Keeping up with peers in India: A new social interactions model of perceived needs*

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 April 2017

Abstract

We propose a new nonlinear model of social interactions. The model allows point identification of peer effects as a function of group means, even with group level fixed effects. The model is robust to measurement problems resulting from only observing a small number of members of each group, and therefore can be estimated using standard survey datasets. We apply our method to a national consumer expenditure survey dataset from India. We find that each additional rupee spent by one's peer group increases one's own perceived needs by roughly 0.5 rupees. This implies that if I and my peers each increase spending by 1 rupee, that has the same effect on my utility as if I alone increased spending by only 0.5 rupees. Our estimates have important tax policy implications, since the larger these peer effects are, the smaller are the welfare gains associated with tax cuts or mean income growth.

1 Introduction

Identification of models with peer effects typically rely on either exogenous variation in group composition (Duflo, Dupas and Kremer, 2011; Carrell, Fullerton and West, 2009) or size (Lee, 2007), or detailed network data (Bramoullé, Djebbari and Fortin, 2009; de Giorgi, Frederiksen, and Pistaferri, 2016).[1] In this paper, we introduce a model with random or fixed

*We thank Niharika Singh, XX for their thoughts and comments on this work. We also thank seminar participants at XX. Norris is grateful for financial support from the Social Sciences and Humanities Research Council of Canada through its Doctoral Fellowship Awards.

[1] A smaller literature uses exogenous variation in group size for identification, though this can be seen as a sort of exogenous variation in group composition.

effects at the group level as well as peer effects in group means, but that can be estimated on cross-sectional data consisting of endogenously-formed and non-overlapping groups. We show how identification is a result of nonlinearities on the right hand side of the model, and demonstrate consistency even when n_g, the number of observations per group, is small.

Consumer demand estimation is a natural application of our model. Engel curves have long been known to be nonlinear (Deaton and Muellbauer, 1980), and there is evidence for peer effects in consumption (Boneva, 2013). In our keeping-up-with-the-Joneses-style model, one's perceived required expenditures, or "needs," depend on (among other things) the average expenditures of others in one's peer group. The higher are these perceived needs, the more one needs to spend to attain the same level of utility. In accordance with the empirical evidence in Luttmer (2005), we find that consumers lose utility from feeling poorer when their peers get richer. However, because our demand model is derived from an implicit utility function using the standard tools of consumer choice, we can go one step further and use the estimates to quantify the importance of peer effects in money-metric utility terms. We implement the model with consumer expenditure survey data from India, and find that each additional rupee spent by peers increases perceived needs by roughly 0.5 rupees. These results provide a structural explanation for the Easterlin (1974) paradox of low correlation between growth in aggregate incomes and growth in reported well-being: if perceived needs ratchet up along with incomes, aggregate income gains do less to improve utility than idiosyncratic income gains. It also implies that income or consumption taxes that apply to all group members are roughly half as expensive in terms of social welfare than one would calculate in the absence of peer effects.[2]

To fix ideas, a typical peer effect model relates outcome y_i for person i in group g with covariate x_i by

$$y_i = \overline{y}_g a + x_i b + u_i, \tag{1}$$

where u_i is an error term uncorrelated with the covariate, the pair (a, b) are parameters to estimate, and $\overline{y}_g$ is the population mean value of y_j over all people j in person i's peer group g (see, e.g., Manski 1993, 2000 and Brock and Durlauf 2001). In contrast, the type of model we consider has the form

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

where $\widehat{y}_g$ is an estimate of $\overline{y}_g$, the term v_g is a group level fixed or random effect, and ε_{gi} is an additional error that arises due to estimation error from the econometrician needing to

[2] We do not address whether there are similar consumption externalities for government-provided goods, which is necessary to understand how transferring money from consumers to the government affects utility.

replace $\overline{y}_g$ in the true model with an estimate $\widehat{y}_g$.

Equation (2) differs from equation (1) in three important ways. First, the model contains a group-level error v_g. We show identification of the model even if v_g is a group level fixed effect. In contrast, typical models like equation (1) cannot be identified in the presence of fixed effects, unless one has specialized data including observable network structures like "intransitive triads." See, e.g., Bramoullé, Djebbari, and Fortin (2009), Jochmans and Weidner (2016), and de Giorgi, Frederiksen, and Pistaferri, (2016).

Second, our model is nonlinear. In our empirical application, this nonlinearity is an unavoidable consequence of utility maximization. However, it is precisely this nonlinear structure that enables identification in the presence of fixed effects. In particular, the coefficient a can be identified from the $x_i\widehat{y}_g$ interaction in the quadratic term of equation (2), which is not eliminated when we first difference to remove v_g.

Third, because we only have survey data with a modest number of observations for each group, we do not assume we can observe the true $\overline{y}_g$ even asymptotically. We therefore replace $\overline{y}_g$ with its estimate $\widehat{y}_g$, and this introduces the additional error term ε_{gi} that is correlated with $\overline{y}_g a + x_i b$ and its square. Part of the novelty of our methodology comes from overcoming these correlations to construct valid moment conditions used for GMM estimation of the model. Our model is potentially applicable to many contexts with nonlinear peer effects, and may be of particular use when the researcher only observes a relatively small number of members of each peer group (for example, with typical government survey data).

Our analysis proceeds as follows. In the next subsections we review the literature on peer effects and needs, and show how our model incorporates peer effects and needs into demand functions.

After a literature review, in Section 2 we prove identification of the above generic peer effects model and provide associated estimators. In Section 3 we apply the generic model to a single annual cross section of Indian National Sample Survey (NSS) data, taking y_i to be expenditures on luxuries and x to be total expenditures. These results show peer effects are present, but do not relate them to utility. We also analyze the effects of own and peer expenditures on answers to a life satisfaction question (which we interpret as a crude proxy for utility) from a separate data set. Taken together, these preliminary analyses indicate that increased luxury expenditures by one's peers increases one's own luxury spending and reduces one's own level of reported utility.

Both economic theory and these non-structural empirical results results motivate our construction of a structural model in which needs depend on the spending of one's peers. This model is described in Section 4. Exploiting revealed preference theory, in Section 5 we derive quantity demand functions associated with this utility model. These demand functions have a similar structure to equation (2), though y_i and $\widehat{y}_g$ become quantity vectors instead of scalars, the parameter a is replaced by a matrix of own and cross equation peer effects, and what appear as constant parameters above are replaced with nonlinear functions of prices and observable demand shifters. Given this analogous structure to our generic model, we prove identification of these demand functions using the same techniques as before, and we provide an associated estimator.

In Section 6 we implement this structural demand model and provide associated welfare analyses, now using multiple annual NSS annual cross-sections of household-level expenditure data. Section 7 concludes. Throughout, we relegate formal derivations and proofs to the appendix.

1.1 Peer Effects in Consumption and Needs

There is a long literature that connects utility and well-being to peer income or consumption levels (see, e.g., Frank 1999, 2012). The Easterlin (1974) paradox asserts an empirical connection between well-being and national average incomes. Though the strength of this connection is debated (Stevenson and Wolfers 2008), the correlation between utility and national-level consumption, ceteris paribus, is negative. Ravina (2008) and Clark and Senik (2010) regress self-reported utility on own budgets and national average budgets, and other correlated aggregate measures like inequality, and find that the negative correlation still stands. Similar results hold for much smaller reference groups; Luttmer (2005) finds that an increase of the average income in one's neighbors reduces self-reported well being.

The possible mechanisms for this are varied. Veblen (1899) effects make consumers value consumption of visible status goods. Reference-dependent utility functions hinge preferences on own-endowments (Tversky and Kahneman 1979). More recent work on these models has led to reference-dependence that is "other-regarding," where utilities depend on reference points that are driven by other agents' decisions or endowments. Models of "keeping up with the Joneses" have one's own consumption feel smaller when one's peers consume more. Surveys of these literature include Kahneman (1992) and Clark, Frijters, and Shields (2008).[3] In our paper, we will model the consumption of our peers as affecting what we perceive as our consumption "needs."

A more recent literature connects consumption choices to peer consumption levels, al-

[3] A smaller literature in macro focuses on intertemporal models of aggregate behaviour, 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). De Giorgi, Frederiksen and Pistaferri (2016) are an interesting middle ground, investigating how intertemporal allocations for individuals depend on peer consumption.

though these analyses are essentially nonstructural. For example, Chao and Schor (1998) regress individual cosmetics spending on group-average cosmetics spending and find positive responses, linking their findings to Veblen effects. Boneva (2013) regresses household quantity demand vectors on household budgets (total expenditures) and on the average budgets of reference groups, using the randomized Progresa rollout to instrument for group averages. De Giorgi, Frederiksen and Pistaferri (2016) show that consumption choices depend on peer consumption levels, using neighbors-of-neighbors as instruments. All these papers suggest that the magnitudes of peer effects in consumption choices are large.

