\overline{\mathbf{q}}_g$ with the leave-two-out estimate $\widehat{\mathbf{q}}_{g,-ii'} = \frac{1}{n_g - 2} \sum_{l \in g, l \neq i, i'} \mathbf{q}_l$, which then introduces an additional error term into the above equation.

Define group level instruments $\mathbf{r}_g$ as in the generic model. In particular, $\mathbf{r}_g$ can include group averages of x_i and of $\mathbf{z}_i$, using data from individuals i that are sampled in other time periods than the one currently being used for Engel curve identification. Define a vector of instruments $\mathbf{r}_{gii'}$ that contains the elements $\mathbf{r}_g, x_i, \mathbf{z}_i, x_{i'}, \mathbf{z}_{i'}$, and squares and cross products of these elements. We then, analogous to the generic model, obtain unconditional moments

$$\begin{aligned} 0 &= E\{[(q_{ji} - q_{ji'}) - \left((x_i - \boldsymbol{\gamma}'\mathbf{z}_i)^2 - (x_{i'} - \boldsymbol{\gamma}'\mathbf{z}_{i'})^2 \right) m_j \\ &\quad - (\delta_j - 2m_j \boldsymbol{\alpha}'\widehat{\mathbf{q}}_{g,-ii'}) \left((x_i - \boldsymbol{\gamma}'\mathbf{z}_i) - (x_{i'} - \boldsymbol{\gamma}'\mathbf{z}_{i'}) \right) - \mathbf{c}_j'(\mathbf{z}_i - \mathbf{z}_{i'})]\mathbf{r}_{gii'}\} \end{aligned} \tag{14}$$

for $j = 1, ..., J$. Identification of the parameters $\boldsymbol{\alpha}$, $\boldsymbol{\gamma}$, $\boldsymbol{\delta}$, $\mathbf{m}$, and $\mathbf{C}$ from these moments then directly follows, given sufficient variation in the covariates and instruments. See the Appendix for details.

The above analysis identifies Engel curves (i.e., demands holding prices fixed) so here $\mathbf{p} = (p_1, ...p_J)$ is just a vector of constants. Before proceeding to the full demand system, it is worth noting that most of the parameters of interest can be identified just from Engel curves without price variation. In particular, given $\boldsymbol{\alpha}$, $\boldsymbol{\gamma}$, $\boldsymbol{\delta}$, $\mathbf{m}$, and $\mathbf{C}$, we can identify $b_j = p_j \delta_j$ for each element b_j of $\mathbf{b}$ and $d_j = p_j m_j e^{\mathbf{b}' \ln \mathbf{p}}$ for each element d_j of $\mathbf{d}$. Also, recall $\widetilde{F}_i(\mathbf{p}) = \boldsymbol{\alpha}'\overline{\mathbf{q}}_g + \boldsymbol{\gamma}'\mathbf{z}_i$, so we can also identify and estimate the value of the relative cost of needs $\widetilde{F}_i(\mathbf{p})$ for any individual i in the given price regime $\mathbf{p}$. However, without price variation we cannot fully identify the matrix $\mathbf{A}$ of own and cross peer effects $\mathbf{A}$. An exception would be if $\mathbf{A}$ were diagonal, which would mean, e.g., that a consumer's perceived needs for luxuries could depend on the mean levels of luxuries consumed by his peers, but not on the mean levels of other goods consumed by his peers. If $\mathbf{A}$ were diagonal, then it could be identified from the Engel curve by $a_{jj} = \alpha_j / p_j$ for each element a_{jj} on the diagonal of $\mathbf{A}$ and each element of α_j of $\boldsymbol{\alpha}$.

To identify $\mathbf{A}$ without restricting it to be diagonal, we need data on multiple price regimes. Let t denote the time period that the above analysis applies to, and let $\mathbf{p}_t$ be the vector of prices in that time period. So we have identified is $\boldsymbol{\alpha}_t = \mathbf{A}'\mathbf{p}_t$. Assume we identify the above Engel curve model in J different time periods. Let $\mathbf{P}$ be the matrix consisting of columns $\mathbf{p}_t$ for $t = 1, ..., J$. Assuming sufficient price variation so that $\mathbf{P}$ is nonsingular, we can then identify $\mathbf{A}$ by $\mathbf{A}'$=$(\boldsymbol{\alpha}_1, ..., \boldsymbol{\alpha}_J)\mathbf{P}^{-1}$.

4.2 Estimating Engel Curves

Having shown that the Engel curves are identified from the moments of equation (14), we can directly can construct GMM estimators corresponding to these moments for $j = 1, ..., J-1$. As is standard in the estimation of continuous demand systems, we only need to estimate the model for goods $j = 1, ..., J-1$. The parameters for the last good J are then obtained from the adding up identity that $q_{Ji} = \left(x_i - \sum_{j=1}^{J-1} p_j q_{ji}\right)/p_J$.

As discussed for the generic model, GMM could either be done directly using the moments of equation (14), treating each (i, i') pair within each group as the unit of observation (and constructing clustered standard errors to account for the correlations among these observations within each group), or by aggregating the moments up to the group level as in equation (22).

The vector of instruments $\mathbf{r}_{gii'}$ can include arbitrary functions of $\mathbf{r}_g$, x_i, $\mathbf{z}_i$, $x_{i'}$, and $\mathbf{z}_{i'}$, where as before $\mathbf{r}_g$ could include group level averages of functions of x and $\mathbf{z}$, constructed using data from some other time period, or some other data source (e.g., census data), than the one that the data used for Engel curve estimation come from. Also as in the generic case, outside instruments $\mathbf{r}_g$ may not be necessary for identification, but we include them here because they may help efficiency, and because they are available in our empirical analysis data set.

Again mimicking the generic case, based on the above equations a sensible set of instruments $\mathbf{r}_{gii'}$ might be $(x_i - x_{i'})$, $(z_{ki} - z_{ki'})$, $(x_i - x_{i'})\,\mathbf{r}_g$, $(z_{ki} - z_{ki'})\,\mathbf{r}_g$, $(x_i^2 - x_{i'}^2)$, and $(z_{ki}^2 - z_{ki'}^2)$ for $k = 1, 2, ..., K$, where $\mathbf{r}_g$ equals the sample means of x and $\mathbf{z}$ constructed using data from other time periods. This constitutes a sufficient number of instruments, but if desired additional valid instruments would include more cross terms such as $(z_{ki}x_i - z_{ki'}x_{i'})$ and $(z_{1i}z_{2i} - z_{1i'}z_{2i'})$.

4.3 Estimating the Full Demand Model

Based on our identification results, we can estimate all of the parameters of our full demand system, assuming we have data from at least $J-1$ time periods. The moments we have for estimation are given by equation (14). Substituting $\boldsymbol{\alpha} = \mathbf{A}'\mathbf{p}$, $\boldsymbol{\gamma} = \mathbf{C}'\mathbf{p}$, $\delta_j = b_j/p_j$, and

$m_j = e^{-\mathbf{b}' \ln \mathbf{p}} d_j / p_j$ into this equation, and adding a t subscript to denote time, we have

$$0 = E\{\mathbf{r}_{gtii'}[(q_{ji} - q_{ji'}) - \left((x_i - \mathbf{p}_t' \mathbf{C}\mathbf{z}_i)^2 - (x_{i'} - \mathbf{p}_t' \mathbf{C}\mathbf{z}_{i'})^2 \right) \frac{d_j}{p_{jt} e^{\mathbf{b}' \ln \mathbf{p}_t}} \tag{15}$$

$$- \left(\frac{b_j}{p_{jt}} - \frac{2d_j}{p_{jt} e^{\mathbf{b}' \ln \mathbf{p}_t}} \mathbf{p}_t' \mathbf{A} \widehat{\mathbf{q}}_{tg,-ii'} \right) ((x_i - \mathbf{p}_t' \mathbf{C}\mathbf{z}_i) - (x_{i'} - \mathbf{p}_t' \mathbf{C}\mathbf{z}_{i'})) - \mathbf{c}_j' (\mathbf{z}_i - \mathbf{z}_{i'})]\} \tag{16}$$

where the vector of instruments $\mathbf{r}_{gtii'}$ is defined below. Equation (15) holds for goods $j = 1, ..., J$ and for all observed pairs of consumers i and i' in each group g in each period t.

Let $\widehat{x}_{(t)g}$ denote the sample mean of x_i over individuals i in group g in all time periods *except* time period t. So, e.g., if we have data from three time periods then $\widehat{x}_{(2)g}$ would be the average of x_i for all individuals in group g in time periods 1 and 3. Define $\widehat{\mathbf{z}}_{(t)g}$ analogously. Let $\mathbf{r}_{gt}$ be the vector of elements $\widehat{x}_{(t)g}$, $\widehat{\mathbf{z}}_{(t)g}$, and $\mathbf{p}_t$. The instrument vector $\mathbf{r}_{gtii'}$ then consists of the elements $(x_i - x_{i'})$, $(z_{ki} - z_{ki'})$, $(x_i - x_{i'})\, \mathbf{r}_{gt}$, $(z_{ki} - z_{ki'})\, \mathbf{r}_{gt}$, $(x_i^2 - x_{i'}^2)$, and $(z_{ki}^2 - z_{ki'}^2)$ for $k = 1, 2, ..., K$.

For estimation of these moments by GMM, let the unit of observation be each observed pair of consumers i and i' in each group g, in each period t. The total number of moments is the number of elements of $\mathbf{r}_{gtii'}$ times $J - 1$, because equation (15) applies to each good and each instrument. As discussed in the previous subsection, we only need to estimate the model for goods $j = 1, ..., J - 1$, because the parameters of the Jth equation are determined by the adding up constraint that expenditures on all goods sum to total expenditures x.

As before, for inference we need to apply clustered standard errors, where each cluster is defined as all of the pairs of observations in all time periods for each group g. We cluster over time as well as across individuals to allow for possible serial correlation in the errors.

4.4 Estimating the Demand Model With Random Effects

The identification and associated estimation discussed so far is based on fixed effects. However, as with the generic model, a great deal of information is lost by differencing out the fixed effects. We now consider adding additional random effects assumptions to the demand model, to thereby provide additional moments for GMM estimation that do not entail differencing.

For the demand system with random effects, as in the generic model we need to separate the quadratic from the linear terms in $\overline{\mathbf{q}}_g$, so rewrite the Engel curve model of equation (12) as

$$\mathbf{q}_i = \mathbf{m}\boldsymbol{\alpha}' \overline{\mathbf{q}}_g \overline{\mathbf{q}}_g' \boldsymbol{\alpha} + (2 (\boldsymbol{\gamma}' \mathbf{z}_i - x_i) \mathbf{m}\boldsymbol{\alpha}' - \boldsymbol{\delta}\boldsymbol{\alpha}' + \mathbf{A}) \overline{\mathbf{q}}_g + (x_i - \boldsymbol{\gamma}' \mathbf{z}_i)^2 \mathbf{m} + (x_i - \boldsymbol{\gamma}' \mathbf{z}_i) \boldsymbol{\delta} + \mathbf{C}\mathbf{z}_i + \mathbf{v}_g + \mathbf{u}_i.$$

Equivalently, for each good $j \in \{1, ..., J\}$, this equation is

$$q_{ji} = \boldsymbol{\alpha}'\overline{\mathbf{q}}_g\overline{\mathbf{q}}_g'\boldsymbol{\alpha}m_j + \left(2(\boldsymbol{\gamma}'\mathbf{z}_i - x_i)m_j\boldsymbol{\alpha}' - \delta_j\boldsymbol{\alpha}' + \mathbf{A}_j'\right)\overline{\mathbf{q}}_g + (x_i - \boldsymbol{\gamma}'\mathbf{z}_i)^2 m_j + (x_i - \boldsymbol{\gamma}'\mathbf{z}_i)\,\delta_j + \mathbf{c}_j'\mathbf{z}_i + v_{gj} + u_{ji}$$

where $\mathbf{A}_j'$ is the jth row of $\mathbf{A}$. Unlike in the fixed effects demand model, we cannot simplify by replacing $\mathbf{v}_g$ with $\widetilde{\mathbf{v}}_g$, because with random effects we assume that $\mathbf{v}_g$ is independent of $(x, \mathbf{z}, \mathbf{u})$, and this independence would not hold for $\widetilde{\mathbf{v}}_g$.

Next, exactly analogous to the generic random effects model, replace the quadratic term $\overline{\mathbf{q}}_g\overline{\mathbf{q}}_g'$ with $\widehat{\mathbf{q}}_{g,-ii'}\mathbf{q}_{i'}'$ and replace the linear term $\overline{\mathbf{q}}_g$ with $\widehat{\mathbf{q}}_{g,-ii'}$. Let $\mathbf{r}_g$ be defined as in the fixed effects Engel curve model, and let the vector of instruments $\mathbf{r}_{gi}$ contain the elements $\mathbf{r}_g$, $x_i, \mathbf{z}_i$, and squares and cross products of these elements (note that unlike the fixed effects case, $\mathbf{r}_{gi}$ here does not contain functions of $x_{i'}$ and $\mathbf{z}_{i'}$). We show in the Appendix that the following unconditional moment holds:

$$E\left[\mathbf{r}_{gi}\left(q_{ji} - m_j\boldsymbol{\alpha}'\widehat{\mathbf{q}}_{g,-ii'}\boldsymbol{\alpha}'\mathbf{q}_{i'} + (2m_j\left(x_i - \boldsymbol{\gamma}'\mathbf{z}_i\right)\boldsymbol{\alpha}' + \delta_j\boldsymbol{\alpha}' - \mathbf{A}_j')\widehat{\mathbf{q}}_{g,-ii'}\right.\right.$$
$$\left.\left. - m_j(x_i - \mathbf{z}_i'\boldsymbol{\gamma})^2 - \delta_j\left(x_i - \boldsymbol{\gamma}'\mathbf{z}_i\right) - \mathbf{c}_j'\mathbf{z}_i - v_{j0}\right)\right] = 0\ ,$$

where v_{j0} is a constant term for each j. As discussed in the Appendix, if we plug in $\boldsymbol{\alpha} = \mathbf{A}'\mathbf{p}$, $\boldsymbol{\gamma} = \mathbf{C}'\mathbf{p}$, $\delta_j = b_j/p_j$, and $m_j = e^{-\mathbf{b}'\ln\mathbf{p}}d_j/p_j$, all of the structural parameters $\mathbf{A}$, $\mathbf{C}$, $\mathbf{b}$, and $\mathbf{d}$ can be identified and estimated from these Engel curve moments, given sufficient variation in the covariates and the instruments. This is different from the fixed effects model. There, the matrix $\mathbf{A}$ was not identified without price variation, because the term $\mathbf{A}\overline{\mathbf{q}}_{\mathbf{g}}$ got differenced out.

Given these instruments and moments, estimation proceeds by GMM just as in the fixed effects case. Also as in that case, if we have multiple time periods, we can add t subscripts and again do GMM as in the fixed effects case, and we could combine the fixed effects and random effects moments into a single large GMM.

5 Empirical sections

To be written

6 Are peer effects wasted?

Our model assumes that there are no compensating benefits associated with the losses from (i.e., costs associated with satisfying) increased needs due to peer effects. However, costs of keeping up with peers are not necessarily all wasted resources. Some fraction of these peer effects could be construed as benefits to society stemming from higher average standards of living. For example, there is a value to society when most people have internet access, even if that makes internet access become a perceived need. Formally, we could include separable component to utility where one benefits from the average standard of living in one's society, which could in turn be related to country or region level fixed costs. As long as one's own behavior makes a negligibly small contribution to this separable component of utility, our numerical analyses are unaffected, though the interpretation of our results changes. We do not take a stand on what fraction of the costs of peer effects are not a wasted from the standpoint of society, though the peer effects of luxuries are unlikely to be useful for the most part.

7 Conclusions

To be written

8 References

Akay, Alpaslan, and Peter Martinsson. (2011) "Does relative income matter for the very poor? Evidence from rural Ethiopia." Economics Letters 110.3: 213-215.

Blackorby and Donaldson, 1994, "Absolute Equivalence Scale Exactness", some book find that thing.

Blume, Lawrence E. Brock, William A., Durlauf, Steven N., Ioannides, Yannis M., 2010. "Identification of Social Interactions," Economics Series 260, Institute for Advanced Studies.

Brock, William A. and Steven N. Durlauf (2001). "Interactions-based Models," Handbook of Econometrics 5: 3297 - 3380 (J.Heckman&E.Leamer,Eds.).Amsterdam: North-Holland

Carlsson, F., Gupta, G., & Johansson-Stenman, O. (2009). Keeping up with the Vaishyas? Caste and relative standing in India. Oxford Economic Papers, 61(1), 52-73.

Chao, Angela, and Juliet B. Schor. "Empirical tests of status consumption: Evidence from women's cosmetics." Journal of Economic Psychology 19.1 (1998): 107-131.

Clark, Andrew E., and Claudia Senik. "Who compares to whom? the anatomy of income comparisons in europe." The Economic Journal 120.544 (2010): 573-594.

Clark, A. E., Frijters, P., & Shields, M. A. (2008). Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles. Journal of Economic

Literature, 95-144.

Donaldson, David and Krishna Pendakur, 2006, "The Identification of Fixed Costs from Behaviour" Journal of Business and Economics Statistics.

Dupor, Bill, and Wen-Fang Liu. "Jealousy and equilibrium overconsumption." The American Economic Review 93.1 (2003): 423-428.

Fontaine, X., & Yamada, K. (2012). Economic Comparison and Group Identity: Lessons from India.

Frank, Robert H. (2012) "The Easterlin Paradox Revisited" Emotion 12.6 (2012): 1188-1191.

Frank, Robert H., Adam Seth Levine, and Oege Dijk. "Expenditure cascades." Cornell University mimeograph. Ithaca: Cornell University (2005).

Frank, Robert H., (1999), Luxury Fever: Money and Happiness in an Era of Excess, The Free Press.

Gali, J. (1994). Keeping up with the Joneses: Consumption externalities, portfolio choice, and asset prices. Journal of Money, Credit and Banking, 26(1), 1-8.

Geary R.C. (1949), "A note on a constant utility index of the cost of living", Review of Economic Studies, 18: 65-66.

Gorman, W. M. (1976). Tricks with utility functions. Essays in economic analysis, 211-243.