Taken together, these results suggest that a structural model of peer effects in consumption choices should start with a utility function that depends both on own-consumption and on peer consumption. That is, direct utility depends on $\mathbf{q}_i$ and $\overline{\mathbf{q}}_g$, where g denotes the peer group of consumer i, $\mathbf{q}_i$ is the vector of quantities of goods consumed by consumer i, and $\overline{\mathbf{q}}_g$ is the mean consumption vector of all consumer's in group g. This in turn implies indirect utility functions of the form

$$u_i = \widetilde{V}(\mathbf{p}, x_i, \mathbf{z}_i, \overline{\mathbf{q}}_g),$$

where u_i is utility, or well-being of consumer i, $\mathbf{p}$ is the vector of prices of goods, $\mathbf{z}_i$ is a vector of characteristics that affect tastes, and $x_i = \mathbf{p}'\mathbf{q}_i$ is consumer i's budget, or total expenditures.

In our model, peer consumption $\overline{\mathbf{q}}_g$ affects needs. In the context of utility and cost functions, "needs" are fixed costs, representing the minimum quantity vector one requires to start getting utility. The idea that preferences have fixed costs that need to be met before expenditures start increasing utility goes back to Samuelson (1947). Samuelson defined the quantity vector $\mathbf{f}_i$ as the "necessary set" of goods. The cost of buying these necessary goods, i.e. needs, is $\mathbf{p}'\mathbf{f}_i$. Samuelson then defined $x_i - \mathbf{p}'\mathbf{f}_i$ as "supernumerary income," one's remaining income after subtracting off the cost of these needs. Utility is then obtained by spending supernumerary income.

The classic 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 structure for dealing with heterogeneity in needs is typically used to account for demographic characteristics z.

In Gorman's model, utility depends on x_i only through the term $x_i - \mathbf{p}'\mathbf{f}_i$, where $\mathbf{f}_i = \mathbf{f}(\mathbf{z}_i)$. Blackorby and Donaldson (1994) show that models of this type have a desirable property for social welfare calculations, which they call Absolute Equivalence Scale Exactness (AESE).[4] In

[4] Blackorby and Donaldson (1994) start from a more general model where utility has the form $u_i = h\left[V\left(\mathbf{p}, x_i - H\left(\mathbf{p}, \mathbf{z}_i\right)\right), \mathbf{z}_i\right]$ for some functions h and H. In the Gorman model context where $H\left(\mathbf{p}, \mathbf{z}_i\right) = \mathbf{p}'\mathbf{f}_i$,

particular, changes in the cross-population sum of income - needs (also known as equivalent income), $\sum_i x_i - \mathbf{p}'\mathbf{f}_i$, are a dollar measure of changes in social welfare. Increases in societal income are thus straightforward to turn into utility terms; a technology change that raises everyone's after tax income by 10% is offset by the associated change in $\mathbf{p}'\mathbf{f}_i$, reflecting the social cost of keeping up with the Joneses.

While our goal consists of quantifying these costs of keeping up with peers, it should be noted that these costs are not necessarily all wasted resources. For example, positive externalities such as arise from widespread internet access would also appear as "needs" in our model. However, we will show evidence from direct regressions of self-reported well-being on group expenditure that indicates that the relationship is negative, at least on average. In any case, it is useful to quantify these costs of peer effects, and separate them from the direct impacts of income increases on consumer's utility.

Our model begins with Blackorby and Donaldson, but then adds peer effects by including $\overline{\mathbf{q}}_g$ in the needs vector $\mathbf{f}_i$, so

$$u_i = V\left(\mathbf{p}, x_i - \mathbf{p}'\mathbf{f}_i\right) \quad \text{with} \quad \mathbf{f}_i = \mathbf{f}(\mathbf{z}_i, \overline{\mathbf{q}}_g) \tag{3}$$

For simplicity we take the function $\mathbf{f}$ to be linear, so

$$\mathbf{f}_i = \mathbf{A}\overline{\mathbf{q}}_g + \mathbf{C}\mathbf{z}_i \tag{4}$$

for some matrices of parameters $\mathbf{A}$ and $\mathbf{C}$.

We do not assume utility u_i is observable, but if it were, then one could directly estimate equation (3). Luttmer (2005) estimates a model where self reported well-being (on a 1 to 7 scale) is a function of $x_i - \delta\mathbf{p}'\overline{\mathbf{q}}_g$. If we interpret this discrete self reported well-being as a crude measure of utility u_i, then Luttmer's model corresponds to a special case of our model in which the matrix $\mathbf{A}$ equals the scalar δ times the identity matrix.

Our main structural analysis does not assume utility is observable. We instead derive quantity demand equations from (3) that express demand as a function of observables only, allowing us to back out the parameters $\mathbf{A}$ and $\mathbf{C}$. Specifically, applying Roy's (1947) identity to equation (3) yields demand functions of the form

$$\mathbf{q}_i = \mathbf{g}(\mathbf{p}, x_i - \mathbf{p}'\mathbf{f}_i) + \mathbf{f}_i \tag{5}$$

their results imply that if h is independent of z_i, then $x - \mathbf{p}'\mathbf{f}_i$ will be a money metric for utility, and all transformations h that are independent of z_i result in money metrics that are not additive in $\mathbf{p}'\mathbf{f}_i$. They also derive results on identification associated with this model, showing when social welfare functions can be constructed based on differences between budgets and needs.

where the functional form of the vector valued function $\mathbf{g}$ depends only on the functional form of V. We will identify and estimate demand functions given by equations (5) and (4), with the addition of error terms that also include random or fixed effects.

The link between this demand model and the simpler peer effects model described earlier can be seen by replacing the simple linear term $x_i + \overline{y}_g a$ from before with $x_i - \mathbf{p}' \left(\mathbf{A}\overline{\mathbf{q}}_g + \mathbf{C}\mathbf{z}_i \right)$ and taking $\mathbf{g}$ to be quadratic in this term. As in the simpler peer effects model, we will need to replace $\overline{\mathbf{q}}_g$ with an estimate $\widehat{\mathbf{q}}_g$, and deal with the same identification and estimation issues discussed earlier.

The class of demand specifications given by equations (5) has a shape invariance in quantities property described by Pendakur (2005). It has the testable implication that, for any price vector p, the quantity demand curve associated with any good j has the same shape across consumers, differing only by horizontal and vertical translations (in x and q, respectively). This property is analogous to the more well known shape invariance in budget shares popularized by Pendakur (1999), Blundell, Chen and Kristensen (2007) and Lewbel (2010). A long literature in empirically modeling consumer demand shows that demands are nonlinear and well-approximated by polynomials (see, e.g., Deaton and Muellbauer 1980, Banks, Blundell and Lewbel 1998, and Blundell, Chen and Kristensen 2007), mirroring our identification conditions and modelling assumption.

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, originally described by Manski (1993, 2000), and expanded on by Brock and Durlauf (2001), and Blume, Brock, Durlauf, and Ioannides (2010). Our model has a specific behaviorally derived structure that overcomes the reflection problem.

In some peer effects models, network information is available that can help identification. For example, Bramoullé, Djebbari, and Fortin (2009) show identification of peer effects in social networks with certain types of interconnections called intransitive triads, essentially using data from friends of friends as instruments. Devezies et. al, (2006) and Lee (2007) use variation in group sizes to aid identification. In our context, making use of standard consumption survey data, we do not have any information on who friends of friends are, and variation in group size also 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.[5]

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, exogenous 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 assume only that players are basing their behavior on the true group means. This 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 show how all of that could be identified in the presence of all of the other obstacles to identification that we face. These obstacles include the reflection problem, only observing a small number of members of each group, 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.

[5] Furthermore, identification from variation in group sizes generally has power only for relatively small groups like classrooms. In our context with groups containing thousands of individuals it is unlikely that this approach would work even if we had outside knowledge of true group sizes.

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 most 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.

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 students 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 identifiable, 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 obstacles to identification in our general context, to show how we overcome these obstacles, 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, and where only a small number of individuals are observed in each group.

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 $G \rightarrow \infty$. In our data we will only observe a small number n_g of the individuals in each peer group g, so asymptotics assuming $n_g \rightarrow \infty$ would be a poor approximation. 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

$$y_i = h\left(\theta \mid \overline{y}_g, x_i\right) + v_g + u_i, \tag{6}$$

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 the peer effects we want to identify. Note that $\overline{x}_g$ does not appear explicitly in this model, but, 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[6]

$$h(\theta \mid \widehat{y}_g, x_i) = \left(\overline{y}_g a + x_i b\right)^2 d + \left(\overline{y}_g a + x_i b\right), \tag{7}$$

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 $\varepsilon_{gi} = h\left(\theta \mid \overline{y}_g, x_i\right) - h(\theta \mid \widehat{y}_g, x_i)$, and the model becomes

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

where

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

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 can 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 employ 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.

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

[6]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

$$y_i - y_{i'} = h\left(\theta \mid \widehat{y}_{g,-ii'}, x_i\right) - h\left(\theta \mid \widehat{y}_{g,-ii'}, x_{i'}\right) + u_i - u_{i'} + \varepsilon_{gi} - \varepsilon_{gi'}. \tag{8}$$

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

$$E\left(u_i - u_{i'} + \varepsilon_{gi} - \varepsilon_{gi'} \mid x_i, x_{i'}\right) = 0, \tag{9}$$

which we can then use to construct moments for estimation of equation (8).

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 (8) contains functions of $\widehat{y}_{g,-ii'}$, x_i, and $x_{i'}$ as regressors, and equation (9) shows that we can use functions of x_i and $x_{i'}$ as instruments (equivalently, x_i and $x_{i'}$ are exogenous regressors). An obvious candidate instrument for $\widehat{y}_{g,-ii'}$ 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 other 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 a different data set.

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 for $\widehat{y}_{g,-ii'}$, constructed as above either from other datasets or from other time periods. Combining these with equations (8) and (9) 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

$$E\left[\left(y_i - y_{i'} - h\left(\theta \mid \widehat{y}_{g,-ii'}, x_i\right) + h\left(\theta \mid \widehat{y}_{g,-ii'}, x_{i'}\right)\right)\mathbf{r}_{gii'}\right] = 0. \tag{10}$$

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 (7) for the function h into equation above, 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).

In addition to extending the above model to allow for a vector of covariates $\mathbf{x}_i$, in the Appendix we also show how the model extends to a J vector of outcomes $\mathbf{y}_i$, replacing the scalar a with a J by J matrix of own and cross equation peer effects. Our utility-derived demand model will also entail a vector of outcomes with a matrix of own and cross peer effects.

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 additional moments that do not entail differencing. The moments obtained under fixed effects remain valid under the additional random effects assumptions. 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 moments that do not require differencing can be obtained by exploiting the random effects independence of v_g from x_i.

For random effects it will be convenient to rewrite the quadratic model, equations (6) and (7), as

$$y_i = \overline{y}_g^2 a^2 d + (a + 2x_i abd)\,\overline{y}_g + \left(x_i b + x_i^2 b^2 d\right) + v_g + u_i. \tag{11}$$

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 are not differencing, we must 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 (11) 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 $E(\widetilde{\varepsilon}_{gii'} | x_i, \mathbf{r}_g) = -da^2 Var\,(v_g)$. Our constructions in estimating the group mean eliminates correlation of the error $\widetilde{\varepsilon}_{gii'}$ with x_i. But $\widetilde{\varepsilon}_{gii'}$ still does not have conditional mean zero, because both $\widehat{y}_{g,-ii'}$ and $y_{i'}$ contain v_g, so the mean of the product of $\widehat{y}_{g,-ii'}$ and $y_{i'}$ includes the variance of v_g.

It follows from the above that

$$E\left[y_i - \widehat{y}_{g,-ii'}y_{i'}a^2d - (a + 2x_iabd)\,\widehat{y}_{g,-ii'} - \left(x_ib + x_i^2b^2d\right) - v_0 \mid x_i, \mathbf{r}_g\right] = 0 \qquad (12)$$

where $v_0 = E(v_g) - da^2Var(v_g)$ 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

$$E\left[\left(y_i - \widehat{y}_{g,-ii'}y_{i'}a^2d - (a + 2x_iabd)\,\widehat{y}_{g,-ii'} - \left(x_ib + x_i^2b^2d\right) - v_0\right)\mathbf{r}_{gi}\right] = 0 \qquad (13)$$

which we can estimate using GMM exactly as before. The moments from the fixed effects model, equation (10), remain valid under random effects, so both equations (10) and (13) could be combined in a single GMM estimator to increase asymptotic efficiency.

As with the fixed effects model, in the Appendix we extend the above model to allow for a vector of covariates $\mathbf{x}_i$, and to allow for a J vector of outcomes $\mathbf{y}_i$, replacing the scalar a with a J by J matrix of own and cross equation peer effects.

3 Nonstructural Analysis: Well-Being, Consumption and Luxuries

Do group level peer effect consumption externalities exist? Do they make people worse off? Before developing our full utility derived structural model, we present some non-structural empirical findings addressing these questions. The first analysis applies the generic model of the previous sections to our India data, and shows that peer effects are present in the consumption of luxuries. The second, using data from a separate India data set, is a simple regression of self reports of well-being on own and peer expenditures. We find that higher peer income is associated with lower subjective well-being, indicating that peer effects have the sign that our theory predicts.

3.1 National Sample Survey data

For our main analyses, we use household consumption data from rounds 59 to 62 of the National Sample Survey (NSS) of India (conducted in 2003 to 2006). Table 1 gives data on household consumption from round 61 of the NSS. We consider only households that are between the 1st and 99th percentiles of household expenditure in each state/year. We use only urban households whose state-identifier is not masked and with 12 or fewer members whose head is aged 20 or more. We define groups as the 3-way cross of: education of household head (in 3 levels: uneducated/illiterate; completed primary; completed secondary); religion/caste (in 3 levels: scheduled tribe/scheduled caste Hindu; non-SC/ST Hindu; non-Hindu[7]); and geographic district (575 districts across 33 states). Our estimation procedure requires at least 3 members in each group, so we drop groups that have fewer than 3 households. Our resulting dataset has 70,217 distinct households in 3259 groups, giving an average group size of less than 22 households. Our estimator uses all household-pairs within each group, and we have a total of 3,009,614 such pairs. We provide summary statistics at the level of the household, and at the level of the household-pairs used for estimation.

The NSS collects item-level household spending for 76 items, and collects quantities for roughly half of these. We consider only the 48 nondurable consumption items, and compute total expenditure x_i as the sum of spending on these nondurable consumption items. We automate the classification of items into luxuries versus necessities by regressing the budget shares of each of these 48 nondurable items on the log of total expenditure, and classify those items with positive slopes as luxuries and the rest as necessities. Total expenditures, and its components of luxury and neccessity spending, are expressed in units of average household expenditure in round 59, so the average total expenditure of 1.15 reported in Table 1 shows that household spending was 15% higher in round 61 than in round 59. Our later analyses make use of all four rounds of data. Round 61 was chosen for the preliminary analysis because it was the round with most observations. Roughly one-quarter of household spending is classified as luxury spending (0.30/1.15). Prices are constructed from unit values at the item level by taking the median at the state-round level, then aggregated up to the level of luxuries and necessities with a Laspeyres index. We also tried using district-level prices, and found similar results.

3.2 Generic Model Estimates

Our first empirical exercise is to use the round 61 NSS data described in Table 1 to estimate the fixed and random effects models of the previous sections. Here, y_i is expenditures on luxuries, $\overline{y}_g$ is the true group-mean expenditure on luxuries, $\widehat{y}_g$ is the observed sample average, and x_i is total expenditures.

We provide estimates using random-effects unconditional moments (13) and fixed-effects unconditional moments (10). Define $\overline{x}_{g,-t}$ to be the group-average expenditure in other time periods (that is, average expenditures in group g in rounds 59, 60 and 62). Fixed-effects in-

[7] We note that in India, non-Hindus are primarily Muslim.

struments $\mathbf{r}_{gii'}$ are: $\overline{x}_{g,-t}, (x_i - x_{ii'}), (x_i - x_{ii'})\overline{x}_{g,-t}, (x_i^2 - x_{ii'}^2), (z_i - z_k), (z_i - z_k)\overline{x}_{g,-t}, z_g, z_g(x_i - x_{i'}), 1$. Random-effects instruments r_{gi} are: $\overline{x}_{g,-t}, x_i, x_i\overline{x}_{g,-t}, x_i^2, z_i, 1$. GMM estimates of the parameters are given in Table 3.

In the random effects model, the coefficients are relatively stable no matter what controls we include. In contrast, the fixed effects model estimates are less stable, varying more across different sets of controls. As one would expect, the random effects model has lower standard errors, however, a Hausman test comparing the two models decisively rejects the random effects model (controlling for group characteristics), so random effects estimates may be biased.

In both models, the estimated values of b and d are positive. As a result, the first and second derivatives of luxury consumption with respect to total expenditures x_i are positive, which is sensible for luxury goods.

The estimated peer effects coefficient a is moderate-sized under random effects, and larger with fixed effects. In both models, peer effects have the sign we expect. A 100 rupee increase in peer group luxury expenditures makes people behave as if they are around 20 rupees poorer in the random effects model, and over 50 rupees poorer in the fixed effects model (again controlling for group level characteristics).

The overall effect of group-average luxury spending is jointly determined by all the parameters, so we calculate the summary measure in $\partial y/\partial \overline{y}_g$. This is the estimated derivative of luxury demand with respect to group-average luxury spending, evaluated at the average total expenditure level (1.15, shown in Table 1). This derivative is positive as expected in both the random effects and fixed effects models, though the point estimate is not statistically significant in the fixed effects model.

While the results here are consistent with our theoretical model, this analysis has several shortcomings. First, it only shows how peer's spending affects one's own spending, but it cannot tell us if these spillovers are bad in the sense of lowering one's utility when one's peers spend more (the results do suggest this is the case, since they show that one acts as if one is poorer when one's peers spend more). Second, although we control for prices by including them as covariates, the model does not do so in a way that is consistent with utility maximization, because the model is not derived from utility theory. Third, the model does not allow for the possibility that group-average *non*-luxury spending affects luxury demands. And fourth, it is not possible to derive welfare implications or interpretations of the resulting estmates.

In order to address the first of these issues, and to provide additional guidance for constructing a formal model of utility that solves all the other shortcomings, we now turn to a brief analysis of well-being data from a different survey. Dealing with the remaining issues will require our full structural model.

3.3 Subjective well-being and peer consumption

Our generic model estimates above are consistent with a model in which increased peer consumption decreases the utility one gets from consuming a given level of luxuries, as suggested by our theoretical model of needs. But it's possible that increased peer consumption could increase rather than decrease one's own utility, e.g., through network effects (as when peers owning cell phones increases the utility of one's own cell phone). That is, we cannot be sure from the generic model that the spillover effects of peer expenditures constitute negative rather than positive externalities.