Hu, Y. & J-L Shiu, 2011. "Nonparametric Identification Using Instrumental Variables: Sufficient Conditions For Completeness," Economics Working Paper Archive 581, The Johns Hopkins University,Department of Economics.

Kahneman, D. (1992). Reference points, anchors, norms, and mixed feelings. Organizational behavior and human decision processes, 51(2), 296-312.

Kalyanaram, G., & Winer, R. S. (1995). Empirical generalizations from reference price research. Marketing Science, 14(3 supplement), G161-G169.

Kőszegi, B., & Rabin, M. (2006). A model of reference-dependent preferences. The Quarterly Journal of Economics, 121(4), 1133-1165.

Maurer, J., & Meier, A. (2008). Smooth it Like the 'Joneses'? Estimating Peer-Group Effects in Intertemporal Consumption Choice. The Economic Journal, 118(527), 454-476.

Manski, C. F. (1993). Identification of Endogenous Social Effects: The Reflection Problem. The Review of Economic Studies. 60: 531-542.

Manski, Charles F. (2000). "Economic analysis of social interactions," Journal of Economic Perspectives 14: 115-136.

Pendakur, Krishna, 2002, "Semiparametric Estimation of Lifetime Child Costs" Journal of Appied Econometrics.

Pendakur, Krishna, 1999, "Semiparametric Estimation and Testing of Base-Independent Equivalence Scales" Journal of Econometrics.

Rabin, M. (1998). Psychology and economics. Journal of economic literature, 36(1), 11-46.

Ravallion, Martin, and Michael Lokshin. "Who cares about relative deprivation?" Journal of Economic Behavior & Organization 73.2 (2010): 171-185.

Ravina, E. (2007). Habit persistence and keeping up with the Joneses: evidence from micro data.

Samuelson, P. A. (1947). "Some Implications of Linearity." The Review of Economic Studies, 15(2), 88-90.

Stone, R (1954), "Linear Expenditure System and Demand Analysis: An Application to the Pattern of British Demand", Economic Journal, 64:511-527.

Veblen, T. B. (1899). The Theory of the Leisure Class, An Economic Study of Institutions. London: Macmillan Publishers.

Yamada, K., & Sato, M. (2010). Another Avenue for Anatomy of Income Comparisons: Evidence from Hypothetical Choice Experiments. Journal of Economic Behavior and Organization.

Appendix: Derivations

8.1 Generic Model Identification and Estimation

Let y_i denote an outcome and $\mathbf{x}_i$ denote a K vector of regressors x_{ki} for an individual i. Let $i \in g$ denote that the individual i belongs to group g. For each group g, assume we observe $n_g = \sum_{i \in g} 1$ individuals, where n_g is a small fixed number which does *not* go to infinity. Let $\overline{y}_g = E(y_i \mid i \in g)$, $\widehat{y}_{g,-ii'} = \sum_{l \in g, l \neq i,i'} y_l/(n_g - 2)$, and $\varepsilon_{yg,-ii'} = \widehat{y}_{g,-ii'} - \overline{y}_g$, so $\overline{y}_g$ is the true group mean outcome and $\widehat{y}_{g,-ii'}$ is the observed leave-two-out group average outcome in our data, and $\varepsilon_{yg,-ii'}$ is the estimation error in the leave-two-out sample group average. Define $\overline{\mathbf{x}}_g = E(\mathbf{x}_i \mid i \in g)$, $\overline{\mathbf{x}\mathbf{x}'_g} = E(\mathbf{x}_i\mathbf{x}'_i \mid i \in g)$, and similarly define $\widehat{\mathbf{x}}_{g,-ii'}$, $\widehat{\mathbf{x}\mathbf{x}'}_{g,-ii'}$, $\varepsilon_{\mathbf{x}g,-ii'}$ and $\varepsilon_{\mathbf{xx}g,-ii'}$ analogously to $\widehat{y}_{g,-ii'}$, and $\varepsilon_{yg,-ii'}$.

Consider the following single equation model (the multiple equation analog is discussed later). For each individual i in group g, let

$$y_i = \left(\overline{y}_g a + \mathbf{x}'_i\mathbf{b}\right)^2 d + \left(\overline{y}_g a + \mathbf{x}'_i\mathbf{b}\right) + v_g + u_i \tag{17}$$

where v_g is a group level fixed effect and u_i is an idiosyncratic error. The goal here is identification and estimation of the effects of $\overline{y}_g$ and x_i on y_i, which means identifying the coefficients a, $\mathbf{b}$, and d.

We could have written the seemingly more general model

$$y_i = \left(\overline{y}_g a + \mathbf{x}_i'\mathbf{b} + c\right)^2 d + \left(\overline{y}_g a + \mathbf{x}_i'\mathbf{b} + c\right) k + v_g + u_i$$

where c and k are additional constants to be estimated. However, it can be shown (see the appendix for details), that by suitably redefining the fixed effect v_g and the constants a, $\mathbf{b}$, and d, that this equation is equivalent either to equation (17) or to $y_i = \left(\overline{y}_g a + \mathbf{x}_i'\mathbf{b}\right)^2 + v_g + u_i$. Since this latter equation is strictly easier to identify and estimate, and is irrelevant for our empirical application, we will rule it out and therefore without loss of generality replace the more general model with equation (17).

Next observe that, regardless of what we assume about within group or between group sample sizes, if this model were linear (i.e., $d = 0$), then we would not be able to identify the effect of $\overline{y}_g$ on y_i, i.e., we would not be able to identify the peer effect. This is because, if $d = 0$, then there is no way to separate $\overline{y}_g$ from the group level fixed effect v_g. All values of a would be observationally equivalent, by suitable redefinitions of v_g. This is a manifestation of the reflection problem, which we overcome by a combination of nonlinearity and functional form restrictions.

We assume that the number of groups G goes to infinity, but we do NOT assume that n_g goes to infinity, so $\widehat{y}_{g,-ii'}$ is not a consistent estimator of $\overline{y}_g$. We instead treat $\varepsilon_{yg,-ii'} = \widehat{y}_{g,-ii'} - \overline{y}_g$ as measurement error in $\widehat{y}_{g,-ii'}$, which is not asymptotically negligible. This makes sense for data like ours where only a small number of individuals are observed within each peer group. This may also be a sensible assumption in many standard applications where true peer groups are small. For example, in a model where peer groups are classrooms, failure to observe a few children in a class of one or two dozen students may mean that the observed class average significantly mismeasures the true class average.

Formally, our first identification theorem makes assumptions A1 to A3 below.

Assumption A1: Each individual i in group g satisfies equation (17). $\mathbf{x}_i$ is a K-dimensional vector of covariates. For each $k \in \{1, ..., K\}$, for each group g with $i \in g$ and $i' \in g$, $\Pr\left(\mathbf{x}_{ik} = \mathbf{x}_{i'k}\right) > 0$. Unobserved v_g are group level fixed effects. Unobserved errors u_i are independent across groups g and have $E(u_i \mid \text{all } \mathbf{x}_{i'} \text{ having } i' \in g \text{ where } i \in g) = 0$. The number of observed groups $G \to \infty$. For each observed group g, we observe a sample of $n_g \geq 3$ observations of $y_i, \mathbf{x}_i$.

Assumption A1 essentially defines the model. Note that Assumption A1 does not require that $n_g \to \infty$. We can allow the observed sample size n_g in each group g to be fixed, or to change with the number of groups G. The true number of individuals comprising each group is unknown and could be finite.

Assumption A2: The coefficients a, $\mathbf{b}$, d are unknown constants satisfying $d \neq 0$, $\mathbf{b} \neq 0$, and $[1 - a(2\mathbf{b}'\overline{\mathbf{x}}_g d + 1)]^2 - 4a^2 d[d\mathbf{b}'\overline{\mathbf{x}\mathbf{x}'_g}\mathbf{b} + \mathbf{b}'\overline{\mathbf{x}}_g + v_g] \geq 0$.

In Assumption A2, as discussed above $d \neq 0$ is needed to avoid the reflection problem. Having $\mathbf{b} \neq 0$ is necessary since otherwise we would have nothing exogenous in the model. Finally, note that the inequality in Assumption A2 takes the form of a simple lower or upper bound (depending on the sign of d) on each fixed effect v_g. This inequality must hold to ensure that an equilibrium exists for each group, thereby avoiding Tamer's (2003) potential incoherence problem. To see this, plug equation (17) for y_i into $\overline{y}_g = E(y_i \mid i \in g)$. This yields a quadratic in $\overline{y}_g$, which, if $a \neq 0$, has the solution

$$\overline{y}_g = \frac{1 - a(2\mathbf{b}'\overline{\mathbf{x}}_g d + 1) \pm \sqrt{[1 - a(2\mathbf{b}'\overline{\mathbf{x}}_g d + 1)]^2 - 4a^2 d[d\mathbf{b}'\overline{\mathbf{x}\mathbf{x}'_g}\mathbf{b} + \mathbf{b}'\overline{\mathbf{x}}_g + v_g]}}{2a^2 d} \tag{18}$$

if the inequality in Assumption A2 is satisfied (while if a does equal zero, then the model will be trivially identified because in that case there aren't any peer effects). We do not take a stand on which root of equation (18) is chosen by consumers, we just make the following assumption.

Assumption A3: Individuals within each group agree on an equilibrium selection rule.

For identification, we need to remove the fixed effect from equation (17), which we do by subtracting off another individual in the same group. For each $(i, i') \in g$, consider pairwise difference

$$\begin{aligned} y_i - y_{i'} &= 2ad\overline{y}_g \mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'}) + d\mathbf{b}'(\mathbf{x}_i\mathbf{x}'_i - \mathbf{x}_{i'}\mathbf{x}'_{i'})\mathbf{b} + \mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'}) + u_i - u_{i'} \\ &= 2ad\widehat{y}_{g,-ii'}\mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'}) + d\mathbf{b}'(\mathbf{x}_i\mathbf{x}'_i - \mathbf{x}_{i'}\mathbf{x}'_{i'})\mathbf{b} + \mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'}) \\ &\quad + u_i - u_{i'} - 2ad\varepsilon_{yg,-ii'}\mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'}), \end{aligned} \tag{19}$$

where the second equality is obtained by replacing $\overline{y}_g$ on the right hand side with $\widehat{y}_{g,-ii'} - \varepsilon_{yg,-ii'}$. In addition to removing the fixed effects v_g, the pairwise difference also removed the linear term $a\overline{y}_g$, and the squared term $da^2\overline{y}_g^2$. The second equality in equation (19) shows that $y_i - y_{i'}$ is linear in observable functions of data, plus a composite error term $u_i - u_{i'} - 2ad\varepsilon_{yg,-ii'}\mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'})$ that contains both $\varepsilon_{yg,-ii'}$ and $u_i - u_{i'}$. By Assumption

A1, $u_i - u_{i'}$ is conditionally mean independent of $\mathbf{x}_i$ and $\mathbf{x}_{i'}$. It can also be shown (see the Appendix) that

$$\varepsilon_{yg,-ii'} = 2ad\overline{y}_g\mathbf{b}'\boldsymbol{\varepsilon}_{\mathbf{x}g,-ii'} + \mathbf{b}'\boldsymbol{\varepsilon}_{\mathbf{xx}g,-ii'}\mathbf{b}d + \mathbf{b}'\boldsymbol{\varepsilon}_{\mathbf{x}g,-ii'} + \widehat{u}_{g,-ii'}.$$

where

$$\boldsymbol{\varepsilon}_{\mathbf{x}g,-ii'} = \frac{1}{n_g - 2}\sum_{l\in g, l\neq i,i'}(\mathbf{x}_l - \overline{\mathbf{x}}_g)\,;\ \boldsymbol{\varepsilon}_{\mathbf{xx}g,-ii'} = \frac{1}{n_g - 2}\sum_{l\in g, l\neq i,i'}\left(\mathbf{x}_l\mathbf{x}_l' - \overline{\mathbf{xx}'}_g\right).$$

Substituting this expression into equation (19) gives an expression for $y_i - y_{i'}$ that is linear in $\widehat{y}_{g,-ii'}(\mathbf{x}_i - \mathbf{x}_{i'})$, $(\mathbf{x}_i\mathbf{x}_i' - \mathbf{x}_{i'}\mathbf{x}_{i'}')$, $(\mathbf{x}_i - \mathbf{x}_{i'})$, and a composite error term.

In addition to the conditionally mean independent errors $u_i - u_{i'}$ and $\widehat{u}_{g,-ii'}$, the components of this composite error term include $\boldsymbol{\varepsilon}_{\mathbf{x}g,-ii'}$ and $\boldsymbol{\varepsilon}_{\mathbf{xx}g,-ii'}$, which are measurement errors in group level mean regressors. If we assumed that the number of individuals in each group went to infinity, then these epsilon errors would asymptotically shrink to zero, and the the resulting identification and estimation would be simple. In our case, these errors do not go to zero, but one might still consider estimation based on instrumental variables. This will be possible with further assumptions on the data.

In the next assumption we allow for the possibility of observing group level variables $\mathbf{r}_g$ that may serve as instruments for $\widehat{y}_{g,-ii'}$. Such instruments may not be necessary, but if such instruments are available (as they will be in our later empirical application), they can help both in weakening sufficient conditions for identification and for later improving estimation efficiency.

Assumption A4: Let $\mathbf{r}_g$ be a vector (possibly empty) of observed group level instruments that are independent of each u_i. Assume $E\left((\mathbf{x}_i - \overline{\mathbf{x}}_g) \mid i \in g, \overline{\mathbf{x}}_g, \overline{\mathbf{xx}'}_g, v_g, \mathbf{r}_g\right) = 0$, $E\left(\left(\mathbf{x}_i\mathbf{x}_i' - \overline{\mathbf{xx}'}_g\right) \mid i \in g, \mathbf{r}_g\right) = 0$, and that $\mathbf{x}_i - \overline{\mathbf{x}}_g$ and $\mathbf{x}_i\mathbf{x}_i' - \overline{\mathbf{xx}'}_g$ are independent across individuals i.

Assumption A4 corresponds to (but is a little stronger than) standard instrument validity assumptions. A sufficient condition for the equalities in Assumption A4 to hold is let $\boldsymbol{\varepsilon}_{ix} = \mathbf{x}_i - \overline{\mathbf{x}}_g$ be independent across individuals, and assume that $E(\boldsymbol{\varepsilon}_{ix} \mid \overline{\mathbf{x}}_g, \overline{\mathbf{xx}'}_g, v_g, \mathbf{r}_g$ for $i \in g) = 0$ and $E\left(\boldsymbol{\varepsilon}_{ix}\boldsymbol{\varepsilon}_{ix}' \mid \overline{\mathbf{x}}_g, \mathbf{r}_g \text{ for } i \in g\right) = E\left(\boldsymbol{\varepsilon}_{ix}\boldsymbol{\varepsilon}_{ix}' \mid i \in g\right)$. To see this, we have

$$\begin{aligned}
E(\mathbf{x}_i\mathbf{x}_i' - \overline{\mathbf{xx}'}_g \quad &| \quad i \in g, \overline{\mathbf{x}}_g, \mathbf{r}_g) = E[(\boldsymbol{\varepsilon}_{ix} + \overline{\mathbf{x}}_g)(\boldsymbol{\varepsilon}_{ix} + \overline{\mathbf{x}}_g)' \mid i \in g, \overline{\mathbf{x}}_g, \mathbf{r}_g] - \overline{\mathbf{xx}'}_g \\
&= \quad E(\boldsymbol{\varepsilon}_{ix}\boldsymbol{\varepsilon}_{ix}' \mid i \in g, \overline{\mathbf{x}}_g, \mathbf{r}_g) + E(\mathbf{x}_i | i \in g)E(\mathbf{x}_i' | i \in g) - E(\mathbf{x}_i\mathbf{x}_i' | i \in g) \\
&= \quad E(\boldsymbol{\varepsilon}_{ix}\boldsymbol{\varepsilon}_{ix}' \mid i \in g, \overline{\mathbf{x}}_g, \mathbf{r}_g) - E(\boldsymbol{\varepsilon}_{ix}\boldsymbol{\varepsilon}_{ix}' | i \in g)
\end{aligned}$$

A simpler but stronger sufficient condition would just be that $\boldsymbol{\varepsilon}_{ix}$ are independent across individuals i and independent of group level variables $\overline{\mathbf{x}}_g, \overline{\mathbf{x}\mathbf{x}'}_g, v_g, \mathbf{r}_g$. Essentially, this corresponds to saying that any individual i in group g has a value of $\mathbf{x}_i$ that is a randomly drawn deviation around their group mean level $\overline{\mathbf{x}}_g$. The first two equalities in A4 are used to show that $E(\varepsilon_{yg,-ii'} \mid \mathbf{r}_g) = 0$, and the independence of measurement errors across individuals is used to show $E(\varepsilon_{yg,-ii'}(\mathbf{x}_i - \mathbf{x}_{i'}) \mid \mathbf{r}_g, \mathbf{x}_i, \mathbf{x}_{i'}) = (\mathbf{x}_i - \mathbf{x}_{i'})E(\varepsilon_{yg,-ii'} \mid \mathbf{r}_g) = 0$, so that $\mathbf{x}_i$ and $\mathbf{x}_{i'}$ are valid instruments. Given Assumptions A1 and A4, one can directly verify that

$$E\left[y_i - y_{i'} - (2ad\widehat{y}_{g,-ii'}\mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'}) + d\mathbf{b}'(\mathbf{x}_i\mathbf{x}_i' - \mathbf{x}_{i'}\mathbf{x}_{i'}')\mathbf{b} + \mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'})) \mid \mathbf{r}_g, \mathbf{x}_i, \mathbf{x}_{i'}\right] = 0. \tag{20}$$

Under Assumptions A1 to A4, $(\mathbf{x}_i - \mathbf{x}_{i'})E(\widehat{y}_{g,-ii'}|\mathbf{r}_g, \mathbf{x}_i, \mathbf{x}_{i'})$ is linearly independent of $(\mathbf{x}_i - \mathbf{x}_{i'})$ and $(\mathbf{x}_i\mathbf{x}_i' - \mathbf{x}_{i'}\mathbf{x}_{i'}')$ with a positive probability. These conditional moments could therefore be used to identify the coefficients $2ad\mathbf{b}$, $b_1d\mathbf{b}$,...$b_Kd\mathbf{b}$, and $\mathbf{b}$, which we could then immediately solve for the three unknowns a, $\mathbf{b}$, d. Note that we have $K+2$ parameters which need to be estimated, and even if no $\mathbf{r}_g$ are available, we have $2K$ instruments $\mathbf{x}_i$ and $\mathbf{x}_{i'}$. The level of $\mathbf{x}_i$ as well as the difference $\mathbf{x}_i - \mathbf{x}_{i'}$ may be useful as an instrument (and nonlinear functions of $\mathbf{x}_i$ can be useful), because (18) shows that $\overline{y}_g$ and hence $\widehat{y}_{g,-ii'}$ is nonlinear in $\overline{\mathbf{x}}_g$, and $\mathbf{x}_i$ is correlated with $\overline{\mathbf{x}}_g$ by $\mathbf{x}_i = \boldsymbol{\varepsilon}_{ix} + \overline{\mathbf{x}}_g$.