To directly check the sign of these spillover effects on utility, we would like to estimate the correlation between utility and peer expenditures, conditioning on one's own expenditure level. While we cannot directly observe utility, here we make use of a proxy, which is a reported ordinal measure of life satisfaction.

Table 2 summarizes 3236 observations from the 5^{th} (2006) and 6^{th} (2014) waves of the World Values Survey. These are the two waves with the most consistent income reporting. In each year the surveyor asks the question, "All things considered, how satisfied are you with your life as a whole these days?"Answers are on a 5-point ordinal scale in the 5^{th} wave, and a 10-point scale in the 6^{th}, which we collapse to a 5-point scale.

Neither wave of the survey reports actual income or consumption expenditures. What this survey does report is position on a ten-point income distribution that corresponds to the deciles of the national income distribution. We use this response to impute individual total expenditure levels by taking the corresponding decile-specific expenditure mean from the NSS data. We also obtain group level total expenditures from the NSS data. For this analysis we define groups by religion (Hindu vs non-Hindu), education level (less than primary, primary, secondary or more) and state of residence (20 states and state groupings). These are much larger, more coarsely defined groups than we use for all of our other analyses. Much larger groups are needed here because the WVS sample size is much smaller than the NSS, and because we have no asymptotic theory to deal with small group sizes in this part of the analysis.

Our measures of total expenditures are deflated using the CPI index for India; in the Appendix we show analogous regressions with both nominal expenditure and Laspeyres-deflated expenditure. Average expenditure is 2200 rupees per month (which deflates to 1999 rupees), or about 50 US dollars. This is lower than the average for India at this time, which appears to be due to sample composition issues in the WVS. For example, only 1.6% of households in the WVS are in the top decile of income.

Table 4 presents estimates of well-being as a function of both own total expenditures and group total expenditures, specified as

$$u_i = \beta_1 \widehat{x}_{igt} + \beta_2 \overline{\widehat{x}}_{gt} + Z_{igt}\alpha + \gamma_g + \phi_t + \varepsilon_{igt}, \tag{14}$$

where u_i is the z-normalized well-being indicator, $\widehat{x}_{igt}$ is imputed individual expenditures, $\overline{\widehat{x}}_{gt}$ is imputed group expenditures, Z_{igt} is vector of individual level controls, γ_g is a group level fixed effect (recall that groups are defined within states, so this effectively includes a state fixed effect as well), and ϕ_t is a year fixed effect. We also repeat this analysis using an ordered logit specification.

Results in the first column of Table 4 imply that a 100 rupees increase in individual expenditures $\widehat{x}_{igt}$ increases satisfaction by 0.13 standard deviations, while a 100 rupees increase in group expenditure $\overline{\widehat{x}}_{gt}$ *decreases* satisfaction by 0.19 standard deviations. The ratio of these effects, $-\beta_1/\beta_2 = 13/19 = 0.68$, says that one must increase one's own expenditures by 68 rupees to compensate for the loss of utility that results from a 100 rupees increase in group expenditure levels. This is similar to Luttmer's (2005) finding of "neighbours as negatives,"where increases in group income holding individual income constant reduces individual's reported well-being. The roughly comparable estimate in Luttmer (2005) is 0.76.

These estimates suggest that group average consumption is negatively related to individual utility, consistent with our theory. Our theory will also imply magnitudes less than one (like the .68 found here), because, as we show later, magnitudes greater than one can result in no equilibrium existing.

Since well being is reported on an ordinal scale, to check the robustness of these results, we estimate the same regression as an ordered logit (see Column 5 of Table 4). The results are qualitatively the same, showing that our results are not being determined by the normalizations implicit in z-scoring the satisfaction responses.

We next consider adding an interaction term $\beta_3 \overline{\widehat{x}}_{gt} \widehat{x}_{igt}$ to equation (14). This is relevant both for the specification of our structural model, and in economic terms, since it would matter if, e.g., utility were more responsive to group expenditure for rich people than for poor people. In columns 2 and 6 of Table 4, we see no evidence that this interaction is important.

Taken together, these regressions suggest that utility is increasing in household expenditure, decreasing in group average expenditure, and that the magnitude of the latter effect is independent of the level of household expenditure. This suggests validity of the absolute equivalence-scale exactness assumption from Blackorby and Donaldson (1994), which we discuss in the next section and employ in our structural analysis later.

The final sets of regressions in Table 4 look at whether or not spillovers are the same across religious groups and across education groups. Consider first β_2, the magnitude of spillovers for Hindus versus non-Hindus. Here we see mild evidence that the effects are larger in magnitude for non-Hindus than for Hindus. Turning to education, we see some evidence that uneducated people and more educated people have smaller estimated effects than do people with primary education. Because religion and education are correlated and because the sample size here is relatively small, multicollinearity prevents us from learning much from a model that includes both religion and education simultaneously.

The relative effects of group vs individual expenditures, $-\beta_1/\beta_2$, is quite imprecisely estimated for all subgroups. The difference in this value between Hindus and nonHindus is not statistically significant. Across education groups, we see weak evidence that the lowest education group has the smallest effect, and the middle education group has the largest effect.

Taken together the nonstructural analyses provided here and in the previous subsection provide empirical evidence supporting the theory that underlies our formal, structural analysis. The next step is to construct a model of utility that is consistent with what we observe here, accomodates price and demographic heterogeneity, and enables analyses of the welfare implications of peer effects. We will then prove identification of, and subsequently estimate, the demand functions for goods that is derived from this model of utility.

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

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, let x_i be the total budget for commodities of consumer i, and let $\mathbf{z}_i$ be a K vector of observed characteristics 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 $\mathbf{p}'\mathbf{q}_i \leq x_i$. 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.

As derived in section 1.1, we assume preferences can be represented by an indirect utility function of the form

$$u_i = V(\mathbf{p}, x_i - \mathbf{p}'\mathbf{f}_i) \quad \text{with} \quad \mathbf{f}_i = \mathbf{A}\overline{\mathbf{q}}_g + \mathbf{C}\mathbf{z}_i \tag{15}$$

for some J by J parameter matrix $\mathbf{A}$ and some J by K parameter matrix $\mathbf{C}$. 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 may 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 expect 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.

For welfare calculations, we need to compare well being across consumers. Define the *equivalent-income* $\widetilde{X}_i$ 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. As discussed earlier, equation (15) is in the class of models that satisfy Blackorby and Donaldson's (1994) Absolute Equivalence Scale Exactness (AESE) property. It follows from their results that $\widetilde{X}_i$ itself cannot be identified, but differences $\widetilde{X}_i - \widetilde{X}_{i'}$ for any two individuals i and i' are given by $\widetilde{X}_i - \widetilde{X}_{i'} = x_i - x_{i'} - \mathbf{p}'(\mathbf{f}_i - \mathbf{f}_{i'})$. A feature of preferences satisfying AESE is that equivalent incomes equal money metric measures of utility, and therefore social welfare functions can be defined as functions of everyone's equivalent incomes $\widetilde{X}_i$. A particularly convenient social welfare function (though one that is not inequality averse) is the simple sum $\sum_i \widetilde{X}_i$. Given estimates of $\mathbf{f}_i$, we can therefore calculate changes in this social welfare function from changes in mean income in the population and corresponding changes in needs $\mathbf{p}' \sum_i \mathbf{f}_i$.

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) shows that any polynomial demand system will have a maximum rank of three, and Lewbel (1989) shows that the simplest tractable class of indirect utility functions that yields rank three polynomials in x is $V(\mathbf{p}, x) = (x - R(\mathbf{p}))^{1-\lambda} B(\mathbf{p}) / (1-\lambda) - D(\mathbf{p})$ for some constant λ and some differentiable functions R, B and D.

The most commonly assumed rank three models are quadratic (see the above references, and Pollak and Wales 1980), which corresponds to $\lambda = 2$, implying $V(\mathbf{p}, x) = -(x - R(\mathbf{p}))^{-1} B(\mathbf{p}) - D(\mathbf{p})$. Combining this with AESE and our parameterization of the needs $\mathbf{f}_i$ gives the model

$$u_i = V(\mathbf{p}, x_i - \mathbf{p}'\mathbf{f}_i) = -\left(x_i - R(\mathbf{p}) - \mathbf{p}'\mathbf{A}\overline{\mathbf{q}}_g - \mathbf{p}'\mathbf{C}\mathbf{z}_i\right)^{-1} B(\mathbf{p}) - D(\mathbf{p})$$

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

Applying Roy's (1947) identity to this indirect utility function then yields the vector of demand functions

$$\mathbf{q}_i = \left(x_i - R(\mathbf{p}) - \mathbf{p}'(\mathbf{A}\overline{\mathbf{q}}_g + \mathbf{C}\mathbf{z}_i)\right)^2 \frac{\nabla D(\mathbf{p})}{B(\mathbf{p})} + \left(x_i - R(\mathbf{p}) - \mathbf{p}'(\mathbf{A}\overline{\mathbf{q}}_g + \mathbf{C}\mathbf{z}_i)\right) \frac{\nabla B(\mathbf{p})}{B(\mathbf{p})} + \nabla R(\mathbf{p}) + \mathbf{A}\overline{\mathbf{q}}_g + \mathbf{C}\mathbf{z}_i. \tag{16}$$

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 or random 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 $l \in g$. In the fixed effects model, the group level fixed effect $\mathbf{v}_g$ is permitted to correlate with other regressors like $\mathbf{p}$ and $\overline{\mathbf{q}}_g$. We also consider a random effects model, where much greater efficiency is gained by adding the restrictions that $\mathbf{v}_g$ satisfies some independence assumptions.

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.

Convenient yet flexible specifications of the price functions are $R(\mathbf{p}) = \mathbf{p}^{1/2\prime}\mathbf{R}\mathbf{p}^{1/2}$ where $\mathbf{R}$ is a symmetric matrix, $\ln B(\mathbf{p}) = \mathbf{b}' \ln \mathbf{p}$ with $\mathbf{b}'\mathbf{1} = 1$, and $D(\mathbf{p}) = \mathbf{d}' \ln \mathbf{p}$ with $\mathbf{d}'\mathbf{1} = 0$. Substituting these error terms and price functions into the above demand function yields, for each good j, the demand model

$$q_{ji} = Q_j\left(\mathbf{p}, x_i, \overline{\mathbf{q}}_g, \mathbf{z}_i\right) + v_{jg} + u_{ji}$$

where

$$Q_j\left(\mathbf{p}, x_i, \overline{\mathbf{q}}_g, \mathbf{z}_i\right) = X_i^2 e^{-\mathbf{b}' \ln \mathbf{p}} \frac{d_j}{p_j} + X_i \frac{b_j}{p_j} + R_{jj} + \sum_{k \neq j} R_{jk} \sqrt{p_k / p_j} + \mathbf{A}_j' \overline{\mathbf{q}}_g + \mathbf{C}_j' \mathbf{z}_i, \tag{17}$$

$\mathbf{A}_j'$ is row j of $\mathbf{A}$, $\mathbf{C}_j'$ is row j of $\mathbf{C}$, and where we define for convenience *deflated income*

$$X_i = X(\mathbf{p}, x_i, \overline{\mathbf{q}}_g, \mathbf{z}_i) = x_i - \mathbf{p}^{1/2\prime}\mathbf{R}\mathbf{p}^{1/2} - \mathbf{p}'\mathbf{A}\overline{\mathbf{q}}_g - \mathbf{p}'\mathbf{C}\mathbf{z}_i. \tag{18}$$

Deflated income X_i is convenient for simplifying notation, but does not have the welfare significance of equivalent income $\widetilde{X}_i$. However, if prices are held fixed, then $\widetilde{X}_i - \widetilde{X}_{i'} = X_i - X_{i'}$, so we can use either one interchangeably for calculating changes in needs or in money metric utility.

In all of the above, different consumers will in general be observed in different time periods. Prices vary by time, and also vary geographically. Assume that our data spans T different price regimes (time periods and/or geographic regions). Each individual i is observed in some particular price regime $t \in \{1, 2, ..., T\}$, and we add a t subscript to every group level variable above.

The goal will be estimation of the set of parameters $\{\mathbf{A}, \mathbf{C}, \mathbf{R}, \mathbf{d}, \mathbf{b}\}$. Our particular interest will be in identifying equivalent income $\widetilde{X}(\mathbf{p}_t, x_i, \overline{\mathbf{q}}_{gt}, \mathbf{z}_i)$, which forms the basis of our welfare analyses as discussed in the previous section. The only parameters the function $\widetilde{X}(\mathbf{p}_t, x_i, \overline{\mathbf{q}}_{gt}, \mathbf{z}_i)$ depends on are $\mathbf{A}$ and $\mathbf{C}$. Interestingly, under random effects $\mathbf{A}$ and $\mathbf{C}$ can be identified even if the data contain no price variation. In our empirical application, some of the characteristics $\mathbf{z}_i$ are group level attributes, that is, they vary across groups but are the same for all individuals within a group. Where it is relevant to make this distinction, we write $\mathbf{C}$ as $\mathbf{C} = \left(\widetilde{\mathbf{C}} : \mathbf{D}\right)$ for submatrices $\widetilde{\mathbf{C}}$ and $\mathbf{D}$, and replace $\mathbf{C}\mathbf{z}_i$ with $\mathbf{C}\mathbf{z}_i = \widetilde{\mathbf{C}}\widetilde{\mathbf{z}}_i + \mathbf{D}\widetilde{\mathbf{z}}_g$, where $\widetilde{\mathbf{z}}_i$ is the vector of characteristics that vary across individuals in a group and $\widetilde{\mathbf{z}}_g$ are group level characteristics. Under fixed effects, $\mathbf{A}$ and $\widetilde{\mathbf{C}}$, but not $\mathbf{D}$, can be identified even when the data contains no price variation.

There is one more extension to the above model that we consider in our estimates, but do not include above to save on notation. We allow a few discrete characteristics (religion/caste and education) to interact with $\overline{\mathbf{q}}_g$. This is equivalent to letting $\mathbf{A}$ vary with these discrete characteristics. Identification of the model with this extension follows immediately from identification of the model with $\mathbf{A}$ constant, since the the same assumptions used to identify the above model with fixed $\mathbf{A}$ can just be applied separately for each value of these characteristics.

Our demand model is the system of equations given by plugging equation (18) into equation (17) for all goods j. In the Appendix we provide assumptions and associated proofs that all of the parameters of the model, $\mathbf{A}$, $\mathbf{C}$, $\mathbf{R}$, $\mathbf{d}$, and $\mathbf{b}$, are identified under either fixed effects or random effects. The methods of identification and proofs proceed in two steps. First, we consider the model without price variation, constructing the Engel curves that correspond to our demand system. Using techniques entirely analogous to the corresponding generic model (in which moments are constructed that can be used for estimation via GMM), we show identification of these Engel curves. We then show that, by identifying the Engel curve parameters in some different price regimes, we can with sufficient price variation recover all of the parameters of the full demand model. For estimation, it is not necessary to perform the estimation in these two steps. Instead, one can just directly construct moments for GMM estimation of the full demand system. In the next sections we summarize this estimation procedure. Formal derivations of the resulting limiting distributions are provided in the Appendix.

5 Implementing the Demand System

Here we show how the parameters in the system of demand equations (17) can be estimated, using GMM, based on the identification and estimation results derived in the Appendix. As in the generic model, we use pairwise differencing to prove identification and to construct the corresponding GMM estimator. We then show how additional moments to increase efficiency can be constructed by making additional random effects assumptions regarding $\mathbf{v}_{gt}$.

Based on the identification results provided in the Appendix, we can estimate all of the parameters of our demand system assuming we have data from at least J different price regimes (time periods and/or regions), which we will index by t. 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 $q_{Ji} = \left(x_i - \sum_{j=1}^{J-1} p_{jt}q_{ji}\right)/p_{Jt}$.

Most of our analyses will be based on $J = 2$, with the two goods being luxuries and necessities, and where $\mathbf{A}$ and $\mathbf{R}$ are both diagonal. With $J = 2$ we only need to estimate the demand equation for good 1, for which equations (17) and equation (18) (assuming $\mathbf{A}$ and $\mathbf{R}$ are diagonal) simplify to

$$\begin{aligned} q_{1i} = {} & X_i^2 e^{-b_1 \ln p_{1t} - (1-b_1)\ln p_{2t}} d_1/p_{1t} + X_i b_1/p_{1t} \\ & + R_{11} + A_{11}\overline{q}_{g1t} + \mathbf{C}_1'\mathbf{z}_i + v_{jgt} + u_{ji}, \end{aligned} \tag{19}$$

where deflated income X_i is given by

$$\begin{aligned} X_i = {} & X(\mathbf{p}_t, x, \overline{\mathbf{q}}_{gt}, \mathbf{z}_i) = x_i - R_{11}p_{1t} - R_{22}p_{2t} \\ & - \left(A_{11}\overline{q}_{g1t} + A_{22}\overline{q}_{g2t} + \mathbf{C}_1'\mathbf{z}_i\right) p_{1t} - \mathbf{C}_2'\mathbf{z}_i p_{2t}. \end{aligned} \tag{20}$$

5.1 Estimating the Demand System With Fixed Effects

As in the generic model, the main complication with estimation arises from the fact that the model is nonlinear, and the error from estimation of $\overline{\mathbf{q}}_g$ can generally be correlated with other variables and with the model error terms. As in the generic model, to construct valid moments with fixed effects we difference across all pairs of individuals i and i' in each group g, and construct both appropriate instruments and an appropriate estimator for $\overline{\mathbf{q}}_g$ that eliminates these unwanted correlations. The estimator for the unobservable true group mean $\overline{\mathbf{q}}_g$ is, as in the generic model, the leave-two-out estimated average

$$\widehat{\mathbf{q}}_{gt,-ii'} = \frac{1}{n_{gt}-2} \sum_{l \in g,t,l \neq i,i'} \mathbf{q}_l.$$

Given any pair of individuals i and i' in group g in time and district t, the variable $\widehat{\mathbf{q}}_{gt,-ii'}$ is simply the average of $\mathbf{q}_l$ over all of the observed consumers l in g and t, except for the individuals i and i'. This definition assumes that n_{gt}, the number of individuals observed in our data in each group g in each time and district t, is three or more.