The above derivations outline how we obtain identification, while the formal proof is given in Theorem 1 below (details are provided in the Appendix). To simplify estimation, we construct unconditional rather than conditional moments for identification and later estimation. Let $\mathbf{r}_{gii'}$ denote a vector of any chosen functions of $\mathbf{r}_g$, $\mathbf{x}_i$, and $\mathbf{x}_{i'}$, which we will take as an instrument vector. It then follows immediately from equation (20) that

$$E\left[\left(y_i - y_{i'} - (1 + 2ad\widehat{y}_{g,-ii'})\sum_{k=1}^{K} b_k(x_{ki} - x_{ki'}) - d\sum_{k=1}^{K}\sum_{k'=1}^{K} b_k b_{k'}(x_{ki}x_{k'i} - x_{ki'}x_{k'i'})\right)\mathbf{r}_{gii'}\right] = 0. \tag{21}$$

Let

$$\begin{aligned} L_{1gii'} &= (y_i - y_{i'}), \quad L_{2kgii'} = (x_{ki} - x_{ki'}), \\ L_{3kgii'} &= \widehat{y}_{g,-ii'}(x_{ki} - x_{ki'}), \quad L_{4kk'gii'} = x_{ki}x_{k'i} - x_{ki'}x_{k'i'} \end{aligned}$$

Equation (21) is linear in these L variables and so could be estimated by GMM. This linearity also means they can be aggregated up to the group level as follows. Define

$$\Gamma_g = \{(i, i') \mid i \text{ and } i' \text{ are observed}, i \in g, i' \in g, i \neq i'\}$$

So Γ_g is the set of all observed pairs of individuals i and i' in the group g. For $\ell \in \{1, 2k, 3k, 4kk' \mid k, k' = 1, ..., K\}$, define vectors

$$\mathbf{Y}_{\ell g} = \frac{\sum_{(i,i')\in\Gamma_g} L_{\ell gii'}\mathbf{r}_{gii'}}{\sum_{(i,i')\in\Gamma_g} 1}$$

Then averaging equation (21) over all $(i, i') \in \Gamma_g$ gives the unconditional group level moment vector

$$E\left(\mathbf{Y}_{1g} - \sum_{k=1}^{K} b_k \mathbf{Y}_{2kg} - 2ad\sum_{k=1}^{K} b_k \mathbf{Y}_{3kg} - d\sum_{k=1}^{K}\sum_{k'=1}^{K} b_k b_{k'} \mathbf{Y}_{4kk'g}\right) = 0. \tag{22}$$

Suppose the instrumental vector $\mathbf{r}_{gii'}$ is q dimensional. Denote the $q \times (K^2 + 2K)$ matrix $\mathbf{Y}_g = (\mathbf{Y}_{21g}, ...\mathbf{Y}_{2Kg}, \mathbf{Y}_{31g}, ...\mathbf{Y}_{3Kg}, \mathbf{Y}_{411g}, \cdots, \mathbf{Y}_{4KKg})$. The following assumption ensures that we can identify the coefficients in this equation.

Assumption A5: $E(\mathbf{Y}_g')E(\mathbf{Y}_g)$ is nonsingular.

Theorem 1. Given Assumptions A1, A2, A3, A4, and A5, the coefficients a, $\mathbf{b}$, d are identified.

As noted earlier, Assumptions A1 to A4 should generally suffice for identification. Assumption A5 is used to obtain more convenient identification based on unconditional moments. Assumption A5 is itself stronger than necessary, since it would suffice to identify arbitrary coefficients of the $\mathbf{Y}$ variables, ignoring all of the restrictions among them that are given by equation (22).

Given the identification in Theorem 1, based on equation (22) we can immediately construct a corresponding group level GMM estimator

$$\begin{aligned}\left(\widehat{a}, \widehat{b}_1, ...\widehat{b}_K, \widehat{d}\right) \quad = \quad & \arg\min \left[\frac{1}{G}\sum_{g=1}^{G}\left(\mathbf{Y}_{1g} - \sum_{k=1}^{K} b_k \mathbf{Y}_{2kg} - 2ad\sum_{k=1}^{K} b_k \mathbf{Y}_{3kg} - d\sum_{k=1}^{K}\sum_{k'=1}^{K} b_k b_{k'} \mathbf{Y}_{4kk'g}\right)\right]' \\ & \cdot\widehat{\Omega}\left[\frac{1}{G}\sum_{g=1}^{G}\left(\mathbf{Y}_{1g} - \sum_{k=1}^{K} b_k \mathbf{Y}_{2kg} - 2ad\sum_{k=1}^{K} b_k \mathbf{Y}_{3kg} - d\sum_{k=1}^{K}\sum_{k'=1}^{K} b_k b_{k'} \mathbf{Y}_{4kk'g}\right)\right] \quad (23)\end{aligned}$$

for some positive definite moment weighting matrix $\widehat{\Omega}$. In equation (23), each group g corresponds to a single observation, the number of observations within each group is assumed to be fixed, and recall we have assumed the number of groups G goes to infinity. Since this equation has removed the v_g terms, there is no remaining correlation across the group level errors, and therefore standard cross section GMM inference will apply. Also, with the number of observed individuals within each group held fixed, there is no loss in rates of convergence by aggregating up to the group level in this way.

One could alternatively apply GMM to equation (21), where the unit of observation would then be each pair (i, i') in each group. However, when doing inference one would then need to use clustered standard errors, treating each group g as a cluster, to account for the correlation that would, by construction, exist among the observations within each group. In this case,

$$\left(\widehat{a}, \widehat{b}_1, ...\widehat{b}_K, \widehat{d}\right) = \arg\min \left(\frac{\sum_{g=1}^{G} \sum_{(i,i')\in\Gamma_g} \mathbf{m}_{gii'}}{\sum_{g=1}^{G} \sum_{(i,i')\in\Gamma_g} 1}\right)' \widehat{\Omega} \left(\frac{\sum_{g=1}^{G} \sum_{(i,i')\in\Gamma_g} \mathbf{m}_{gii'}}{\sum_{g=1}^{G} \sum_{(i,i')\in\Gamma_g} 1}\right), \tag{24}$$

where

$$\mathbf{m}_{gii'} = L_{1gii'}\mathbf{r}_{gii'} - \sum_{k=1}^{K} b_k L_{2kgii'}\mathbf{r}_{gii'} - 2ad \sum_{k=1}^{K} b_k L_{3kgii'}\mathbf{r}_{gii'} - d \sum_{k=1}^{K} \sum_{k'=1}^{K} b_k b_{k'} L_{4kk'gii'}\mathbf{r}_{gii'}.$$

The remaining issue is how to select the vector of instruments $\mathbf{r}_{gii'}$, the elements of which are functions of $\mathbf{r}_g, \mathbf{x}_i, \mathbf{x}_{i'}$ chosen by the econometrician. Based on equation (21), $\mathbf{r}_{gii'}$ should include the differences $x_{ki} - x_{ki'}$ and $x_{ki}x_{k'i} - x_{ki'}x_{k'i'}$ for all k, k' from 1 to K, and should include terms that will correlate with $\widehat{y}_{g,-ii'}(x_{ki} - x_{ki'})$. Using equation (18) as a guide for what determines $\overline{y}_g$ and hence what should correlate with $\widehat{y}_{g,-ii'}$, suggests that $\mathbf{r}_{gii'}$ could include, e.g., $x_{ki}(x_{ki} - x_{ki'})$ or $x_{ki}^{1/2}(x_{ki} - x_{ki'})$.

We might also have available additional instruments $\mathbf{r}_g$ that come from other data sets. A strong set of instruments for $\widehat{y}_{g,-ii'}(x_{ki} - x_{ki'})$ could be $(x_{ki} - x_{ki'})\mathbf{r}_g$, where $\mathbf{r}_g$ is a vector of one or more group level variables that are correlated with $\overline{y}_g$, but still satisfy Assumption A4. One such possible $\mathbf{r}_g$ is a vector of group means of functions of $\mathbf{x}$ that are constructed using individuals that are observed in the same group as individual i, but in a different time period of our survey. For example, we might let $\mathbf{r}_g$ include $\widehat{\mathbf{x}}_{gt\cdot} = \sum_{s\neq t} \sum_{i\in gs} \mathbf{x}_i / \sum_{s\neq t} \sum_{i\in gs} 1$ where s indicates the period and t is the current period. In our empirical application, since the data take the form of repeated cross sections rather than panels, different individuals are observed in each time period. So $\widehat{\mathbf{x}}_{gt\cdot}$ is just an estimate of the group mean of $\overline{\mathbf{x}}_g$, but based on data from time periods other than one used for estimation. This produces the necessary uncorrelatedness (instrument validity) conditions in Assumption A4. The relevance of these instruments (the nonsingularity condition in Assumption A5) will hold as long as group level moments of functions of $\mathbf{x}$ in one time period are correlated with the same group level moments in other periods.

In our later empirical application, what corresponds to the vector $\mathbf{x}_i$ here includes the total expenditures, age, and other charecteristics of a consumer i, so Assumptions A4 and A5 will hold if the distribution of income and other characteristics within groups are sufficiently similar across time periods, while the specific individuals within each group who are sampled change over time. The nonlinearity of $\overline{y}_g$ in equation (18) shows that additional nonlinear functions of $\widehat{\mathbf{x}}_{gt\cdot}$, could also be valid and potentially useful additional instruments.

8.2 Proof of Theorem 1

We first show that we may without loss of generality assume $c = 0$ and $k = 1$ the single equation generic model. Suppose that

$$y_i = \left(\overline{y}_g a + \mathbf{x}_i'\mathbf{b} + c\right)^2 d + \left(\overline{y}_g a + \mathbf{x}_i'\mathbf{b} + c\right) k + v_g + u_i$$

One can readily check that this model can be rewritten as

$$y_i = \left(\overline{y}_g a + \mathbf{x}_i'\mathbf{b}\right)^2 d + (2cd + k)\left(\overline{y}_g a + \mathbf{x}_i'\mathbf{b}\right) + c^2 d + ck + v_g + u_i.$$

If $2cd + k \neq 0$ then this equation is identical to equation (17), replacing the fixed effect v_g with the fixed effect $\widetilde{v}_g = c^2 d + ck + v_g$, and replacing the constants a, $\mathbf{b}$, d, with constants $\widetilde{a}$, $\widetilde{\mathbf{b}}$, $\widetilde{d}$ defined by $\widetilde{a} = (2cd + k)\, a$, $\widetilde{\mathbf{b}} = (2cd + k)\,\mathbf{b}$, and $\widetilde{d} = d/\left(2cd + k\right)^2$. If $2cd + k = 0$, then by letting $\widetilde{v}_g = c^2 d + ck + v_g$ this equation becomes $y_i = \left(\overline{y}_g a + \mathbf{x}_i'\mathbf{b}\right)^2 d + \widetilde{v}_g + u_i$, which is the case we have already ruled out.

We next derive the equilibrium of $\overline{y}_g$. Expanding equation (17), we have

$$y_i = \overline{y}_g^2 d a^2 + a(2d\mathbf{x}_i'\mathbf{b} + 1)\overline{y}_g + \mathbf{b}'\mathbf{x}_i\mathbf{x}_i'\mathbf{b}d + \mathbf{x}_i'\mathbf{b} + v_g + u_i \tag{25}$$

Taking the within group expected value of this expression gives

$$\overline{y}_g = \overline{y}_g^2 d a^2 + a(2d\mathbf{b}'\overline{\mathbf{x}}_g + 1)\overline{y}_g + d\mathbf{b}'\overline{\mathbf{x}\mathbf{x}'}_g\mathbf{b} + \mathbf{b}'\overline{\mathbf{x}}_g + v_g. \tag{26}$$

so the equilibrium value of $\overline{y}_g$ must satisfy this equation for the model to be coherent. If $a = 0$, then we get $\overline{y}_g = d\mathbf{b}'\overline{\mathbf{x}\mathbf{x}'}_g\mathbf{b} + \mathbf{b}'\overline{\mathbf{x}}_g + v_g$ which exists and is unique. If $a \neq 0$, meaning that peer effects are present, then equation (26) is a quadratic with roots

$$\overline{y}_g = \frac{1 - a(2\mathbf{b}'\overline{\mathbf{x}}_g d + 1) \pm \sqrt{[1 - a(2\mathbf{b}'\overline{\mathbf{x}}_g d + 1)]^2 - 4a^2 d[d\mathbf{b}'\overline{\mathbf{x}\mathbf{x}'}_g\mathbf{b} + \mathbf{b}'\overline{\mathbf{x}}_g + v_g]}}{2a^2 d}.$$

The equilibrium of $\overline{y}_g$ therefore exists under Assumption A2 and is unique under Assumption A3. Note that regardless of whether $a = 0$ or not, $\overline{y}_g$ is always a function of $\overline{\mathbf{x}}_g$, $\overline{\mathbf{x}\mathbf{x}'}_g$, and v_g.

We now derive an expression for the measurement error $\varepsilon_{yg,-ii'}$. From equation (25), we have the group average

$$\widehat{y}_{g,-ii'} = \overline{y}_g^2 da^2 + a(2d\mathbf{b}'\widehat{\mathbf{x}}_{g,-ii'} + 1)\overline{y}_g + \mathbf{b}'\widehat{\mathbf{x}\mathbf{x}'}_{g,-ii'}\mathbf{b}d + \mathbf{b}'\widehat{\mathbf{x}}_{g,-ii'} + v_g + \widehat{u}_{g,-ii'}.$$

Subtracting equation (26) then gives the measurement error

$$\begin{aligned}
\varepsilon_{yg,-ii'} &= \widehat{y}_{g,-ii'} - \overline{y}_g = \frac{1}{n_g - 2} \sum_{l \neq i,i',l \in g} [2ad\overline{y}_g \mathbf{b}'(\mathbf{x}_l - \overline{\mathbf{x}}_g) + \mathbf{b}'(\mathbf{x}_l\mathbf{x}_l' - \overline{\mathbf{x}\mathbf{x}'_g})\mathbf{b}d + \mathbf{b}'(\mathbf{x}_l - \overline{\mathbf{x}}_g) + u_l] \\
&= 2ad\overline{y}_g \mathbf{b}'\varepsilon_{\mathbf{x}g,-ii'} + \mathbf{b}'\varepsilon_{\mathbf{x}\mathbf{x}g,-ii'}\mathbf{b}d + \mathbf{b}'\varepsilon_{\mathbf{x}g,-ii'} + \widehat{u}_{g,-ii'}.
\end{aligned}$$

Given the above results, we can now proceed with identification of the parameters. Substituting the above into the $y_i - y_{i'}$ gives

$$y_i - y_{i'} = 2ad\widehat{y}_{g,-ii'}\mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'}) + d\mathbf{b}'(\mathbf{x}_i\mathbf{x}_i' - \mathbf{x}_{i'}\mathbf{x}_{i'}')\mathbf{b} + \mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'}) + U_{ii'},$$

where

$$U_{ii'} = u_i - u_{i'} - 2ad(2ad\overline{y}_g\mathbf{b}'\boldsymbol{\varepsilon}_{\mathbf{x}g,-ii'} + \mathbf{b}'\boldsymbol{\varepsilon}_{\mathbf{x}\mathbf{x}g,-ii'}\mathbf{b}d + \mathbf{b}'\boldsymbol{\varepsilon}_{\mathbf{x}g,-ii'} + \widehat{u}_{g,-ii'})\mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'}).$$

Under Assumption A4, for each $i \in g$, $E\left((\mathbf{x}_i - \overline{\mathbf{x}}_g) \mid \overline{\mathbf{x}}_g, \overline{\mathbf{x}\mathbf{x}'}_g, v_g, \mathbf{r}_g\right) = 0$, and with its independence across individuals, we have

$$\begin{aligned}
& E\left(\overline{y}_g \boldsymbol{\varepsilon}_{\mathbf{x}g,-ii'}(\mathbf{x}_i - \mathbf{x}_{i'})' \mid \mathbf{r}_g, \mathbf{x}_i, \mathbf{x}_{i'}\right) \\
=\ & E\left(\overline{y}_g E(\boldsymbol{\varepsilon}_{\mathbf{x}g,-ii'} \mid \overline{\mathbf{x}}_g, \overline{\mathbf{x}\mathbf{x}'}_g, v_g, \mathbf{r}_g, \mathbf{x}_i, \mathbf{x}_{i'})(\mathbf{x}_i - \mathbf{x}_{i'})'\right) \\
=\ & E\left(\overline{y}_g E\left(\varepsilon_{\mathbf{x}g,-ii'} \mid \overline{\mathbf{x}}_g, \overline{\mathbf{x}\mathbf{x}'}_g, v_g, \mathbf{r}_g, \boldsymbol{\varepsilon}_{i\mathbf{x}g}, \boldsymbol{\varepsilon}_{i'\mathbf{x}g}\right)(\mathbf{x}_i - \mathbf{x}_{i'})'\right) = 0.
\end{aligned}$$

Together with $E\left(\varepsilon_{\mathbf{x}\mathbf{x}g,-ii'}(\mathbf{x}_i - \mathbf{x}_{i'}) \mid \mathbf{r}_g, \mathbf{x}_i, \mathbf{x}_{i'}\right) = 0$, $E\left(\varepsilon_{\mathbf{x}g,-ii'}(\mathbf{x}_i - \mathbf{x}_{i'}) \mid \mathbf{r}_g, \mathbf{x}_i, \mathbf{x}_{i'}\right) = 0$, and $E(\widehat{u}_{g,-ii'}(\mathbf{x}_i - \mathbf{x}_{i'})) = 0$, we have $E(U_{ii'}|\mathbf{r}_g, \mathbf{x}_i, \mathbf{x}_{i'}) = 0$ and hence,

$$E\left[y_i - y_{i'} - (2ad\widehat{y}_{g,-ii'}\mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'}) + d\mathbf{b}'(\mathbf{x}_i\mathbf{x}_i' - \mathbf{x}_{i'}\mathbf{x}_{i'}')\mathbf{b} + \mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'}))\,|\mathbf{r}_g, \mathbf{x}_i, \mathbf{x}_{i'}\right] = 0$$

For $\ell \in \{1, 2k, 3k, 4kk' \mid k, k' = 1, ..., K\}$, define vectors $\mathbf{Y}_{\ell g}$ as Section 4 and we have the group level moment condition

$$E\left(\mathbf{Y}_{1g} - \sum_{k=1}^{K} b_k\mathbf{Y}_{2kg} - 2ad\sum_{k=1}^{K} b_k\mathbf{Y}_{3kg} - d\sum_{k=1}^{K}\sum_{k'=1}^{K} b_k b_{k'}\mathbf{Y}_{4kk'g}\right) = 0. \tag{27}$$

Then, using the nonsingularity in Assumption A5, we have a, $\mathbf{b}$, d identified from

$$(\mathbf{b}', 2ad\mathbf{b}', db_1\mathbf{b}', \cdots, db_K\mathbf{b}')' = \left[E(\mathbf{Y}_g')E(\mathbf{Y}_g)\right]^{-1} \cdot E(\mathbf{Y}_g')E\left(\mathbf{Y}_{1g}\right),$$

where $\mathbf{Y}_g = \left(\mathbf{Y}_{21g}, ...\mathbf{Y}_{2Kg}, \mathbf{Y}_{31g}, ...\mathbf{Y}_{3Kg}, \mathbf{Y}_{411g}, \cdots, \mathbf{Y}_{4KKg}\right)$.