For each pair of consumers i and i' in each group g in each price regime t, the moments we have for estimation are

$$0 = E\{\mathbf{r}_{gtii'}\left[q_{ji} - q_{ji'} - Q_j\left(\mathbf{p}_t, x_i, \widehat{\mathbf{q}}_{gt,-ii'}, \mathbf{z}_i\right) + Q_j\left(\mathbf{p}_t, x_{i'}, \widehat{\mathbf{q}}_{gt,-ii'}, \mathbf{z}_{i'}\right)\right]\} \tag{21}$$

where the vector of instruments $\mathbf{r}_{gtii'}$ is defined below. Equation (21) holds for goods $j = 1, ..., J-1$ 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 observed individuals i in group g in all time periods *except* the time period of price regime t. 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 $k = 1, 2, ..., K$. 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 $(z_{1i}z_{2i} - z_{1i'}z_{2i'})$.

For estimation of these moments by GMM, let the unit of observation be each observed pair of consumers i and i' in each group and price regime gt. The total number of moments is the number of elements of $\mathbf{r}_{gtii'}$ times $J-1$, because equation (21) applies to each good and each instrument. As with the generic model, one must use clustered standard errors, where each group is a cluster, to account for the correlations that, by construction, will exist among the i and i' pairs that comprise each observation within each group. Clustering over time as well as across individuals allows for possible serial correlation in the errors. See the Appendix for details regarding the construction and properties of this estimator.

5.2 Estimating the Demand Model With Random Effects

As in the generic model, a great deal of information is lost by differencing out the fixed effects. We now consider adding additional assumptions to the demand model, in particular that $\mathbf{v}_{gt}$ is independent of $x_i, \mathbf{p}_t, \mathbf{z}_i$ and $\overline{\mathbf{q}}_{gt}$, which allows us to treat $\mathbf{v}_{gt}$ as random effects. These assumptions provide stronger moments for GMM estimation that do not entail differencing.

In the fixed effects model, differencing removed $\mathbf{v}_{gt}$, but it also removed the matrix $\overline{\mathbf{q}}_{gt}\overline{\mathbf{q}}'_{gt}$ that appears inside the squared deflated income term X^2. In the random effects estimator we do not difference, so $\overline{\mathbf{q}}_{gt}\overline{\mathbf{q}}'_{gt}$ does not drop out, and we must deal with estimation error in this quadratic term. Our solution takes the same form as in the random effects generic model. We use $\widehat{\mathbf{q}}_{gt,-ii'}$ to estimate $\overline{\mathbf{q}}_{gt}$ wherever it appears linearly, and for every i' in the same gt as i, we use $\widehat{\mathbf{q}}_{gt,-ii'}\mathbf{q}'_{i'}$ to estimate the product $\overline{\mathbf{q}}_{gt}\overline{\mathbf{q}}'_{gt}$.

Specifically, based on equation (18), define $\widehat{X}_{1tg,-ii'}$ and $\widehat{X}_{2tg,-ii'}$ (estimates of X and X^2, respectively) by

$$\widehat{X}_{1tg,-ii'} = x_i - \mathbf{p}_t^{1/2\prime}\mathbf{R}\mathbf{p}_t^{1/2} - \mathbf{p}'_t\mathbf{A}\widehat{\mathbf{q}}_{gt,-ii'} - \mathbf{p}'_t\mathbf{C}\mathbf{z}_i. \tag{22}$$

and

$$\begin{aligned}\widehat{X}_{2tg,-ii'} &= \mathbf{p}'_t\mathbf{A}\widehat{\mathbf{q}}_{gt,-ii'}\mathbf{q}'_{i'}\mathbf{A}'\mathbf{p}_t - 2\left(x_i - \mathbf{p}_t^{1/2\prime}\mathbf{R}\mathbf{p}_t^{1/2} - \mathbf{p}'_t\mathbf{C}\mathbf{z}_i\right)\mathbf{p}'_t\mathbf{A}\widehat{\mathbf{q}}_{gt,-ii'} \\ &\quad + \left(x_i - \mathbf{p}_t^{1/2\prime}\mathbf{R}\mathbf{p}_t^{1/2} - \mathbf{p}'_t\mathbf{C}\mathbf{z}_i\right)^2.\end{aligned} \tag{23}$$

Equations (22) and (23) are nothing more than replacing $\overline{\mathbf{q}}_{gt}$ with $\widehat{\mathbf{q}}_{gt,-ii'}$ and replacing $\overline{\mathbf{q}}_{gt}\overline{\mathbf{q}}'_{gt}$ with $\widehat{\mathbf{q}}_{gt,-ii'}\mathbf{q}'_{i'}$ in the expressions for X and X^2.

Let $\mathbf{r}_{gti}$ denote the vector of instruments consisting of the elements $\widehat{x}_{(t)g}$, $\widehat{\mathbf{z}}_{(t)g}$, $\mathbf{p}_t$, x_i, $\mathbf{z}_i$, along with squares and cross products of these elements. The instruments here include functions of x_i and $\mathbf{z}_i$ but, unlike the fixed effects case, the instruments do not include functions of $x_{i'}$ and $\mathbf{z}_{i'}$ and so also do not contains functions of differences like $x_i - x_{i'}$ or $\mathbf{z}_i - \mathbf{z}_{i'}$. The resulting moments used to estimate the random effects model are, for goods $j = 1, ..., J-1$,

$$0 = E\left[\mathbf{r}_{gti}\left(q_{ji} - \widehat{X}_{2tg,-ii'}e^{-\mathbf{b}'\ln\mathbf{p}_t}\frac{d_j}{p_{jt}} - \widehat{X}_{1tg,-ii'}\frac{b_j}{p_{jt}} - R_{jj} - \sum_{k\neq j}R_{jk}\sqrt{\frac{p_{kt}}{p_{jt}}} - \mathbf{A}'_j\widehat{\mathbf{q}}_{gt,-ii'} - \mathbf{C}'_j\mathbf{z}_i - v_{jt0}\right)\right] \tag{24}$$

where v_{jt0} is an additional constant term for each good j and price regime t that results from taking an average of the random effects component of the error term. As in the generic random effects model, this nonzero v_{jt0} term is due to the error in estimating the product $\overline{\mathbf{q}}_{gt}\overline{\mathbf{q}}'_{gt}$.

We show in the Appendix that, under suitable assumptions listed there, equation (24) holds after substituting in equations (22) and (23), and these suffice for identifying all of the model parameters $\mathbf{A}$, $\mathbf{C}$, $\mathbf{R}$, $\mathbf{d}$, and $\mathbf{b}$. The associated estimator can be summarized as follows:

1. Start with the demand model $Q_j\left(\mathbf{p}_t, x_i, \overline{\mathbf{q}}_{gt}, \mathbf{z}_i\right)$ for each good j and consumer i, defined by plugging equation (18) into equation (17). Note that this demand model has parameters $\mathbf{A}$, $\mathbf{C}$, $\mathbf{R}$, $\mathbf{d}$, and $\mathbf{b}$.

2. In these demand equations replace the product $\overline{\mathbf{q}}_{gt}\overline{\mathbf{q}}'_{gt}$ with $\widehat{\mathbf{q}}_{gt,-ii'}\mathbf{q}'_{i'}$, and everywhere else $\overline{\mathbf{q}}_{gt}$ appears replace $\overline{\mathbf{q}}_{gt}$ with $\widehat{\mathbf{q}}_{gt,-ii'}$ where i' is another individual in the same group g and the same price regime t as individual i. Call the resulting demand model $\widetilde{Q}_j\left(\mathbf{p}_t, x_i, \widehat{\mathbf{q}}_{gt,-ii'}, \mathbf{q}_{i'}, \mathbf{z}_i\right)$.

3. Estimate the model parameters by GMM, for $j = 1, ..., J-1$, using as moments $E\left[\left(q_{ji} - \widetilde{Q}_j\left(\mathbf{p}_t, x_i, \widehat{\mathbf{q}}_{gt,-ii'}, \mathbf{q}_{i'}, \mathbf{z}_i\right) - v_{jt0}\right)\mathbf{r}_{gti}\right] = 0$. The data consist of every pair of individuals i and i' in each group and price regime gt. Inference is standard GMM with clustered standard errors, where each group g is a cluster. If we add the additional assumption that the random effects $\mathbf{v}_{gt}$ are uncorrelated across price regimes t, then we can take each g in each price regime t as a separate cluster.

There are two options for dealing with the v_{jt0} terms. One is just treat the set of constants v_{jt0} for $j = 1, ..., J-1$ and $t = 1, ..., T$ as $(J-1)T$ additional parameters to estimate. However, this option may be unattractive in terms of loss of efficiency if T is relatively large. Alternatively, in the Appendix we show that, given some additional mild assumptions,

$$v_{jt0} = -\mathbf{p}'_t\mathbf{A}\mathbf{\Sigma}_v\mathbf{A}'\mathbf{p}_t\left(e^{-\mathbf{b}'\ln\mathbf{p}_t}\right)d_j/p_{jt}$$

where $\mathbf{\Sigma}_v = Var(\mathbf{v}_{gt})$, which for this construction is assumed constant over time. Then, instead of estimating $(J-1)T$ parameters in addition to the structural ones, we may substitute this expression for v_{jt0} into the above GMM moments, and thereby only estimate the matrix $\mathbf{\Sigma}_v$, consisting of $J(J+1)/2$ parameters, in addition to the structural parameters.