8.3 Multiple Equation Generic Model With Fixed Effects

Our actual demand application has a vector of J outcomes and a corresponding system of J equations. Extending the generic model to a multiple equation system introduces potential cross equation peer effects, resulting in more parameters to identify and estimate. Let $\mathbf{y}_i = (y_{1i}, ..., y_{Ji})$ be a J-dimensional outcome vector, where y_{ji} denotes the j'th outcome for individual i. Then we extend the single equation generic model to the multi equation that for each good j,

$$y_{ji} = (\overline{\mathbf{y}}_g'\mathbf{a}_j + \mathbf{x}_i'\mathbf{b}_j)^2 d_j + \left(\overline{\mathbf{y}}_g'\mathbf{a}_j + \mathbf{x}_i'\mathbf{b}_j\right) + v_{jg} + u_{ji}, \tag{28}$$

where $\overline{\mathbf{y}}_g = E(\mathbf{y}_i|i \in g)$ and $\mathbf{a}_j = (a_{1j}, ..., a_{Jj})'$ is the associated J-dimensional vector of peer effects for jth outcome (which in our application is the jth good). We now show that analogous derivations to the single equation model gives conditional moments

$$E\left((y_{ji} - y_{ji'} - 2d_j\widehat{\mathbf{y}}_{g,-ii'}'\mathbf{a}_j(\mathbf{x}_i - \mathbf{x}_{i'})'\mathbf{b}_j - d_j\mathbf{b}_j'(\mathbf{x}_i\mathbf{x}_i' - \mathbf{x}_{i'}\mathbf{x}_{i'}')\mathbf{b}_j - (\mathbf{x}_i - \mathbf{x}_{i'})'\mathbf{b}_j) \mid \mathbf{r}_g, \mathbf{x}_i, \mathbf{x}_i'\right) = 0.$$

Construction of unconditional moments for GMM estimation then follows exactly as before. The only difference is that now each outcome equation contains a vector of coefficients $\mathbf{a}_j$ instead of a single a. To maximize efficiency, the moments used for estimating each outcome equation can be combined into a single large GMM that estimates all of the parameters for all of the outcomes at the same time.

From

$$y_{ji} = d_j(\overline{\mathbf{y}}_g'\mathbf{a}_j)^2 + 2\overline{\mathbf{y}}_g'\mathbf{a}_j d_j \mathbf{x}_i'\mathbf{b}_j + \mathbf{b}_j'\mathbf{x}_i\mathbf{x}_i'\mathbf{b}_j d_j + \overline{\mathbf{y}}_g'\mathbf{a}_j + \mathbf{x}_i'\mathbf{b}_j + v_{jg} + u_{ji},$$

we have the equilibrium

$$\overline{y}_{jg} = d_j(\overline{\mathbf{y}}_g'\mathbf{a}_j)^2 + 2d_j\overline{\mathbf{y}}_g'\mathbf{a}_j\overline{\mathbf{x}}_g'\mathbf{b}_j + \mathbf{b}_j'\overline{\mathbf{x}\mathbf{x}'}_g\mathbf{b}_j d_j + \overline{\mathbf{y}}_g'\mathbf{a}_j + \overline{\mathbf{x}}_g'\mathbf{b}_j + v_{jg}$$

and the leave-two-out group average

$$\widehat{y}_{jg,-ii'} = d_j(\overline{\mathbf{y}}_g'\mathbf{a}_j)^2 + 2d_j\overline{\mathbf{y}}_g'\mathbf{a}_j\widehat{\mathbf{x}}_{g,-ii'}'\mathbf{b}_j + \mathbf{b}_j'\widehat{\mathbf{x}\mathbf{x}'}_{g,-i}\mathbf{b}_j d_j + \overline{\mathbf{y}}_g'\mathbf{a}_j + \widehat{\mathbf{x}}_{g,-ii'}'\mathbf{b}_j + v_{jg} + \widehat{u}_{jg,-ii'}.$$

Therefore, the measurement error is

$$\varepsilon_{yjg,-ii'} = \widehat{y}_{jg,-ii'} - \overline{y}_{jg} = 2d_j \overline{\mathbf{y}}_g' \mathbf{a}_j \boldsymbol{\varepsilon}_{xg,-ii'}' \mathbf{b}_j + \mathbf{b}_j' \boldsymbol{\varepsilon}_{xxg,-ii'} \mathbf{b}_j d_j + \boldsymbol{\varepsilon}_{xg,-ii'}' \mathbf{b}_j + \widehat{u}_{jg,-ii'}.$$

Using the same analysis as before,

$$\begin{aligned} y_{ji} - y_{ji'} &= 2d_j \overline{\mathbf{y}}_g' \mathbf{a}_j (\mathbf{x}_i - \mathbf{x}_{i'})' \mathbf{b}_j + d_j \mathbf{b}_j' (\mathbf{x}_i \mathbf{x}_i' - \mathbf{x}_{i'} \mathbf{x}_{i'}') \mathbf{b}_j + (\mathbf{x}_i - \mathbf{x}_{i'})' \mathbf{b}_j + u_{ji} - u_{ji'} \\ &= 2d_j \widehat{\mathbf{y}}_{g,-ii'}' \mathbf{a}_j (\mathbf{x}_i - \mathbf{x}_{i'})' \mathbf{b}_j + d_j \mathbf{b}_j' (\mathbf{x}_i \mathbf{x}_i' - \mathbf{x}_{i'} \mathbf{x}_{i'}') \mathbf{b}_j + (\mathbf{x}_i - \mathbf{x}_{i'})' \mathbf{b}_j + u_{ji} - u_{ji'} \\ &\quad -2d_j \boldsymbol{\varepsilon}_{yg,-ii'}' \mathbf{a}_j (\mathbf{x}_i - \mathbf{x}_{i'})' \mathbf{b}_j. \end{aligned}$$

Therefore, for $j = 1, ..., J$, we have the moment condition

$$E\left((y_{ji} - y_{ji'} - (\mathbf{x}_i - \mathbf{x}_{i'})' \mathbf{b}_j - 2d_j \widehat{\mathbf{y}}_{g,-ii'}' \mathbf{a}_j (\mathbf{x}_i - \mathbf{x}_{i'})' \mathbf{b}_j - d_j \mathbf{b}_j' (\mathbf{x}_i \mathbf{x}_i' - \mathbf{x}_{i'} \mathbf{x}_{i'}') \mathbf{b}_j) | \mathbf{r}_{gii'} \right) = 0.$$

Denote

$$\begin{aligned} L_{1jgii'} &= (y_{ji} - y_{ji'}), \quad L_{2kgii'} = (x_{ki} - x_{ki'}), \\ L_{3jkgii'} &= \widehat{y}_{jg,-ii'} (x_{ki} - x_{ki'}), \quad L_{4kk'gii'} = x_{ki} x_{k'i} - x_{ki'} x_{k'i'}. \end{aligned}$$

For $\ell \in \{1j, 2k, 3jk, 4kk' \mid j = 1, ..., J; k, k' = 1, ..., K\}$, define vectors

$$\mathbf{Y}_{\ell g} = \frac{\sum_{(i,i') \in \Gamma_g} L_{\ell gii'} \mathbf{r}_{gii'}}{\sum_{(i,i') \in \Gamma_g} 1}$$

and the identification comes from the group level unconditional moment equation

$$E\left(\mathbf{Y}_{1jg} - \sum_{k=1}^{K} b_{jk} \mathbf{Y}_{2kg} - 2d_j \sum_{j'=1}^{J} \sum_{k=1}^{K} a_{jj'} b_{jk} \mathbf{Y}_{3j'kg} - d_j \sum_{k=1}^{K} \sum_{k'=1}^{K} b_{jk} b_{jk'} \mathbf{Y}_{4kk'g} \right) = 0,$$

where b_{jk} is the kth element of $\mathbf{b}_j$ and $a_{jj'}$ is the j'th element of $\mathbf{a}_j$.

Let $\mathbf{Y}_g = (\mathbf{Y}_{21g}, ...\mathbf{Y}_{2Kg}, \mathbf{Y}_{311g}, \mathbf{Y}_{312g}, ...\mathbf{Y}_{3JKg}, \mathbf{Y}_{411g}, \cdots, \mathbf{Y}_{4KKg})$. If $E(\mathbf{Y}_g)' E(\mathbf{Y}_g)$ is nonsingular, for each $j = 1, ..., J$, we can identify

$$(\mathbf{b}_j', 2a_{j1} d_j \mathbf{b}_j', ..., 2a_{jJ} d_j \mathbf{b}_j', d_j b_{j1} \mathbf{b}_j', ..., d_j b_{jK} \mathbf{b}_j')' = \left[E(\mathbf{Y}_g)' E(\mathbf{Y}_g) \right]^{-1} \cdot E(\mathbf{Y}_g)' E(\mathbf{Y}_{1jg}).$$

From this, $\mathbf{b}_j$, d_j, and $\mathbf{a}_j$ can be identified for each $j = 1, ..., J$.

For a single large GMM that estimates all of the parameters for all of the outcomes at the same time, we construct the group level GMM estimation based on

$$\left(\widehat{\mathbf{a}}_1', ..., \widehat{\mathbf{a}}_J', \widehat{\mathbf{b}}_1', ...\widehat{\mathbf{b}}_J', \widehat{d}_1, ..., \widehat{d}_J\right)' = \arg\min \left(\frac{1}{G}\sum_{g=1}^{G}\mathbf{m}_g\right)' \widehat{\Omega} \left(\frac{1}{G}\sum_{g=1}^{G}\mathbf{m}_g\right),$$

where $\widehat{\Omega}$ is some positive definite moment weighting matrix and

$$\mathbf{m}_g = \begin{pmatrix} \mathbf{Y}_{11g} \\ \vdots \\ \mathbf{Y}_{1Jg} \end{pmatrix} - \begin{pmatrix} \sum_{k=1}^{K} b_{1k}\mathbf{Y}_{2kg} \\ \vdots \\ \sum_{k=1}^{K} b_{Jk}\mathbf{Y}_{2kg} \end{pmatrix} - 2\begin{pmatrix} d_1 \sum_{j'=1}^{J}\sum_{k=1}^{K} a_{1j'}b_{1k}\mathbf{Y}_{3j'kg} \\ \vdots \\ d_J \sum_{j'=1}^{J}\sum_{k=1}^{K} a_{Jj'}b_{Jk}\mathbf{Y}_{3j'kg} \end{pmatrix} - \begin{pmatrix} d_1 \sum_{k=1}^{K}\sum_{k'=1}^{K} b_{1k}b_{1k'}\mathbf{Y}_{4kk'g} \\ \vdots \\ d_J \sum_{k=1}^{K}\sum_{k'=1}^{K} b_{Jk}b_{Jk'}\mathbf{Y}_{4kk'g} \end{pmatrix}$$

is a $qJ-$dimensional vector.

Alternatively, we can construct the individual level GMM estimation using the group clustered standard errors

$$\left(\widehat{\mathbf{a}}_1', ..., \widehat{\mathbf{a}}_J', \widehat{\mathbf{b}}_1', ...\widehat{\mathbf{b}}_J', \widehat{d}_1, ..., \widehat{d}_J\right)' = \arg\min \left(\frac{\sum_{g=1}^{G}\sum_{(i,i')\in\Gamma_g}\mathbf{m}_{gii'}}{\sum_{g=1}^{G}\sum_{(i,i')\in\Gamma_g}1}\right)' \widehat{\Omega} \left(\frac{\sum_{g=1}^{G}\sum_{(i,i')\in\Gamma_g}\mathbf{m}_{gii'}}{\sum_{g=1}^{G}\sum_{(i,i')\in\Gamma_g}1}\right),$$

where

$$\begin{aligned}\mathbf{m}_{gii'} \quad = \quad & \begin{pmatrix} L_{11gii'}\mathbf{r}_{gii'} \\ \vdots \\ L_{1Jgii'}\mathbf{r}_{gii'} \end{pmatrix} - \begin{pmatrix} \sum_{k=1}^{K} b_{1k}L_{2kgii'}\mathbf{r}_{gii'} \\ \vdots \\ \sum_{k=1}^{K} b_{Jk}L_{2kgii'}\mathbf{r}_{gii'} \end{pmatrix} - 2\begin{pmatrix} d_1 \sum_{j'=1}^{J}\sum_{k=1}^{K} a_{1j'}b_{1k}L_{3j'gii'}\mathbf{r}_{gii'} \\ \vdots \\ d_J \sum_{j'=1}^{J}\sum_{k=1}^{K} a_{Jj'}b_{Jk}L_{3j'gii'}\mathbf{r}_{gii'} \end{pmatrix} \\ & - \begin{pmatrix} d_1 \sum_{k=1}^{K}\sum_{k'=1}^{K} b_{1k}b_{1k'}L_{4kk'gii'}\mathbf{r}_{gii'} \\ \vdots \\ d_J \sum_{k=1}^{K}\sum_{k'=1}^{K} b_{Jk}b_{Jk'}L_{4kk'gii'}\mathbf{r}_{gii'} \end{pmatrix}.\end{aligned}$$

8.4 Multiple Equation Generic Model With Random Effects

Here we provide the derivation of equation (7), thereby showing validity of the moments used for random effects estimation. As with fixed effects, we here extend the model to allow a vector of covariates $\mathbf{x}_i$. We begin by rewriting the generic model with vector $\mathbf{x}_i$, equation (17).

$$y_i = \overline{y}_g^2 a^2 d + a\left(1 + 2\mathbf{b}'\mathbf{x}_i d\right)\overline{y}_g + \mathbf{b}'\mathbf{x}_i + \mathbf{b}'\mathbf{x}_i\mathbf{x}_i'\mathbf{b}d + v_g + u_i, \tag{29}$$

We now add the assumption that v_g is independent of $\mathbf{x}$ and u, making it a random effect. Taking the expectation of this expression given being in group g gives

$$\overline{y}_g = \overline{y}_g^2 da^2 + a(2d\mathbf{b}'\overline{\mathbf{x}}_g + 1)\overline{y}_g + d\mathbf{b}'\overline{\mathbf{x}\mathbf{x}'_g}\mathbf{b} + \mathbf{b}'\overline{\mathbf{x}}_g + \mu, \tag{30}$$

where $\mu = E(v_g)$. Hence, the group mean $\overline{y}_g$ is an implicit function of $\overline{\mathbf{x}}_g$ and $\overline{\mathbf{x}\mathbf{x}'_g}$.

Define measurement errors $\varepsilon_{\mathbf{x}l} = \mathbf{x}_l - \overline{\mathbf{x}}_g$, $\varepsilon_{\mathbf{xx}l} = \mathbf{x}_l\mathbf{x}'_l - \overline{\mathbf{x}\mathbf{x}'_g}$, and $\varepsilon_{yg,-ii'} = \widehat{y}_{g,-ii'} - \overline{y}_g$.

For any $i' \in g$, the measurement error $\varepsilon_{yi'} = y_{i'} - \overline{y}_g$ is

$$\begin{aligned} \varepsilon_{yi'} &= 2ad\overline{y}_g\mathbf{b}'(\mathbf{x}_{i'} - \overline{\mathbf{x}}_g) + d\mathbf{b}'(\mathbf{x}_{i'}\mathbf{x}'_{i'} - \overline{\mathbf{x}\mathbf{x}'_g})\mathbf{b} + \mathbf{b}'(\mathbf{x}_{i'} - \overline{\mathbf{x}}_g) + u_{i'} + v_g \\ &= 2ad\overline{y}_g\mathbf{b}'\varepsilon_{\mathbf{x}i'} + d\mathbf{b}'\varepsilon_{\mathbf{xx}i'}\mathbf{b} + \mathbf{b}'\varepsilon_{\mathbf{x}i'} + u_{i'} + v_g - \mu. \end{aligned}$$

and so the measurement error $\varepsilon_{yg,-ii'} = \widehat{y}_{g,-ii'} - \overline{y}_g$ is

$$\varepsilon_{yg,-ii'} = \widehat{y}_{g,-ii'} - \overline{y}_g = 2ad\overline{y}_g\mathbf{b}'\varepsilon_{\mathbf{x}g,-ii'} + \mathbf{b}'\varepsilon_{\mathbf{xx}g,-ii'}\mathbf{b}d + \mathbf{b}'\varepsilon_{\mathbf{x}g,-ii'} + \widehat{u}_{g,-ii'} + v_g - \mu.$$

Next define $\widetilde{\varepsilon}_{gii'}$ by

$$\widetilde{\varepsilon}_{gii'} = \left(\overline{y}_g^2 - \widehat{y}_{g,-ii'}y_{i'}\right)a^2 d + a\left(1 + 2\mathbf{b}'\mathbf{x}_i d\right)\left(\overline{y}_g - \widehat{y}_{g,-ii'}\right),$$

so

$$y_i = \widehat{y}_{g,-ii'}y_{i'}a^2 d + a\left(1 + 2\mathbf{b}'\mathbf{x}_i d\right)\widehat{y}_{g,-ii'} + \mathbf{b}'\mathbf{x}_i + \mathbf{b}'\mathbf{x}_i\mathbf{x}'_i\mathbf{b}d + v_g + u_i + \widetilde{\varepsilon}_{gii'}. \tag{31}$$

Then

$$\begin{aligned} \widetilde{\varepsilon}_{gii'} &= \left(\overline{y}_g^2 - (\overline{y}_g + \varepsilon_{yg,-ii'})(\overline{y}_g + \varepsilon_{yi'})\right)a^2 d - a\left(1 + 2\mathbf{b}'\mathbf{x}_i d\right)\varepsilon_{yg,-ii'} \\ &= -(\varepsilon_{yg,-ii'} + \varepsilon_{y,i'})\overline{y}_g a^2 d - \varepsilon_{yg,-ii'}\varepsilon_{y,i'}a^2 d - a\left(1 + 2\mathbf{b}'\mathbf{x}_i d\right)\varepsilon_{yg,-ii'}. \end{aligned}$$

Make the following assumptions.

Assumption C1: For any individual l, v_g is independent of $(\mathbf{x}_l, \overline{\mathbf{x}}_g, \overline{\mathbf{x}\mathbf{x}'_g})$, the error term u_l, and measurement errors $\varepsilon_{\mathbf{x}l}$ and $\varepsilon_{\mathbf{xx}l}$.

Assumption C2: For each individual l in group g, conditional on $(\overline{\mathbf{x}}_g, \overline{\mathbf{x}\mathbf{x}'_g})$ the measurement errors $\varepsilon_{\mathbf{x}l}$ and $\varepsilon_{\mathbf{xx}l}$ are independent across individuals and have zero means.

Assumption C3: For each group g, v_g is independent across groups with $E(v_g|\mathbf{x}, \overline{\mathbf{x}}_g, \overline{\mathbf{x}\mathbf{x}'_g}) = \mu$ and we have the conditional homoskedasticity that $Var(v_g|\mathbf{x}, \overline{\mathbf{x}}_g, \overline{\mathbf{x}\mathbf{x}'_g}) = \sigma^2$.

Let $v_0 = \mu - da^2\sigma^2$. It follows from these assumptions that, for any $l \neq i$, $E(\overline{y}_g\varepsilon_{yl}|\mathbf{x}_i, \overline{\mathbf{x}}_g, \overline{\mathbf{x}\mathbf{x}'_g}) =$

0 and $E(\varepsilon_{yl}\mathbf{x}_i|\mathbf{x}_i, \overline{\mathbf{x}}_g, \overline{\mathbf{xx}'_g}) = 0$. Hence, $E(\widetilde{\varepsilon}_{gii'}|x_i, \overline{\mathbf{x}}_g, \overline{\mathbf{xx}'_g}) = -da^2 E(\varepsilon_{yg,-ii'}\varepsilon_{y,i'}|\mathbf{x}_i, \overline{\mathbf{x}}_g, \overline{\mathbf{xx}'_g}) = -da^2 Var\,(v_g)$ and

$$E(v_g + u_i + \widetilde{\varepsilon}_{gii'} \mid \overline{\mathbf{x}}_g, \overline{\mathbf{xx}'_g}, \mathbf{x}_i) = \mu - da^2\sigma^2 = v_0. \tag{32}$$

By construction $v_g + u_i + \widetilde{\varepsilon}_{gii'}$ is also independent of $\mathbf{r}_g$. Given this, equation (7) then follows from equations (31) (32).

8.5 Identification of the Demand System With Fixed Effects

We begin by considering the Engel curve model, without price variation. As

$$\begin{aligned}
\mathbf{q}_i &= x_i^2\mathbf{m} + \mathbf{m}\left(\boldsymbol{\alpha}'\overline{\mathbf{q}}_g\right)^2 + \left(\boldsymbol{\gamma}'\mathbf{z}_i\mathbf{z}_i'\boldsymbol{\gamma}\right)\mathbf{m} - 2\mathbf{m}\boldsymbol{\alpha}'\overline{\mathbf{q}}_g x_i - 2\mathbf{m}\boldsymbol{\gamma}'\mathbf{z}_i x_i + 2\mathbf{m}\boldsymbol{\alpha}'\overline{\mathbf{q}}_g\boldsymbol{\gamma}'\mathbf{z}_i \\
&\quad + \left(x_i - \boldsymbol{\alpha}'\overline{\mathbf{q}}_g - \boldsymbol{\gamma}'\mathbf{z}_i\right)\boldsymbol{\delta} + \mathbf{A}\overline{\mathbf{q}}_{\mathbf{g}} + \mathbf{C}\mathbf{z}_i + \mathbf{v}_g + \mathbf{u}_i,
\end{aligned}$$

we have

$$\begin{aligned}
\overline{\mathbf{q}}_g &= \overline{x_g^2}\mathbf{m} + \mathbf{m}\left(\boldsymbol{\alpha}'\overline{\mathbf{q}}_g\right)^2 + \left(\boldsymbol{\gamma}'\overline{\mathbf{zz}'}_g\boldsymbol{\gamma}\right)\mathbf{m} - 2\mathbf{m}\boldsymbol{\alpha}'\overline{\mathbf{q}}_g\overline{x}_g - 2\mathbf{m}\boldsymbol{\gamma}'\overline{x\mathbf{z}}_g + 2\mathbf{m}\boldsymbol{\alpha}'\overline{\mathbf{q}}_g\boldsymbol{\gamma}'\overline{\mathbf{z}}_g \\
&\quad + \left(\overline{x}_g - \boldsymbol{\alpha}'\overline{\mathbf{q}}_g - \boldsymbol{\gamma}'\overline{\mathbf{z}}_g\right)\boldsymbol{\delta} + \mathbf{A}\overline{\mathbf{q}}_{\mathbf{g}} + \mathbf{C}\overline{\mathbf{z}}_g + \mathbf{v}_g; \\
\widehat{\mathbf{q}}_{g,-ii'} &= \widehat{x^2}_{g,-ii'}\mathbf{m} + \mathbf{m}\left(\boldsymbol{\alpha}'\overline{\mathbf{q}}_g\right)^2 + \left(\boldsymbol{\gamma}'\widehat{\mathbf{zz}'}_{g,-ii'}\boldsymbol{\gamma}\right)\mathbf{m} - 2\mathbf{m}\boldsymbol{\alpha}'\overline{\mathbf{q}}_g\widehat{x}_{g,-ii'} - 2\mathbf{m}\boldsymbol{\gamma}'\widehat{\mathbf{z}x}_{g,-ii'} \\
&\quad + 2\mathbf{m}\boldsymbol{\alpha}'\overline{\mathbf{q}}_g\boldsymbol{\gamma}'\widehat{\mathbf{z}}_{g,-ii'} + \left(\widehat{x}_{g,-ii'} - \boldsymbol{\alpha}'\overline{\mathbf{q}}_g - \boldsymbol{\gamma}'\widehat{\mathbf{z}}_{g,-ii'}\right)\boldsymbol{\delta} + \mathbf{A}\overline{\mathbf{q}}_{\mathbf{g}} + \mathbf{C}\widehat{\mathbf{z}}_{g,-ii'} + \mathbf{v}_g + \widehat{\mathbf{u}}_{g,-ii'}.
\end{aligned}$$

Hence,

$$\begin{aligned}
\boldsymbol{\varepsilon}_{qg,-ii'} &= \widehat{\mathbf{q}}_{g,-ii'} - \overline{\mathbf{q}}_g = -2\mathbf{m}\boldsymbol{\alpha}'\overline{\mathbf{q}}_g(\varepsilon_{xg,-ii'} - \boldsymbol{\gamma}'\boldsymbol{\varepsilon}_{zg,-ii'}) + \varepsilon_{x^2g,-ii'}\mathbf{m} + \boldsymbol{\gamma}'\boldsymbol{\varepsilon}_{zzg,-ii'}\boldsymbol{\gamma}\mathbf{m} \\
&\quad -2\mathbf{m}\boldsymbol{\gamma}'\boldsymbol{\varepsilon}_{zxg,-ii'} + \boldsymbol{\delta}\varepsilon_{xg,-ii'} + (\mathbf{C} - \boldsymbol{\delta}\boldsymbol{\gamma}')\boldsymbol{\varepsilon}_{zg,-ii'} + \widehat{\mathbf{u}}_{g,-ii'}.
\end{aligned}$$

The pairwise difference gives

$$\begin{aligned}
\mathbf{q}_i - \mathbf{q}_{i'} &= -2\mathbf{m}\boldsymbol{\alpha}'\overline{\mathbf{q}}_g[(x_i - x_{i'}) - \boldsymbol{\gamma}'(\mathbf{z}_i - \mathbf{z}_{i'})] + (x_i^2 - x_{i'}^2)\mathbf{m} + [\boldsymbol{\gamma}'\left(\mathbf{z}_i\mathbf{z}_i' - \mathbf{z}_{i'}\mathbf{z}_{i'}'\right)\boldsymbol{\gamma}]\mathbf{m} \\
&\quad -2\mathbf{m}\boldsymbol{\gamma}'(\mathbf{z}_i x_i - \mathbf{z}_{i'}x_{i'}) + \boldsymbol{\delta}(x_i - x_{i'}) + (\mathbf{C} - \boldsymbol{\delta}\boldsymbol{\gamma}')(\mathbf{z}_i - \mathbf{z}_{i'}) + \mathbf{u}_i - \mathbf{u}_{i'} \\
&= -2\mathbf{m}\boldsymbol{\alpha}'\widehat{\mathbf{q}}_{g,-ii'}[(x_i - x_{i'}) - \boldsymbol{\gamma}'(\mathbf{z}_i - \mathbf{z}_{i'})] + (x_i^2 - x_{i'}^2)\mathbf{m} + [\boldsymbol{\gamma}'\left(\mathbf{z}_i\mathbf{z}_i' - \mathbf{z}_{i'}\mathbf{z}_{i'}'\right)\boldsymbol{\gamma}]\mathbf{m} \\
&\quad -2\mathbf{m}\boldsymbol{\gamma}'(\mathbf{z}_i x_i - \mathbf{z}_{i'}x_{i'}) + \boldsymbol{\delta}(x_i - x_{i'}) + (\mathbf{C} - \boldsymbol{\delta}\boldsymbol{\gamma}')(\mathbf{z}_i - \mathbf{z}_{i'}) + \mathbf{U}_{ii'},
\end{aligned}$$

where the composite error is

$$\mathbf{U}_{ii'} = \mathbf{u}_i - \mathbf{u}_{i'} + 2\mathbf{m}\boldsymbol{\alpha}'\boldsymbol{\varepsilon}_{qg,-ii'}[(x_i - x_{i'}) - \boldsymbol{\gamma}'(\mathbf{z}_i - \mathbf{z}_{i'})].$$

For identification and estimation, we need following assumptions.

Assumption B1: Each individual i in group g satisfies equation (12). Unobserved errors $\mathbf{u}_i$'s are independent across groups and have zero mean conditional on all $(x_l, \mathbf{z}_l)$ for $l \in g$, and $\mathbf{v}_g$ are unobserved group level fixed effects. The number of observed groups $G \to \infty$. For each observed group g, a sample of n_g observations of $\mathbf{q}_i, x_i, \mathbf{z}_i$ is observed. Each sample size n_g is fixed and does not go to infinity. The true number of individuals comprising each group is unknown and could be finite.

Assumption B2: The coefficients $\mathbf{A}, \mathbf{C}, \mathbf{b}, \mathbf{d}$ are unknown constants satisfying $\mathbf{b}'\mathbf{1} = 1$, $\mathbf{d}'\mathbf{1} = 0$, $\mathbf{d} \neq 0$, and there exist solutions of $\overline{\mathbf{q}}_g$ such that

$$\begin{aligned}
\overline{\mathbf{q}}_g &= \overline{x_g^2}\mathbf{m} + \mathbf{m}\left(\boldsymbol{\alpha}'\overline{\mathbf{q}}_g\right)^2 + \left(\boldsymbol{\gamma}'\overline{\mathbf{z}\mathbf{z}'}_g\boldsymbol{\gamma}\right)\mathbf{m} - 2\mathbf{m}\boldsymbol{\alpha}'\overline{\mathbf{q}}_g\overline{x}_g - 2\mathbf{m}\boldsymbol{\gamma}'\overline{x\mathbf{z}}_g + 2\mathbf{m}\boldsymbol{\alpha}'\overline{\mathbf{q}}_g\boldsymbol{\gamma}'\overline{\mathbf{z}}_g \\
&\quad + \left(\overline{x}_g - \boldsymbol{\alpha}'\overline{\mathbf{q}}_g - \boldsymbol{\gamma}'\overline{\mathbf{z}}_g\right)\boldsymbol{\delta} + \mathbf{A}\overline{\mathbf{q}}_g + \mathbf{C}\overline{\mathbf{z}}_g + \mathbf{v}_g. \qquad (33)
\end{aligned}$$

Assumption B1 just defines the model. In Assumption B2, if q is a scalar, then the solution exists when

$$\begin{aligned}
&\left(1 + 2mAp\overline{x}_g - 2mAp\boldsymbol{\gamma}'\overline{\mathbf{z}}_g - A + pA\delta\right)^2 \\
&-4m\left(Ap\right)^2\left(m\overline{x_g^2} + m\boldsymbol{\gamma}'\overline{\mathbf{z}\mathbf{z}'}_g\boldsymbol{\gamma} - 2m\boldsymbol{\gamma}'\overline{x\mathbf{z}}_g + \overline{x}_g\delta - \delta\boldsymbol{\gamma}'\overline{\mathbf{z}}_g + C\overline{\mathbf{z}}_g + v_g\right) \geq 0
\end{aligned}$$

as $m \neq 0(d \neq 0)$ is needed to avoid the reflection problem. Assumption B2 ensures that an equilibrium exists for each group, thereby avoiding Tamer's (2003) potential incoherence problem. To see this, if $A \neq 0$, $\overline{\mathbf{q}}_g$ has the solution

$$\begin{aligned}
\overline{q}_g &= \frac{1}{2m\left(Ap\right)^2}\left(1 + 2mAp\overline{x}_g - 2mAp\boldsymbol{\gamma}'\overline{\mathbf{z}}_g - A + pA\delta\right) \pm \left[\left(1 + 2mAp\overline{x}_g - 2mAp\boldsymbol{\gamma}'\overline{\mathbf{z}}_g - A + pA\delta\right)^2\right. \\
&\quad \left. -4m\left(Ap\right)^2\left(m\overline{x_g^2} + m\boldsymbol{\gamma}'\overline{\mathbf{z}\mathbf{z}'}_g\boldsymbol{\gamma} - 2m\boldsymbol{\gamma}'\overline{x\mathbf{z}}_g + \overline{x}_g\delta - \delta\boldsymbol{\gamma}'\overline{\mathbf{z}}_g + C\overline{\mathbf{z}}_g + v_g\right)\right]^{1/2}, \qquad (34)
\end{aligned}$$

while if A does equal zero, then the model will be trivially identified because in that case there aren't any peer effects. From equation (34), we can see $\overline{\mathbf{q}}_g$ is an implicit function of $\overline{x_g^2}, \overline{x}_g, \overline{\mathbf{z}}_g, \overline{\mathbf{z}\mathbf{z}'}_g, \overline{x\mathbf{z}}_g$, and $\mathbf{v}_g$. In the case of multiple equilibria, we do not take a stand on which root of equation (33) is chosen by consumers, we just make the following assumption.

Assumption B3: Individuals within each group agree on an equilibrium selection rule.

Assumption B4: Within each group g, the vector $(x_i, \mathbf{z}_i)$ is a random sample drawn from a distribution that has mean $(\overline{x}_g, \overline{\mathbf{z}}_g) = E\left((x_i, \mathbf{z}_i) \mid i \in g\right)$ and variance $\Sigma_{x\mathbf{z}g} = \begin{pmatrix} \sigma_{xg}^2 & \sigma_{x\mathbf{z}g} \\ \sigma_{x\mathbf{z}g}' & \Sigma_{\mathbf{z}g} \end{pmatrix}$ where $\sigma_{xg}^2 = Var(x_i \mid i \in g)$, $\sigma_{x\mathbf{z}g} = Cov(x_i, \mathbf{z}_i \mid i \in g)$ and $\Sigma_{\mathbf{z}g} = Var(\mathbf{z}_i \mid i \in g)$. Denote

$\varepsilon_{ix} = x_i - \overline{x}_g$ and $\boldsymbol{\varepsilon}_{iz} = \mathbf{z}_i - \overline{\mathbf{z}}_g$. Assume $E\left((\varepsilon_{ix}, \boldsymbol{\varepsilon}_{iz})|\overline{\mathbf{z}}_g, \overline{x\mathbf{z}}_g, \overline{\mathbf{z}\mathbf{z}'}_g, \overline{x}_g, \overline{x^2_g}, \mathbf{v}_g, \mathbf{r}_g\right) = 0$ and is independent across individual i's.