Formal details, derivations, and limiting distribution theory of this estimator are provided in the Appendix.

6 Structural Demand System: Empirical Results

In this section, we estimate the structural demand model defined by plugging equation (18) into equation (17) for each good j, and adding error terms u and group level fixed or random effects v to each equation in each time period. We begin with a $J = 2$ goods model, which therefore only requires estimating the demand for good $j = 1$, defined by substituting equation (20) into equation (19). Demand for good $j = 2$ is then given by the adding up constraint defined earlier.

Like most standard modern continuous demand models (see, e.g., Banks, Blundell, and Lewbel 1997 or Lewbel and Pendakur, 2009), our theoretical model includes a quadratic function of prices given by the matrix $\mathbf{R}$, to allow for general cross price effects. However, in our data the geometric mean of prices turns out to be highly collinear with some individual prices, leading to a severe multicollinearity problem when $\mathbf{R}$ is not diagonal. We therefore restrict $\mathbf{R}$ to be diagonal. Note that, because of the presence of additional price functions in our model, imposing the constraint that $\mathbf{R}$ be diagonal is not restrictive when $J \leq 3$, in the sense that our model remains Diewert-flexible (see, e.g., Diewert 1974) in own and cross price effects even with this restriction.

Our fixed and random effects estimators are as described in the previous section. With $J = 2$, our two goods are luxuries and necessities, and so we are estimating just one equation for luxury demands, which is similar in spirit to the generic model estimates we reported earlier. However, even with $J = 2$ our structural model estimates differ from the generic model in three important ways. First, our structural model provides a welfare measure of the peer effect spillovers through the needs component of the utility function. The matrix $\mathbf{A}$ gives the magnitude of these spillovers. If these are positive, then an increase in group-average spending increases needs, and affects utility in the same way as a decrease in household expenditure. This is indeed what we find, and is also consistent with Luttmer (2005) and with our WVS data life satisfaction estimates reported earlier. Second our structural model incorporates demographic and price variation, in a way that is consistent with utility maximization. Note that this required a nonlinear model, in that just including functions of prices and demographics as a regressors in a linear quantity demand function is inconsistent with utility maximization (see, e.g., Blundell, Duncan and Pendakur 1998 and Pendakur 2005). Third, we allow the demands for goods to depend separately on the group average spending of each good, e.g., luxury demands can depend on both group-average luxuries spending and group-average necessities spending.

6.1 Structural Model Data

In our baseline empirical work, we consider only non-urban Hindu non-Scheduled Caste/Scheduled Tribe (SC/ST) households. We report some additional results for samples of non-urban non-Hindu households and samples of SC/ST households. We have groups defined by education (3 categories: illiterate or barely literate; primary or some secondary; completed secondary or more) and by district (575 districts across 33 states), and 4 years of data, allowing each group to be observed up to four times. We require each group to have at least 10 observations in each of at least two time periods. Roughly 18 per cent of households are lost with this restriction. We experimented with including very small groups but this resulted in a substantial decrease in estimation precision. Essentially, very small groups have very noisy estimates of group-averages, and although our measurement error correction is consistent even when including these groups, its efficiency properties are adversely affected.

For our sample of non-urban Hindu non-SC/ST households, these choices result in a total of 1111 distinct groups that are observed in at least 2 time periods each, for a total of 2354 period-groups. Each group is seen in either two, three or four time periods, but most groups are observed only twice. Average size of a group in each period is about two dozen households.

Our observed prices vary by time and by state, so since t indexes price regimes, t ranges from 1 to $T = 4 * 33 = 132$. Each individual i is only observed once, in one price regime and belonging to one group.

Table 5 gives summary statistics for the NSS data used in this part. These are analogous to those shown in Table 1, but here we are including additional regressors and multiple rounds (time periods) of data. Table 5 shows summary statistics for the nine household-level demographic characteristics that comprise our vector z_i. These are household size less 1 divided by 10; the age of the head of the household divided by 120; an indicator that there is a married couple in the household; the natural log of one plus the number of hectares of land owned by the household; an indicator that the household has a ration card for basic foods and fuels; an indicator that the household head is Hindu and Dalit or scheduled (i.e lowest) caste (equal to zero for non-Dalit/SC Hindus); an indicator that the household is Muslim or other religion; and indicators that the highest level of education of the household head is primary or secondary level (they are both zero for uneducated or illiterate household heads).

Table 5 also gives summary statistics on state-level prices $\mathbf{p}$. We construct prices of our demand aggregates as follows. In a first stage, following Deaton (1998), we compute state-item-level local average unit-value prices for the subset of items for which we have quantity data, to equal the state-level sum of spending divided by the state-level sum of quantities. Then, in a second stage, we aggregate these state-item-level unit value prices into state-level luxury and necessity prices using a Stone price index, with weights given by the overall sample average spending on each item. In a typical state and period, these prices are computed as averages of roughly 2000 observations, so we do not attempt to instrument for possible measurement errors in these constructed price regressors.

Table 5 also reports summary statistics for prices and quantities of visible and invisible subcomponents of luxuries and necessities. We use these later on, when we consider the question of whether social interaction effects differ for goods that are visible to other consumers in comparison to those that are not visible. We use the categorization of Roth (2014) to classify goods as visible versus invisible.

6.2 Baseline Structural Model Estimates

Table 6 gives estimates of the spillover parameter matrix $\mathbf{A}$ using fixed effects (FE) moments for the 2-good system (luxuries and necessities). We consider 2 cases here: the left panel, labeled "A same," gives estimates for the case where $\mathbf{A}$ is equal to a scalar, a, times the identity matrix, so $\mathbf{A} = a\mathbf{I}_J$. The right panel of Table 6, labeled "A diagonal" gives estimates for the case where $\mathbf{A}$ is a diagonal matrix. Later we will consider cross-effects, allowing $\mathbf{A}$ to have non-zero off diagonal elements.

In the "A same" case, needs are given by $F_i(\mathbf{p}_t) = \mathbf{p}_t'\mathbf{A}\overline{\mathbf{q}}_{gt} + \mathbf{p}_t'\mathbf{C}\mathbf{z}_i = a\mathbf{p}_t'\overline{\mathbf{q}}_{gt} + \mathbf{p}_t'\mathbf{C}\mathbf{z}_i = a\overline{x}_{gt} + \mathbf{p}_t'\mathbf{C}\mathbf{z}_i$, so we can interpret the scalar a as $\partial F/\partial \overline{x}_{gt} = a$. The estimate of the scalar a in Table 6 is 0.50, meaning that a 100 rupee increase in group-average income $\overline{x}_{gt}$ increases perceived needs (and therefore decreases equivalent income) by 50 rupees. The standard error of a is 0.11 so we can reject $a = 0$, which would correspond to no peer effects. We can also reject $a = 1$ which would correspond to peer effects so large that there are no increases in utility associated with aggregate consumption growth.

This scalar a, obtained using revealed preference theory on consumption data, has a roughly comparable interpretation to the estimate of spillover effects in the WVS life satisfaction model reported in section 3.3. Our estimate there was 0.68. This is also comparable to Luttmer's (2005) estimate of 0.76 using stated well-being data.

In the bottom left of Table 6, we test, and reject, the hypothesis that the two elements on the diagonal of $\mathbf{A}$ are equal to each other. However, when we estimated the model allowing the two elements to differ (see the right panel of Table 6), we obtain estimates that lie far outside the plausible range of $[0, 1]$. These estimates are also very imprecise, with standard errors are roughly triple those in the left panel. The explanation for this imprecision and corresponding "wildness" of the estimates is that, in the FE model, all the parameters in $\mathbf{A}$ are identified just from the $(x_i - x_{i'})\overline{\mathbf{q}}_{gt}$ interaction terms (recall here and below that, in the actual estimates, $\overline{\mathbf{q}}_{gt}$ is unobserved and is replaced by an estimate like $\widehat{\mathbf{q}}_{gt,-ii'}$). In our data, the elements of our estimate of $\overline{\mathbf{q}}_{gt}$ are highly correlated with each other, with a correlation coefficient of 0.85, resulting in a large degree of multicollinearity. The result is that the estimated first element of $\mathbf{A}$ is implausibly low, offset by the second element of $\mathbf{A}$

that is implausibly high by a similar magnitude.

This problem of multicollinarity is considerably reduced in the Random Effects (RE) model, with its stronger assumptions. In particular the RE model contains an additive $\overline{\mathbf{q}}_{gt}$ term which is differenced out in the FE model. This is in addition to a now undifferenced $x_i\overline{\mathbf{q}}_{gt}$ interaction term.

This additional $\overline{\mathbf{q}}_{gt}$ helps to pin down the estimate of $\mathbf{A}$. On the bottom of Table 6, we report the results of a Hausman test comparing the FE and RE models. The additional restrictions of the RE model are not rejected in the "A same" specification, but are rejected in the more general "A diagonal" specification.