To satisfy Assumption B4, we can think of group level variables like $\overline{x}_g$, $\overline{\mathbf{z}}_g$ and $\mathbf{v}_g$ as first being drawn from some distribution, and then separately drawing the individual level variables $(\varepsilon_{ix}, \boldsymbol{\varepsilon}_{iz})$ from some distribution that is unrelated from the group level distribution, to then determine the individual level observables $x_i = \overline{x}_g + \varepsilon_{ix}$ and $\mathbf{z}_i = \overline{\mathbf{z}}_g + \boldsymbol{\varepsilon}_{iz}$. It then follows from Assumption B4 that $E(\varepsilon_{xg,-ii'} \mid x_i, \mathbf{z}_i, x_{i'}, \mathbf{z}_{i'}, \mathbf{r}_g) = 0$ and $E(\boldsymbol{\varepsilon}_{zg,-ii'} \mid x_i, \mathbf{z}_i, x_{i'}, \mathbf{z}_{i'}, \mathbf{r}_g) = 0$. With similar arguments in the generic model, Assumption B4 suffices to ensure that

$$E(\boldsymbol{\varepsilon}_{qg,-ii'}[(x_i - x_{i'}), (\mathbf{z}_i - \mathbf{z}_{i'})']|x_i, x_{i'}, \mathbf{z}_i, \mathbf{z}_{i'}, \mathbf{r}_g) = E(\boldsymbol{\varepsilon}_{qg,-ii'}|\mathbf{r}_g) \cdot [(x_i - x_{i'}), (\mathbf{z}_i - \mathbf{z}_{i'})'] = 0.$$

Then we have the moment condition

$$E\{[\mathbf{q}_i - \mathbf{q}_{i'} + 2\mathbf{m}\boldsymbol{\alpha}'\widehat{\mathbf{q}}_{g,-ii'}(x_i - x_{i'}) - 2\mathbf{m}\boldsymbol{\alpha}'\widehat{\mathbf{q}}_{g,-ii'}\boldsymbol{\gamma}'(\mathbf{z}_i - \mathbf{z}_{i'}) - (x_i^2 - x_{i'}^2)\mathbf{m} - \boldsymbol{\gamma}'\left(\mathbf{z}_i\mathbf{z}_i' - \mathbf{z}_{i'}\mathbf{z}_{i'}'\right)\boldsymbol{\gamma}\mathbf{m} \tag{35}$$
$$+2\mathbf{m}\boldsymbol{\gamma}'(\mathbf{z}_i x_i - \mathbf{z}_{i'} x_{i'}) - \boldsymbol{\delta}(x_i - x_{i'}) + (\boldsymbol{\delta}\boldsymbol{\gamma}' - \mathbf{C})(\mathbf{z}_i - \mathbf{z}_{i'})]|x_i, x_{i'}, \mathbf{z}_i, \mathbf{z}_{i'}, \mathbf{r}_g\} = 0$$

for the Engel curve and

$$\begin{aligned} E\Bigg[\Bigg(&\mathbf{q}_i - \mathbf{q}_{i'} - \frac{\mathbf{b}}{\mathbf{p}_t}(x_i - x_{i'}) + (\frac{\mathbf{b}}{\mathbf{p}_t}\mathbf{C}'\mathbf{p}_t - \mathbf{C}_j)(\mathbf{z}_i - \mathbf{z}_{i'}) \\ &+e^{-\mathbf{b}'\ln\mathbf{p}_t}\frac{\mathbf{d}}{\mathbf{p}_t}[2\mathbf{p}_t'\mathbf{A}\widehat{\mathbf{q}}_{gt,-ii'}(x_i - x_{i'}) - 2\mathbf{p}_t'\mathbf{A}\widehat{\mathbf{q}}_{gt,-ii'}\mathbf{p}_t'\mathbf{C}(\mathbf{z}_i - \mathbf{z}_{i'}) - (x_i^2 - x_{i'}^2)] \\ &+e^{-\mathbf{b}'\ln\mathbf{p}_t}\frac{\mathbf{d}}{\mathbf{p}_t}[2\mathbf{p}_t'\mathbf{C}(\mathbf{z}_i x_i - \mathbf{z}_{i'} x_{i'}) - \mathbf{p}_t'\mathbf{C}\left(\mathbf{z}_i\mathbf{z}_i' - \mathbf{z}_{i'}\mathbf{z}_{i'}'\right)\mathbf{C}'\mathbf{p}_t]\Bigg)|x_i, x_{i'}, \mathbf{z}_i, \mathbf{z}_{i'}, \mathbf{r}_g\Bigg] = 0. \end{aligned} \tag{36}$$

for the full demand curve.

Let the instrument vector $\mathbf{r}_{gii'}$ be any functional form of $\mathbf{r}_g$, $(x_i, \mathbf{z}_i')'$, and $(x_{i'}, \mathbf{z}_{i'}')'$. Denote

$$\begin{aligned} L_{1jgii'} &= (q_{ji} - q_{ji'}),\ \ L_{2jgii'} = \widehat{q}_{jg,-ii'}(x_i - x_{i'}),\ \ L_{3jkgii'} = \widehat{q}_{jg,-ii'}(z_{ki} - z_{ki'}),\ \ L_{4gii'} = x_i^2 - x_{i'}^2, \\ L_{5kk'gii'} &= z_{ki}z_{k'i} - z_{ki'}z_{k'i'},\ \ L_{6kgii'} = z_{ki}x_i - z_{ki'}x_{i'},\ \ L_{7gii'} = x_i - x_{i'},\ \ L_{8kgii'} = z_{ki} - z_{ki'}. \end{aligned}$$

For $\ell \in \{1j, 2j, 3jk, 4, 5kk', 6k, 7, 8k \mid j = 1, ..., J; k, k' = 1, ..., K\}$, define vectors

$$\mathbf{Q}_{\ell g} = \frac{\sum_{(i,i')\in\Gamma_g} L_{\ell gii'}\mathbf{r}_{gii'}}{\sum_{(i,i')\in\Gamma_g} 1}.$$

Then for each good j, the identification is based on

$$\begin{aligned} 0 &= E\Bigg(\mathbf{Q}_{1jg} + 2m_j \sum_{j'=1}^{J} \alpha_{j'}\mathbf{Q}_{2j'g} - 2m_j \sum_{j'=1}^{J}\sum_{k=1}^{K} \alpha_{j'}\gamma_k \mathbf{Q}_{3j'kg} - m_j\mathbf{Q}_{4g} - m_j \sum_{k=1}^{K}\sum_{k'=1}^{K} \gamma_k\gamma_{k'}\mathbf{Q}_{5gkk'} \\ &\quad +2m_j \sum_{k=1}^{K} \gamma_k \mathbf{Q}_{6kg} - \delta_j \mathbf{Q}_{7g} + \sum_{k=1}^{K} (\delta_j\gamma_k - c_{jk})\mathbf{Q}_{8kg}\Bigg). \end{aligned}$$

where γ_k is the kth element of $\boldsymbol{\gamma} = \mathbf{C}'\mathbf{p}$ and c_{jk} is the jth element of $\mathbf{c}_j$.

Assumption B5: $E\left(\mathbf{Q}_g'\right) E\left(\mathbf{Q}_g\right)$ is nonsingular, where

$$\mathbf{Q}_g = \left(\mathbf{Q}_{21g}, ..., \mathbf{Q}_{2Jg}, \mathbf{Q}_{311g}, ..., \mathbf{Q}_{3JKg}, \mathbf{Q}_{4g}, \mathbf{Q}_{511g}, ..., \mathbf{Q}_{5KKg}, \mathbf{Q}_{61g}, ..., \mathbf{Q}_{6Kg}, \mathbf{Q}_{7g}, \mathbf{Q}_{81g}, ..., \mathbf{Q}_{8Kg}\right).$$

Under Assumption B5, we can identify

$$\begin{aligned} &(-2m_j\boldsymbol{\alpha}', 2m_j\alpha_1\boldsymbol{\gamma}', ..., 2m_j\alpha_J\boldsymbol{\gamma}', m_j, m_j\gamma_1\boldsymbol{\gamma}', m_j\gamma_K\boldsymbol{\gamma}', -2m_j\boldsymbol{\gamma}', \delta_j, \mathbf{c}_j' - \delta_j\boldsymbol{\gamma}')' \\ = &\left[E\left(\mathbf{Q}_g'\right) E\left(\mathbf{Q}_g\right)\right]^{-1} E\left(\mathbf{Q}_g'\right) E\left(\mathbf{Q}_{1jg}\right) \end{aligned}$$

for each $j = 1, ..., J$. From this, $\boldsymbol{\gamma}$, $\mathbf{m}$, $\boldsymbol{\alpha}$, $\boldsymbol{\delta}$, and $\mathbf{C}$ are identified. Furthermore, the structural parameters $\mathbf{b}$ and $\mathbf{d}$ are also identified from $b_j = \delta_j p_j$ and $d_j = e^{\mathbf{b}' \ln \mathbf{p}} m_j p_j$.

We have now shown identification of the Engel curve system. To identify the full demand system, let $\mathbf{p}_t$ denote the vector of prices in a single price regime t. Using the groups that are observed facing this set of prices, from above we can identity $\boldsymbol{\alpha}_t$, $\mathbf{C}$, $\mathbf{b}$, and $\mathbf{d}$, where $\boldsymbol{\alpha}_t = \mathbf{A}'\mathbf{p}_t$.

Assumption B6: Data are observed in at least J price regimes $\mathbf{p}_1$, ..., $\mathbf{p}_J$ such that the $J \times J$ matrix $\mathbf{P}$ consisting of rows $\mathbf{p}_1'$, ..., $\mathbf{p}_J'$ is nonsingular.

Given Assumption B6, $\mathbf{A}$ is identified by $\mathbf{A}'=(\boldsymbol{\alpha}_1, ..., \boldsymbol{\alpha}_J)\mathbf{P}^{-1}$. The above proves the following theorem:

Theorem 2. Given Assumptions B1-B5, the parameters $\boldsymbol{\gamma}$, $\mathbf{m}$, $\boldsymbol{\alpha}$, $\boldsymbol{\delta}$ and $\mathbf{C}$ in the Engel curve system (12) are identified. If Assumption B6 also holds, the parameters $\mathbf{A}$, $\mathbf{C}$, $\mathbf{b}$, and $\mathbf{d}$ in the full demand system (11) are identified.

8.6 Fixed Effects Estimation of the Demand System

As is standard in the estimation of continuous demand systems, we only need to estimate the model for goods $j = 1, ..., J-1$. The parameters for the last good J are then obtained from the adding up identity that $q_{Ji} = \left(x_i - \sum_{j=1}^{J-1} p_j q_{ji}\right)/p_J$.

For the Engel curve system, we construct the group level GMM estimation based on

$$\left(\widehat{\boldsymbol{\alpha}}', \widehat{m}_1, ..., \widehat{m}_{J-1}, \widehat{\delta}_1, ..., \widehat{\delta}_{J-1}, \widehat{\mathbf{c}}_1', ...\widehat{\mathbf{c}}_{J-1}'\right)' = \arg\min \left(\frac{1}{G}\sum_{g=1}^{G}\mathbf{m}_g\right)' \widehat{\Omega} \left(\frac{1}{G}\sum_{g=1}^{G}\mathbf{m}_g\right),$$

where $\widehat{\Omega}$ is some positive definite moment weighting matrix and

$$\begin{aligned}
\mathbf{m}_g &= \begin{pmatrix} \mathbf{Q}_{11g} \\ \vdots \\ \mathbf{Q}_{1(J-1)g} \end{pmatrix} + 2\begin{pmatrix} m_1 \sum\limits_{j'=1}^{J} \alpha_{j'}\mathbf{Q}_{2j'g} \\ \vdots \\ m_{J-1} \sum\limits_{j'=1}^{J} \alpha_{j'}\mathbf{Q}_{2j'g} \end{pmatrix} - \begin{pmatrix} m_1\mathbf{Q}_{4g} \\ \vdots \\ m_{J-1}\mathbf{Q}_{4g} \end{pmatrix} \\
&\quad -2\begin{pmatrix} m_1 \sum\limits_{j'=1}^{J}\sum\limits_{k=1}^{K} \alpha_{j'}\gamma_k\mathbf{Q}_{3j'kg} \\ \vdots \\ m_{J-1} \sum\limits_{j'=1}^{J}\sum\limits_{k=1}^{K} \alpha_{j'}\gamma_k\mathbf{Q}_{3j'kg} \end{pmatrix} - \begin{pmatrix} m_1 \sum\limits_{k=1}^{K}\sum\limits_{k'=1}^{K} \gamma_k\gamma_{k'}\mathbf{Q}_{5kk'g} \\ \vdots \\ m_{J-1} \sum\limits_{k=1}^{K}\sum\limits_{k'=1}^{K} \gamma_k\gamma_{k'}\mathbf{Q}_{5kk'g} \end{pmatrix} \\
&\quad +2\begin{pmatrix} m_1 \sum\limits_{k=1}^{K} \gamma_k\mathbf{Q}_{6kg} \\ \vdots \\ m_{J-1} \sum\limits_{k=1}^{K} \gamma_k\mathbf{Q}_{6kg} \end{pmatrix} - \begin{pmatrix} \delta_1\mathbf{Q}_{7g} \\ \vdots \\ \delta_{J-1}\mathbf{Q}_{7g} \end{pmatrix} + \begin{pmatrix} \sum\limits_{k=1}^{K} (\delta_1\gamma_k - c_{1k})\mathbf{Q}_{8kg} \\ \vdots \\ \sum\limits_{k=1}^{K} (\delta_{J-1}\gamma_k - c_{(J-1)k})\mathbf{Q}_{8kg} \end{pmatrix}
\end{aligned}$$

is a $q(J-1)-$dimensional vector.

Or the individual level GMM estimation with group clustered standard error

$$\left(\widehat{\boldsymbol{\alpha}}', \widehat{m}_1, ..., \widehat{m}_{J-1}, \widehat{\delta}_1, ..., \widehat{\delta}_{J-1}, \widehat{\mathbf{c}}_1', ...\widehat{\mathbf{c}}_{J-1}'\right)' = \arg\min \left(\frac{\sum_{g=1}^{G}\sum_{(i,i')\in\Gamma_g}\mathbf{m}_{gii'}}{\sum_{g=1}^{G}\sum_{(i,i')\in\Gamma_g}1}\right)' \widehat{\Omega} \left(\frac{\sum_{g=1}^{G}\sum_{(i,i')\in\Gamma_g}\mathbf{m}_{gii'}}{\sum_{g=1}^{G}\sum_{(i,i')\in\Gamma_g}1}\right),$$

where $\widehat{\Omega}$ is some positive definite moment weighting matrix and

$$\mathbf{m}_{gii'} = \begin{pmatrix} L_{11gii'}\mathbf{r}_{gii'} \\ \vdots \\ L_{1(J-1)gii'}\mathbf{r}_{gii'} \end{pmatrix} + 2\begin{pmatrix} m_1 \sum_{j'=1}^{J} \alpha_{j'} L_{2j'gii'}\mathbf{r}_{gii'} \\ \vdots \\ m_{J-1} \sum_{j'=1}^{J} \alpha_{j'} L_{2j'gii'}\mathbf{r}_{gii'} \end{pmatrix} - \begin{pmatrix} m_1 L_{4gii'}\mathbf{r}_{gii'} \\ \vdots \\ m_{J-1} L_{4gii'}\mathbf{r}_{gii'} \end{pmatrix}$$

$$-2\begin{pmatrix} m_1 \sum_{j'=1}^{J}\sum_{k=1}^{K} \alpha_{j'}\gamma_k L_{3j'kgii'}\mathbf{r}_{gii'} \\ \vdots \\ m_{J-1} \sum_{j'=1}^{J}\sum_{k=1}^{K} \alpha_{j'}\gamma_k L_{3j'kgii'}\mathbf{r}_{gii'} \end{pmatrix} - \begin{pmatrix} m_1 \sum_{k=1}^{K}\sum_{k'=1}^{K} \gamma_k\gamma_{k'} L_{5kk'gii'}\mathbf{r}_{gii'} \\ \vdots \\ m_{J-1} \sum_{k=1}^{K}\sum_{k'=1}^{K} \gamma_k\gamma_{k'} L_{5kk'gii'}\mathbf{r}_{gii'} \end{pmatrix}$$

$$+2\begin{pmatrix} m_1 \sum_{k=1}^{K} \gamma_k L_{6kgii'}\mathbf{r}_{gii'} \\ \vdots \\ m_{J-1} \sum_{k=1}^{K} \gamma_k L_{6kgii'}\mathbf{r}_{gii'} \end{pmatrix} - \begin{pmatrix} \delta_1 L_{7gii'}\mathbf{r}_{gii'} \\ \vdots \\ \delta_{J-1} L_{7gii'}\mathbf{r}_{gii'} \end{pmatrix} + \begin{pmatrix} \sum_{k=1}^{K} (\delta_1\gamma_k - c_{1k}) L_{8kgii'}\mathbf{r}_{gii'} \\ \vdots \\ \sum_{k=1}^{K} (\delta_{J-1}\gamma_k - c_{(J-1)k}) L_{8kgii'}\mathbf{r}_{gii'} \end{pmatrix}.$$

For the full demand system, the GMM estimation uses each different value of gt as a different group, so the total number of groups is $N = \sum_{g=1}^{G}\sum_{t=1}^{T} 1$ where the sum is over all values gt can take on. Define

$$\Gamma_{gt} = \{(i, i') \mid i \text{ and } i' \text{ are observed}, i \in gt, i' \in gt, i \neq i'\}$$

So Γ_{ngt} is the set of all observed pairs of individuals i and i' in the group g at period t.Let the instrument vector $\mathbf{r}_{gtii'}$ be any functional form of $\mathbf{r}_{gt}$, $(x_i, \mathbf{z}_i')'$, and $(x_{i'}, \mathbf{z}_{i'}')'$. Denote

$$\begin{aligned} L_{1jgtii'} &= (q_{ji} - q_{ji'}),\ \ L_{2jgtii'} = \widehat{q}_{jgt,-ii'}(x_i - x_{i'}),\ \ L_{3jkgtii'} = \widehat{q}_{jgt,-ii'}(z_{ki} - z_{ki'}),\ \ L_{4gii'} = x_i^2 - x_{i'}^2, \\ L_{5kk'gtii'} &= z_{ki}z_{k'i} - z_{ki'}z_{k'i'},\ \ L_{6kgtii'} = z_{ki}x_i - z_{ki'}x_{i'},\ \ L_{7gtii'} = x_i - x_{i'},\ \ L_{8kgtii'} = z_{ki} - z_{ki'}. \end{aligned}$$

For $\ell \in \{1j, 2j, 3jk, 4, 5kk', 6k, 7, 8k \mid j = 1, ..., J; k, k' = 1, ..., K\}$, define vectors

$$\mathbf{Q}_{\ell gt} = \frac{\sum_{(i,i')\in\Gamma_{gt}} L_{\ell gtii'}\mathbf{r}_{gtii'}}{\sum_{(i,i')\in\Gamma_{gt}} 1}.$$

We construct the group level GMM estimation

$$\left(\widehat{\mathbf{A}}_1', ..., \widehat{\mathbf{A}}_{J-1}', \widehat{b}_1, ..., \widehat{b}_{J-1}, \widehat{d}_1, ..., \widehat{d}_{J-1}, \widehat{\mathbf{c}}_1', ... \widehat{\mathbf{c}}_{J-1}'\right)' = \arg\min \left(\frac{1}{N}\sum_{g=1}^{G}\sum_{t=1}^{T}\mathbf{m}_{gt}\right)' \widehat{\Omega} \left(\frac{1}{N}\sum_{g=1}^{G}\sum_{t=1}^{T}\mathbf{m}_{gt}\right),$$

where $\widehat{\Omega}$ is some positive definite moment weighting matrix and

$$
\begin{aligned}
\mathbf{m}_{gt} \;=\;& \begin{pmatrix} \mathbf{Q}_{11gt} \\ \vdots \\ \mathbf{Q}_{1(J-1)gt} \end{pmatrix} + 2e^{-\mathbf{b}'\ln \mathbf{p}_t} \begin{pmatrix} \frac{d_1}{p_{1t}} \sum_{j_1=1}^{J} \sum_{j_2=1}^{J} A_{j_1 j_2} p_{j_1 t} \mathbf{Q}_{2j_2gt} \\ \vdots \\ \frac{d_{J-1}}{p_{(J-1)t}} \sum_{j_1=1}^{J} \sum_{j_2=1}^{J} A_{j_1 j_2} p_{j_1 t} \mathbf{Q}_{2j_2gt} \end{pmatrix} - e^{-\mathbf{b}'\ln \mathbf{p}_t} \begin{pmatrix} \frac{d_1}{p_{1t}} \mathbf{Q}_{4gt} \\ \vdots \\ \frac{d_{J-1}}{p_{(J-1)t}} \mathbf{Q}_{4gt} \end{pmatrix} \\
& -2e^{-\mathbf{b}'\ln \mathbf{p}_t} \begin{pmatrix} \frac{d_1}{p_{1t}} \sum_{j_1=1}^{J} \sum_{j_2=1}^{J} \sum_{j_3=1}^{J} \sum_{k=1}^{K} A_{j_1 j_2} p_{j_1 t} c_{j_3 k} p_{j_3 t} \mathbf{Q}_{3j_2kgt} \\ \vdots \\ \frac{d_{J-1}}{p_{(J-1)t}} \sum_{j_1=1}^{J} \sum_{j_2=1}^{J} \sum_{j_3=1}^{J} \sum_{k=1}^{K} A_{j_1 j_2} p_{j_1 t} c_{j_3 k} p_{j_3 t} \mathbf{Q}_{3j_2kgt} \end{pmatrix} - \begin{pmatrix} \frac{b_1}{p_{1t}} \mathbf{Q}_{7gt} \\ \vdots \\ \frac{b_{J-1}}{p_{(J-1)t}} \mathbf{Q}_{7gt} \end{pmatrix} \\
& -e^{-\mathbf{b}'\ln \mathbf{p}_t} \begin{pmatrix} \frac{d_1}{p_{1t}} \sum_{j_1=1}^{J} \sum_{j_2=1}^{J} \sum_{k=1}^{K} \sum_{k'=1}^{K} p_{j_1 t} p_{j_2 t} c_{j_1 k} c_{j_2 k'} \mathbf{Q}_{5kk'gt} \\ \vdots \\ \frac{d_{J-1}}{p_{(J-1)t}} \sum_{j_1=1}^{J} \sum_{j_2=1}^{J} \sum_{k=1}^{K} \sum_{k'=1}^{K} p_{j_1 t} p_{j_2 t} c_{j_1 k} c_{j_2 k'} \mathbf{Q}_{5kk'gt} \end{pmatrix} \\
& +2e^{-\mathbf{b}'\ln \mathbf{p}_t} \begin{pmatrix} \frac{d_1}{p_{1t}} \sum_{j=1}^{J} \sum_{k=1}^{K} c_{jk} p_{jt} \mathbf{Q}_{6kgt} \\ \vdots \\ \frac{d_{J-1}}{p_{(J-1)t}} \sum_{j=1}^{J} \sum_{k=1}^{K} c_{jk} p_{jt} \mathbf{Q}_{6kgt} \end{pmatrix} + \begin{pmatrix} \sum_{j=1}^{J} \sum_{k=1}^{K} (\frac{b_1}{p_{1t}} c_{jk} p_{jt} - c_{1k}) \mathbf{Q}_{8kgt} \\ \vdots \\ \sum_{j=1}^{J} \sum_{k=1}^{K} (\frac{b_{J-1}}{p_{(J-1)t}} c_{jk} p_{jt} - c_{(J-1)k}) \mathbf{Q}_{8kgt} \end{pmatrix}.
\end{aligned}
$$

Or the individual level GMM estimation with group clustered standard error

$$
\begin{aligned}
& \left(\widehat{\mathbf{A}}'_1, ..., \widehat{\mathbf{A}}'_{J-1}, \widehat{b}_1, ..., \widehat{b}_{J-1}, \widehat{d}_1, ..., \widehat{d}_{J-1}, \widehat{\mathbf{c}}'_1, ... \widehat{\mathbf{c}}'_{J-1}\right)' \\
=\;& \arg\min \left(\frac{\sum_{t=1}^{T} \sum_{g=1}^{G} \sum_{(i,i')\in\Gamma_{gt}} \mathbf{m}_{gtii'}}{\sum_{t=1}^{T} \sum_{g=1}^{G} \sum_{(i,i')\in\Gamma_{gt}} 1}\right)' \widehat{\Omega} \left(\frac{\sum_{t=1}^{T} \sum_{g=1}^{G} \sum_{(i,i')\in\Gamma_{gt}} \mathbf{m}_{gtii'}}{\sum_{t=1}^{T} \sum_{g=1}^{G} \sum_{(i,i')\in\Gamma_{gt}} 1}\right),
\end{aligned}
$$

where

$$
\mathbf{m}_{gtii'} = \begin{pmatrix} L_{11gtii'}\mathbf{r}_{gtii'} \\ \vdots \\ L_{1(J-1)gtii'}\mathbf{r}_{gtii'} \end{pmatrix} + 2e^{-\mathbf{b}'\ln\mathbf{p}_t} \begin{pmatrix} \frac{d_1}{p_{1t}} \sum_{j_1=1}^{J}\sum_{j_2=1}^{J} A_{j_1j_2}p_{j_1t}L_{2j_2gtii'}\mathbf{r}_{gtii'} \\ \vdots \\ \frac{d_{J-1}}{p_{(J-1)t}} \sum_{j_1=1}^{J}\sum_{j_2=1}^{J} A_{j_1j_2}p_{j_1t}L_{2j_2gtii'}\mathbf{r}_{gtii'} \end{pmatrix}
$$

$$
-2e^{-\mathbf{b}'\ln\mathbf{p}_t} \begin{pmatrix} \frac{d_1}{p_{1t}} \sum_{j_1=1}^{J}\sum_{j_2=1}^{J}\sum_{j_3=1}^{J}\sum_{k=1}^{K} A_{j_1j_2}p_{j_1t}c_{j_3k}p_{j_3t}L_{3j_2kgtii'}\mathbf{r}_{gtii'} \\ \vdots \\ \frac{d_{J-1}}{p_{(J-1)t}} \sum_{j_1=1}^{J}\sum_{j_2=1}^{J}\sum_{j_3=1}^{J}\sum_{k=1}^{K} A_{j_1j_2}p_{j_1t}c_{j_3k}p_{j_3t}L_{3j_2kgtii'}\mathbf{r}_{gtii'} \end{pmatrix} - \begin{pmatrix} \frac{b_1}{p_{1t}}L_{7gtii'}\mathbf{r}_{gtii'} \\ \vdots \\ \frac{b_{J-1}}{p_{(J-1)t}}L_{7gtii'}\mathbf{r}_{gtii'} \end{pmatrix}
$$

$$
-e^{-\mathbf{b}'\ln\mathbf{p}_t} \begin{pmatrix} \frac{d_1}{p_{1t}} \sum_{j_1=1}^{J}\sum_{j_2=1}^{J}\sum_{k=1}^{K}\sum_{k'=1}^{K} p_{j_1t}p_{j_2t}c_{j_1k}c_{j_2k'}L_{5kk'gtii'}\mathbf{r}_{gtii'} \\ \vdots \\ \frac{d_{J-1}}{p_{(J-1)t}} \sum_{j_1=1}^{J}\sum_{j_2=1}^{J}\sum_{k=1}^{K}\sum_{k'=1}^{K} p_{j_1t}p_{j_2t}c_{j_1k}c_{j_2k'}L_{5kk'gtii'}\mathbf{r}_{gtii'} \end{pmatrix} - e^{-\mathbf{b}'\ln\mathbf{p}_t} \begin{pmatrix} \frac{d_1}{p_{1t}}L_{4gtii'}\mathbf{r}_{gtii'} \\ \vdots \\ \frac{d_{J-1}}{p_{(J-1)t}}L_{4gtii'}\mathbf{r}_{gtii'} \end{pmatrix}
$$

$$
+2e^{-\mathbf{b}'\ln\mathbf{p}_t} \begin{pmatrix} \frac{d_1}{p_{1t}} \sum_{j=1}^{J}\sum_{k=1}^{K} c_{jk}p_{jt}L_{6kgtii'}\mathbf{r}_{gtii'} \\ \vdots \\ \frac{d_{J-1}}{p_{(J-1)t}} \sum_{j=1}^{J}\sum_{k=1}^{K} c_{jk}p_{jt}L_{6kgtii'}\mathbf{r}_{gtii'} \end{pmatrix} + \begin{pmatrix} \sum_{j=1}^{J}\sum_{k=1}^{K}(\frac{b_1}{p_{1t}}c_{jk}p_{jt} - c_{1k})L_{8kgtii'}\mathbf{r}_{gtii'} \\ \vdots \\ \sum_{j=1}^{J}\sum_{k=1}^{K}(\frac{b_{J-1}}{p_{(J-1)t}}c_{jk}p_{jt} - c_{(J-1)k})L_{8kgtii'}\mathbf{r}_{gtii'} \end{pmatrix}.
$$

8.7 Construction of Instruments For Fixed Effects Demand System Estimation

For estimation, we need to establish that the set of instruments $\mathbf{r}_{gt}$ provided in the text are valid. For any matrix of random variables $\mathbf{w}$, we have $\widehat{\mathbf{w}}_{gt\cdot}$ defined by

$$
\widehat{\mathbf{w}}_{gt\cdot} = \frac{\sum_{s\neq t}\sum_{i\in gs}\mathbf{w}_i}{\sum_{s\neq t}\sum_{i\in gs}1}
$$

From Assumption B4, we can write $\widehat{\mathbf{w}}_{gt\cdot} = \overline{\mathbf{w}}_{gt\cdot} + \varepsilon_{wgt\cdot}$, where $\varepsilon_{wgt\cdot}$ is a summation of measurement errors from other periods. Assume now that $\varepsilon_{wgt} \perp (\varepsilon_{wgt\cdot}, \overline{\mathbf{w}}_{gt\cdot})$.

As discussed after assumption B4, we can think of $(x_i, \mathbf{z}_i)$ as being determined by having $(\varepsilon_{ix}, \boldsymbol{\varepsilon}_{iz})$ drawn independently from group level variables. As long as these draws are independent across individuals, and different individuals are observed in each time period, then we will have $\varepsilon_{wgt} \perp (\varepsilon_{wgt\cdot}, \overline{\mathbf{w}}_{gt\cdot})$ for $\mathbf{w}$ being suitable functions of $(x_i, \mathbf{z}_i)$. Alternatively, if we interpret the ε's as being measurement errors in group level variables, then the assumption is that these measurement errors are independent over time. In contrast to the ε's, we assume that true group level variables like $\overline{x}_{gt}$ and $\overline{\mathbf{z}}_{gt}$ are correlated over time, e.g., the true mean group income in one time period is not independent of the true mean group income in other time periods.

Given $\varepsilon_{wgt} \perp (\varepsilon_{wgt\cdot}, \overline{\mathbf{w}}_{gt\cdot})$, we have

$$0 = E(\boldsymbol{\varepsilon}_{qgt,-ii'}[(x_i - x_{i'}) - \boldsymbol{\gamma}'_{gt}(\mathbf{z}_i - \mathbf{z}_{i'})] \mid \widehat{\mathbf{w}}_{gt\cdot}, x_{it}, x_{i't}, \mathbf{z}_{it}, \mathbf{z}_{i't}),$$

because

$$E\left(\overline{\mathbf{q}}_{gt}[(x_i - x_{i'}) - \boldsymbol{\gamma}'_{gt}(\mathbf{z}_i - \mathbf{z}_{i'})](\widehat{\mathbf{x}^*}_{gt,-ii'} - \overline{\mathbf{x}^*}_{gt}) \mid \overline{\mathbf{x}^*}_{gt}, \overline{\mathbf{x}^*\mathbf{x}^{*\prime}}_{gt}, \mathbf{v}_{gt}, \overline{\mathbf{w}}_{gt\cdot}, \varepsilon_{wgt\cdot}, \mathbf{x}^*_{it}, \mathbf{x}^*_{i't}\right) = 0,$$

and

$$\begin{aligned} E\left([(\mathbf{x}^*_i - \mathbf{x}^*_{i'})](\widehat{\mathbf{x}^*}_{gT,-ii'} - \overline{\mathbf{x}^*}_{gt})' \mid \overline{\mathbf{w}}_{gt\cdot}, \varepsilon_{wgt\cdot}, \mathbf{x}^*_{it}, \mathbf{x}^*_{i't}\right) &= 0; \\ E\left([(\mathbf{x}^*_i - \mathbf{x}^*_{i'})](\widehat{\mathbf{x}^*\mathbf{x}^{*\prime}}_{gt,-ii'} - \overline{\mathbf{x}^*\mathbf{x}^{*\prime}}_{gt})' \mid \overline{\mathbf{w}}_{gt\cdot}, \varepsilon_{wgt\cdot}, \mathbf{x}^*_{it}, \mathbf{x}^*_{i't}\right) &= 0, \end{aligned}$$

where $\mathbf{x}^* = (x, \mathbf{z}')'$. It follows that $\left(\widehat{\mathbf{x}^*\mathbf{x}^{*\prime}}_{gt\cdot}, \widehat{\mathbf{x}^*}_{gt\cdot}\widehat{\mathbf{x}^*}'_{gt\cdot}, \widehat{\mathbf{x}^*}_{gt\cdot}\right)$ is a valid instrument for $\widehat{\mathbf{q}}_{gt,-ii'}$.

The full set of proposed instruments is therefore $\mathbf{r}_{gii'} = \mathbf{r}_g \otimes (\mathbf{x}^*_i - \mathbf{x}^*_{i'}, \mathbf{x}^*_i\mathbf{x}^{*\prime}_i - \mathbf{x}^*_{i'}\mathbf{x}^{*\prime}_{i'})$, where

$$\mathbf{r}_g = \left(\widehat{\mathbf{x}^*\mathbf{x}^{*\prime}}_{gt\cdot}, \widehat{\mathbf{x}^*}_{gt\cdot}\widehat{\mathbf{x}^*}'_{gt\cdot}, \widehat{\mathbf{x}^*}_{gt\cdot}, \mathbf{x}^*_i + \mathbf{x}^*_{i'}, x_i^2 + x_{i'}^2, x_i^{1/2} + x_{i'}^{1/2}\right),$$

for the Engel curve system, and $\mathbf{r}_{gtii'} = \mathbf{r}_{gt} \otimes (\mathbf{x}^*_i - \mathbf{x}^*_{i'}, \mathbf{x}^*_i\mathbf{x}^{*\prime}_i - \mathbf{x}^*_{i'}\mathbf{x}^{*\prime}_{i'})$, where

$$\mathbf{r}_{gt} = \mathbf{p}'_t \otimes \left(\widehat{\mathbf{x}^*\mathbf{x}^{*\prime}}_{gt\cdot}, \widehat{\mathbf{x}^*}_{gt\cdot}\widehat{\mathbf{x}^*}'_{gt\cdot}, \widehat{\mathbf{x}^*}_{gt\cdot}, \mathbf{x}^*_i + \mathbf{x}^*_{i'}, x_i^2 + x_{i'}^2, x_i^{1/2} + x_{i'}^{1/2}\right).$$

for the full demand system.

8.8 Derivation of Random Effects Demand System Moments

For the random effects model, the Engel curve system is sufficient to identify all the structure parameters, including the peer effects matrix $\mathbf{A}$. The Engel curve model with random effects is

$$\begin{aligned} \mathbf{q}_i &= x_i^2\mathbf{m} + \mathbf{m}\left(\boldsymbol{\alpha}'\overline{\mathbf{q}}_g\right)^2 + (\boldsymbol{\gamma}'\mathbf{z}_i\mathbf{z}_i'\boldsymbol{\gamma})\,\mathbf{m} - 2\mathbf{m}\boldsymbol{\alpha}'\overline{\mathbf{q}}_g x_i - 2\mathbf{m}\boldsymbol{\gamma}'\mathbf{z}_i x_i + 2\mathbf{m}\boldsymbol{\alpha}'\overline{\mathbf{q}}_g\boldsymbol{\gamma}'\mathbf{z}_i \\ &\quad + \left(x_i - \boldsymbol{\alpha}'\overline{\mathbf{q}}_g - \boldsymbol{\gamma}'\mathbf{z}_i\right)\boldsymbol{\delta} + \mathbf{A}\overline{\mathbf{q}}_g + \mathbf{C}\mathbf{z}_i + \mathbf{v}_g + \mathbf{u}_i, \end{aligned}$$

Therefore,

$$\begin{aligned}
\boldsymbol{\varepsilon}_{qi'} &= \mathbf{q}_{i'} - \overline{\mathbf{q}}_g = -2\mathbf{m}\boldsymbol{\alpha}'\overline{\mathbf{q}}_g(\varepsilon_{xi'} - \boldsymbol{\gamma}'\boldsymbol{\varepsilon}_{zi'}) + \varepsilon_{x^2i'}\mathbf{m} + \boldsymbol{\gamma}'\boldsymbol{\varepsilon}_{zzi'}\boldsymbol{\gamma}\mathbf{m} - 2\mathbf{m}\boldsymbol{\gamma}'\boldsymbol{\varepsilon}_{zxi'} \\
&\quad + \boldsymbol{\delta}\varepsilon_{xi'} + (\mathbf{C} - \boldsymbol{\delta}\boldsymbol{\gamma}')\boldsymbol{\varepsilon}_{zi'} + \mathbf{v}_g - E(\mathbf{v}_g) + \widehat{\mathbf{u}}_{i'}; \\
\boldsymbol{\varepsilon}_{qg,-ii'} &= \widehat{\mathbf{q}}_{g,-ii'} - \overline{\mathbf{q}}_g = -2\mathbf{m}\boldsymbol{\alpha}'\overline{\mathbf{q}}_g(\varepsilon_{xg,-ii'} - \boldsymbol{\gamma}'\boldsymbol{\varepsilon}_{zg,-ii'}) + \varepsilon_{x^2g,-ii'}\mathbf{m} + \boldsymbol{\gamma}'\boldsymbol{\varepsilon}_{zzg,-ii'}\boldsymbol{\gamma}\mathbf{m} \\
&\quad -2\mathbf{m}\boldsymbol{\gamma}'\boldsymbol{\varepsilon}_{zxg,-ii'} + \boldsymbol{\delta}\varepsilon_{xg,-ii'} + (\mathbf{C} - \boldsymbol{\delta}\boldsymbol{\gamma}')\boldsymbol{\varepsilon}_{zg,-ii'} + \mathbf{v}_g - E(\mathbf{v}_g) + \widehat{\mathbf{u}}_{g,-ii'}.
\end{aligned}$$

By rewriting q_{ji} as

$$\begin{aligned}
q_{ji} &= m_j\left(\boldsymbol{\alpha}'\overline{\mathbf{q}}_g\right)^2 + m_j(x_i - \mathbf{z}_i'\boldsymbol{\gamma})^2 - [2m_j\,(x_i - \boldsymbol{\gamma}'\mathbf{z}_i) + \delta_j]\boldsymbol{\alpha}'\overline{\mathbf{q}}_g + \delta_j(x_i - \boldsymbol{\gamma}'\mathbf{z}_i) \\
&\quad + \mathbf{c}_j'\mathbf{z}_i + \mathbf{A}_j'\overline{\mathbf{q}}_g + v_{jg} + u_{ji} \\
&= m_j\boldsymbol{\alpha}'\widehat{\mathbf{q}}_{g,-ii'}\boldsymbol{\alpha}'\mathbf{q}_{i'} + m_j(x_i - \mathbf{z}_i'\boldsymbol{\gamma})^2 - [2m_j\,(x_i - \boldsymbol{\gamma}'\mathbf{z}_i) + \delta_j]\boldsymbol{\alpha}'\widehat{\mathbf{q}}_{g,-ii'} + \delta_j\,(x_i - \boldsymbol{\gamma}'\mathbf{z}_i) \\
&\quad + \mathbf{c}_j'\mathbf{z}_i + \mathbf{A}_j'\widehat{\mathbf{q}}_{g,-ii'} + v_{jg} + u_{ji} + \widetilde{\varepsilon}_{jgii'},
\end{aligned}$$

where

$$\begin{aligned}
\widetilde{\varepsilon}_{jgii'} &= m_j\boldsymbol{\alpha}'(\overline{\mathbf{q}}_g\overline{\mathbf{q}}_g' - \widehat{\mathbf{q}}_{g,-ii'}\mathbf{q}_{i'}')\boldsymbol{\alpha} - [2m_j\,(x_i - \boldsymbol{\gamma}'\mathbf{z}_i) + \delta_j]\boldsymbol{\alpha}'(\overline{\mathbf{q}}_g - \widehat{\mathbf{q}}_{g,-ii'}) + \mathbf{A}_j'(\overline{\mathbf{q}}_g - \widehat{\mathbf{q}}_{g,-ii'}) \\
&= -m_j\boldsymbol{\alpha}'[(\boldsymbol{\varepsilon}_{qg,-ii'} + \boldsymbol{\varepsilon}_{qi'})\overline{\mathbf{q}}_g' + \boldsymbol{\varepsilon}_{qg,-ii'}\boldsymbol{\varepsilon}_{qi'}']\boldsymbol{\alpha} + [2m_j\,(x_i - \boldsymbol{\gamma}'\mathbf{z}_i) + \delta_j]\boldsymbol{\alpha}'\boldsymbol{\varepsilon}_{qg,-ii'} - \mathbf{A}_j'\boldsymbol{\varepsilon}_{qg,-ii'}.
\end{aligned}$$

and letting $U_{jii'} = v_{jg} + u_{ji} + \widetilde{\varepsilon}_{jgii'}$, we have the conditional expectation

$$E(U_{jii'}|\mathbf{z}_i, x_i, \mathbf{r}_g) = E(v_{jg}|\mathbf{z}_i, x_i, \mathbf{r}_g) - m_j\boldsymbol{\alpha}'E(\boldsymbol{\varepsilon}_{qg,-ii'}\boldsymbol{\varepsilon}_{qi'}'|\mathbf{z}_i, x_i, \mathbf{r}_g)\boldsymbol{\alpha} = \mu_j - m_j\boldsymbol{\alpha}'\boldsymbol{\Sigma}_v\boldsymbol{\alpha},$$

where $\mu_j = E(v_{jg}|\mathbf{z}_i, x_i, \mathbf{r}_g)$ and $\boldsymbol{\Sigma}_v = Var(\mathbf{v}_g|\mathbf{z}_i, x_i, \mathbf{r}_g)$. From this, we can construct the conditional moment condition

$$\begin{aligned}
&E\left[q_{ji} - m_j\boldsymbol{\alpha}'\widehat{\mathbf{q}}_{g,-ii'}\boldsymbol{\alpha}'\mathbf{q}_{i'} - m_j(x_i - \mathbf{z}_i'\boldsymbol{\gamma})^2 + [2m_j\,(x_i - \boldsymbol{\gamma}'\mathbf{z}_i) + \delta_j]\boldsymbol{\alpha}'\widehat{\mathbf{q}}_{g,-ii'}\right. \\
&\left.-\delta_j\,(x_i - \boldsymbol{\gamma}'\mathbf{z}_i) - \mathbf{c}_j'\mathbf{z}_i - \mathbf{A}_j'\widehat{\mathbf{q}}_{g,-ii'}|x_i, \mathbf{z}_i, \mathbf{r}_g\right] - v_{j0} = 0\ ,
\end{aligned}$$

where $v_{j0} = \mu_j - m_j\boldsymbol{\alpha}'\boldsymbol{\Sigma}_v\boldsymbol{\alpha}$ is a constant.

Let the instrument vector $\mathbf{r}_{gi}$ be any functional form of $\mathbf{r}_g$ and $(x_i, \mathbf{z}_i')'$. Then for any $i, i' \in g$ with $i \neq i'$, the following unconditional moment condition holds

$$\begin{aligned}
&E\left[\left(q_{ji} - m_j\boldsymbol{\alpha}'\widehat{\mathbf{q}}_{g,-ii'}\boldsymbol{\alpha}'\mathbf{q}_{i'} - m_j(x_i - \mathbf{z}_i'\boldsymbol{\gamma})^2 + [2m_j\,(x_i - \boldsymbol{\gamma}'\mathbf{z}_i) + \delta_j]\boldsymbol{\alpha}'\widehat{\mathbf{q}}_{g,-ii'}\right.\right. \\
&\left.\left.-\delta_j\,(x_i - \boldsymbol{\gamma}'\mathbf{z}_i) - \mathbf{c}_j'\mathbf{z}_i - \mathbf{A}_j'\widehat{\mathbf{q}}_{g,-ii'} - v_{j0}\right)\mathbf{r}_{gi}\right] = 0\ .
\end{aligned}$$

We can sum over all $i' \neq i$ in the group g. Using the property of $\frac{1}{n_g-1}\sum_{i'\in g, i'\neq i}\widehat{q}_{jg,-ii'} = \widehat{q}_{jg,-i}$, then for any $i \in g$,

$$E\{\mathbf{r}_{gi}[q_{ji}-m_j\boldsymbol{\alpha}'\frac{1}{n_g-1}\sum_{i'\in g,i'\neq i}\widehat{\mathbf{q}}_{g,-ii'}\mathbf{q}_{i'}'\boldsymbol{\alpha}-m_jx_i^2-m_j\boldsymbol{\gamma}'\mathbf{z}_i\mathbf{z}_i'\boldsymbol{\gamma}+2m_j\boldsymbol{\gamma}'\mathbf{z}_ix_i+2m_j\boldsymbol{\alpha}'\widehat{\mathbf{q}}_{g,-i}x_i$$
$$-2m_j\boldsymbol{\alpha}'\widehat{\mathbf{q}}_{g,-i}\mathbf{z}_i'\boldsymbol{\gamma}-\delta_jx_i+\widehat{\mathbf{q}}_{g,-i}(\delta_j\boldsymbol{\alpha}-\mathbf{A}_j)+\mathbf{z}_i(\delta_j\boldsymbol{\gamma}-\mathbf{c}_j)-v_{j0}]\}=0$$

Denote

$$\begin{aligned}
L_{1jgi} &= q_{ji},\ \ L_{2jj'g}=\frac{1}{n_g-1}\sum_{i'\in g,i'\neq i}\widehat{q}_{jg,-ii'}q_{j'i'},\quad L_{3gi}=x_i^2,\\
L_{4kk'gi} &= z_{ki}z_{k'i},\quad L_{5kgi}=z_{ki}x_i,\quad L_{6jgi}=\widehat{q}_{jg,-i}x_i,\ \ L_{7jkgi}=\widehat{q}_{jg,-i}z_{ki},,\\
L_{8gi} &= x_i,\ \ L_{9jgi}=\widehat{q}_{jg,-i},\ \ L_{10kgi}=z_{ki},\ \ L_{11gi}=1.
\end{aligned}$$

For $\ell \in \{1j,2jj',3,4kk',5k,6j,7jk,8,9j,10k,11 \mid j,j'=1,...,J;k,k'=1,...,K\}$, define group level vectors

$$\mathbf{H}_{\ell g}=\frac{1}{n_g-1}\sum_{i\in g}L_{\ell gi}\mathbf{r}_{gi}.$$

Then for each good j, the identification is based on

$$\begin{aligned}
&E\Bigg(\mathbf{H}_{1jg}-m_j\sum_{j=1}^{J}\sum_{j'=1}^{J}\alpha_{j'}\alpha_j\mathbf{H}_{2jj'g}-m_j\mathbf{H}_{3g}-m_j\sum_{k=1}^{K}\sum_{k'=1}^{K}\gamma_k\gamma_{k'}\mathbf{H}_{4kk'g}+2m_j\sum_{k=1}^{K}\gamma_k\mathbf{H}_{5kg}\\
&+2m_j\sum_{j'=1}^{J}\alpha_{j'}\mathbf{H}_{6j'g}-2m_j\sum_{j'=1}^{J}\sum_{k=1}^{K}a_{j'}\gamma_k\mathbf{H}_{7j'kg}-\delta_j\mathbf{H}_{8g}+\sum_{j'=1}^{J}(\delta_j\alpha_{j'}-A_{jj'})\mathbf{H}_{9j'g}\\
&+\sum_{k=1}^{K}(\delta_j\gamma_k-c_{jk})\mathbf{H}_{10kg}-v_{j0}\mathbf{H}_{11g}\Bigg)=0.
\end{aligned}$$

Assumption C4: $E\left(\mathbf{H}_g'\right)E\left(\mathbf{H}_g\right)$ is nonsingular, where

$$\begin{aligned}
\mathbf{H}_g \;=\; &(\mathbf{H}_{211g},...,\mathbf{H}_{2JJg},\mathbf{H}_{3g},\mathbf{H}_{411g},...,\mathbf{H}_{4KKg},\mathbf{H}_{51g},...,\mathbf{H}_{5Kg},\mathbf{H}_{61g},...,\\
&\mathbf{H}_{6Jg},\mathbf{H}_{711g},...,\mathbf{H}_{7JKg},\mathbf{H}_{8g},\mathbf{H}_{91g},...,\mathbf{H}_{9Jg},\mathbf{H}_{101g},...,\mathbf{H}_{10Kg},\mathbf{H}_{11g}).
\end{aligned}$$

Under Assumptions C1-C4, we can identify

$$\begin{aligned}
&(m_j\alpha_1\boldsymbol{\alpha}',...,m_j\alpha_J\boldsymbol{\alpha}',m_j,m_j\gamma_1\boldsymbol{\gamma}',...,m_j\gamma_K\boldsymbol{\gamma}',-2m_j\boldsymbol{\gamma}',-2m_j\boldsymbol{\alpha}',2m_j\gamma_1\boldsymbol{\alpha}',...,2m_j\gamma_K\boldsymbol{\alpha}',\\
&\delta_j,-\delta_j\boldsymbol{\alpha}'+\mathbf{A}_j',\mathbf{c}_j'-\delta_j\boldsymbol{\gamma}',v_{j0})'=\left[E\left(\mathbf{H}_g'\right)E\left(\mathbf{H}_g\right)\right]^{-1}E\left(\mathbf{H}_g'\right)E\left(\mathbf{H}_{1jg}\right).
\end{aligned}$$

for each $j=1,...,J$. From this, $\boldsymbol{\gamma}$, $\mathbf{m}$, $\boldsymbol{\alpha}$, $\boldsymbol{\delta}$, $\mathbf{A}$, $\mathbf{C}$ and $\mathbf{v}_0$ are all identified. Furthermore, the structural parameters $\mathbf{b}$ and $\mathbf{d}$ are also identified from $b_j=\delta_jp_j$ and $d_j=e^{\mathbf{b}'\ln\mathbf{p}}m_jp_j$.

Hence, all the structure parameters are identified in the Engel curve system with random effects. This is different from the fixed effects model because the key term for identifying $\mathbf{A}$ is $\mathbf{A}\overline{\mathbf{q}}_{\mathbf{g}}$, which is differenced out in fixed effects model. For estimation, we construct the group level GMM estimation based on

$$\left(\widehat{\mathbf{A}}_1', ..., \widehat{\mathbf{A}}_{J-1}', \widehat{b}_1, ..., \widehat{b}_{J-1}, \widehat{d}_1, ..., \widehat{d}_{J-1}, \widehat{\mathbf{c}}_1', ...\widehat{\mathbf{c}}_{J-1}', \widehat{v}_{1,0}, ...\widehat{v}_{J-1,0}\right)' = \arg\min \left(\frac{1}{G}\sum_{g=1}^{G}\mathbf{m}_g\right)' \widehat{\Omega} \left(\frac{1}{G}\sum_{g=1}^{G}\mathbf{m}_g\right),$$

where $\widehat{\Omega}$ is some positive definite moment weighting matrix and

$$\begin{aligned}
\mathbf{m}_g \;=\; & \begin{pmatrix} \mathbf{H}_{1,1g} \\ \vdots \\ \mathbf{H}_{1(J-1)g} \end{pmatrix} - \begin{pmatrix} m_1 \sum_{j=1}^{J}\sum_{j'=1}^{J} \alpha_{j'}\alpha_j \mathbf{H}_{2jj'g} \\ \vdots \\ m_{J-1} \sum_{j=1}^{J}\sum_{j'=1}^{J} \alpha_{j'}\alpha_j \mathbf{H}_{2jj'g} \end{pmatrix} - \begin{pmatrix} m_1 \mathbf{H}_{3g} \\ \vdots \\ m_{J-1}\mathbf{H}_{3g} \end{pmatrix} \\
& - \begin{pmatrix} m_1 \sum_{k=1}^{K}\sum_{k'=1}^{K} \gamma_k\gamma_{k'} \mathbf{H}_{4kk'g} \\ \vdots \\ m_{J-1} \sum_{k=1}^{K}\sum_{k'=1}^{K} \gamma_k\gamma_{k'} \mathbf{H}_{4kk'g} \end{pmatrix} + 2 \begin{pmatrix} m_1 \sum_{k=1}^{K} \gamma_k \mathbf{H}_{5kg} \\ \vdots \\ m_{J-1} \sum_{k=1}^{K} \gamma_k \mathbf{H}_{5kg} \end{pmatrix} + 2 \begin{pmatrix} m_1 \sum_{j'=1}^{J} \alpha_{j'} \mathbf{H}_{6j'g} \\ \vdots \\ m_{J-1} \sum_{j'=1}^{J} \alpha_{j'} \mathbf{H}_{6j'g} \end{pmatrix} \\
& -2 \begin{pmatrix} m_1 \sum_{j'=1}^{J}\sum_{k=1}^{K} \alpha_{j'}\gamma_k \mathbf{H}_{7j'kg} \\ \vdots \\ m_{J-1} \sum_{j'=1}^{J}\sum_{k=1}^{K} \alpha_{j'}\gamma_k \mathbf{H}_{7j'kg} \end{pmatrix} - \begin{pmatrix} \delta_1 \mathbf{H}_{8g} \\ \vdots \\ \delta_{J-1}\mathbf{H}_{8g} \end{pmatrix} + \begin{pmatrix} \sum_{j'=1}^{J} (\delta_1 \alpha_{j'} - A_{1j'}) \mathbf{H}_{9j'g} \\ \vdots \\ \sum_{j'=1}^{J} (\delta_{J-1} \alpha_{j'} - A_{(J-1)j'}) \mathbf{H}_{9j'g} \end{pmatrix} \\
& + \begin{pmatrix} \sum_{k=1}^{K} (\delta_1 \gamma_k - c_{1k}) \mathbf{H}_{10kg} \\ \vdots \\ \sum_{k=1}^{K} (\delta_{(J-1)} \gamma_k - c_{(J-1)k}) \mathbf{H}_{10kg} \end{pmatrix} - \begin{pmatrix} v_{1,0}\mathbf{H}_{11g} \\ \vdots \\ v_{(J-1),0}\mathbf{H}_{11g} \end{pmatrix}
\end{aligned}$$

is a $q(J-1)-$dimensional vector.