The estimates of $\mathbf{A}$ in the RE model are reported in Table 7. The RE estimate of the scalar a in the "A same" model is 0.55, while for diagonal $\mathbf{A}$ the estimate of the luxuries spillover coefficient (the first element on the diagonal of $\mathbf{A}$) is 0.46 and the necessities spillover coefficient is 0.57. The standard errors of these estimates are around 0.02, far lower than in the fixed effects model. While similar in magnitude, we reject the hypothesis that the two elements of $\mathbf{A}$ are equal, which is not surprising given how small their standard errors are.

The interpretation of these separate coefficients is that the j'th element on the diagonal of $\mathbf{A}$ equals $\partial F/\partial\left(p_{jt}\overline{q}_{jgt}\right)$, which is the response of perceived needs to a 1 rupee increase in average group expenditures on good j, $p_{jt}\overline{q}_{jgt}$. To compare these estimates to the scalar a, suppose group average expenditure $\overline{x}_{gt}$ increased by 100 rupees. Then group average luxury expenditures, $p_{jt}\overline{q}_{jgt}$ would increase by about 30 rupees (since, in Table 5, luxuries are about 30% of total spending), and so the luxuries spillover would be about 14 rupees (0.46 times 30). Similarly the necessities spillover is about 40 rupees (0.57 times 70), yielding total spillovers of 54 rupees, which is very close to the estimates one gets with a scalar a (50 rupees in the FE model or 55 rupees in the RE model).

6.3 Alternative Structural Model Estimates

In the rightmost panel of Table 7, we report RE estimates where the matrix $\mathbf{A}$ is unrestricted, allowing for nonzero cross-effects, e.g., allowing peer group consumption of necessities to directly impact one's demand for luxuries. The estimates display a similar (though less extreme) wildness to that of the FE model with diagonal $\mathbf{A}$. The reason is similar, in that now we are trying to estimate four coefficients primarily from the four multicollinear terms $x_i\overline{q}_{1gt}$, $x_i\overline{q}_{2gt}$, $\overline{q}_{1gt}$, and $\overline{q}_{2gt}$. So although we formally prove identification of the model with a general $\mathbf{A}$, one would either require a larger data set or more relative variation in group quantities and within group total expenditures to obtain reliable estimates .

In Table 8, we turn to the question of whether or not consumption externalities depend on whether or not goods are visible or invisible to one's peers, according to the characterisation of Roth (2014). The idea is peer effects may be larger for visible goods, both because they are more conspicuous, and because of potential Veblen (1899) effects. Here, we would expect larger consumption externalities for visible goods than for invisible goods. We might additionally expect this to be particularly true for luxuries, as opposed to necessities. We now have a demand system with $J = 4$ goods.

The first columns of Table 8 give the fixed- and random-effects estimates of the scalar a in the "A same" model, where now four elements of the diagonal of $\mathbf{A}$ are all constrained to be equal. The estimates of the scalar a are 0.71 and 0.65, respectively. These are rather higher than the 0.50 to 0.55 estimates we obtained with $J = 2$ goods, but are closer to our WVS life satisfaction data model estimate of 0.68, and Luttmer's (2005) estimate of 0.76 using stated well-being data. The estimates of the scalar a with $J = 4$ goods have smaller standard errors than in the case with $J = 2$ goods, because now there are more equations being used to estimate the same parameter.

The last columns of Table 8 give the RE estimates of the "diagonal A" model. As before, we find that luxuries have somewhat smaller externalities than necessities. However, estimated element of $\mathbf{A}$ for visible luxuries is smaller than that for invisible luxuries, though the estimated value for visible necessities is larger than for invisible necessities. So the Veblen or conspicuous consumption story is supported for necessities, but not for luxuries.

In Table 9, we consider how consumption externalities vary across group-level characteristics. In the left-hand panel, we provide fixed effects estimates of the scalar a in the "A same," $J = 2$ goods model on 3 subsamples of the nonurban population: Hindu non-SC/ST (non-Scheduled Caste/Scheduled Tribe) households, SC/ST households, and non-Hindu SC/ST households. We find that these groups differ substantially not just in the estimates of the scalar a, but also in their b_j and d_j coefficients. For Hindu non-SC/ST households, the estimate of a is 0.50 (the same as in our baseline model) but for SC/ST households and for non-Hindu non-SC/ST households, the point estimates of a are close to zero, though with larger standard errors. This suggests that peer effects may vary by caste and religion.

This right hand panel of Table 9 reports differences in the scalar a across three education groups: illiterate/barely literate, primary or some secondary education, and complete secondary or more education. We initially ran this model on three different subsamples based on these education levels, but unlike the case with caste and religion, we found that the b_j and d_j coefficient estimates did not differ much across the groups. We therefore pooled the data, just letting the scalar a be a linear index in the three education levels. Here we find very low and insignificant spillovers for the illiterate/barely literate. In contrast, the estimate of a is 0.56 for the middle education group, and lower (but not significantly different from 0.56) in the highest education group.

6.4 Structural Model Estimates Summary

We draw the following lessons on spillovers from our structural revealed preference based model estimates.

First, while spillovers do appear to differ somewhat between luxuries and necessities, the estimates suggest that the matrix $\mathbf{A}$ of spillovers can be reasonably approximated by a scalar a times the identity matrix (the "A same" model), implying that consumption externalities are close to that scalar a times group-average expenditure. In our data, multicollinearity prevents estimation of completely general $\mathbf{A}$ matrix models.

Second, fixed effects results in a considerable loss of efficiency relative to random effects, and in the "A same" model, the added restrictions implied by random effects over fixed effects are not rejected.

Third, our baseline estimates of spillovers are at or a little above 0.5. However, alternative model specifications, and nonstructural estimates based on reported life satisfaction suggest higher spillovers of up to around 0.7. We also find evidence that particular subgroups may have lower spillovers.

7 Conclusions

We show identification and GMM estimation of a peer effects in a generic quadratic model, where in the data most members of each group might not be observed. The model allows for fixed or random effects, and allows the number of observed individuals in each peer group to be fixed, so we obtain consistent estimates of the model even though peer group means cannot be consistently estimated. Unlike most peer effects models, our model can be estimated from standard survey data where the vast majority of members of each peer group are not observed.

We next provide a utility derived consumer demand model, where one's perceived needs for each commodity depends in part on the average consumption of one's peers. We show how this model can be used for welfare analysis, and in particular to identify what fraction of income increases are spent on "keeping up with the Joneses" type peer effects. This demand model that incorporates needs has a structure analogous to our generic peer effects model, and so can be identified and estimated in the same way

We apply the model to consumption data from India, and find large peer effects. Our estimates imply that an increase in group-average spending of 100 rupees would induce an increase in needs of 50 rupees or more. In this model, an increase in needs is, from the individual consumer's point of view, equivalent to a decrease in total expenditures. These results could therefore at least partly explain the Easterlin (1974) paradox, in that income growth over time, which increases people's consumption budgets, likely results in much less utility growth than standard demand model estimates (that ignore these peer effects) would suggest.

These results also suggest that income or consumption taxes may have far less negative effects on consumer welfare than are implied by standard models. This is because a tax that reduces my expenditures by a dollar will, if applied to everyone in my peer group, have the same effect on my utility as a tax of only 50 cents that ignores the peer effects.In short, the larger these peer effects are, the smaller are the welfare gains associated with tax cuts or mean income growth.

8 References

Akay, A., and P. Martinsson, (2011), "Does relative income matter for the very poor? Evidence from rural Ethiopia," Economics Letters 110(3), 213-215.

Banks, J., R. Blundell, and A. Lewbel, (1997), "Quadratic Engel curves and consumer demand," Review of Economics and Statistics 79, 527–539.

Blackorby, C. and D. Donaldson, (1994), "Measuring the Costs of Children: A Theoretical Framework," in R. Blundell, I. Preston, and I. Walker, eds., The Economics of Household Behaviour (Cambridge University Press) pp. 51-69.

Blume, L. E., W. A. Brock, S. N. Durlauf, and Y. M. Ioannides., (2010) "Identification of Social Interactions," Economics Series 260, Institute for Advanced Studies.

Boneva, T. (2013), "Neighbourhood Effects in Consumption: Evidence from Disaggregated Consumption Data," Cambridge Working Papers in Economics 1328.

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

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

Carrell, S.E., R.L. Fullerton, and J.E. West (2009), "Does your cohort matter? Measuring peer effects in college achievement." Journal of Labor Economics, 27(3), 439-464.

Chao, A. and J. B. Schor, (1998) "Empirical tests of status consumption: Evidence from women's cosmetics," Journal of Economic Psychology 19 (1), 107-131.

Clark, A. E., and C. Senik, (2010), ”Who compares to whom? the anatomy of income comparisons in europe,” The Economic Journal, 120(544), 573-594.

Clark, A. E., P. Frijters, and M. A. Shields, (2008), "Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles," Journal of Economic Literature, 46(1), 95-144.

Diewert, W.E., (1974), "Applications of duality theory," in: M.D. Intriligator and D.A. Kendrick, eds., Frontiers of quantitative economics, vol. II (North-Holland, Amsterdam), 106-171.

Donaldson, D. and K. Pendakur, (2006), ”The Identification of Fixed Costs from Behaviour” Journal of Business and Economics Statistics, 24, 255-265.

Duflo, E., P. Dupas, and M. Kremer, (2011), “Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya,"American Economic Review, 101(5): 1739-74.

Dupor, B., and W-F Liu, (2003), ”Jealousy and equilibrium overconsumption,” The American Economic Review 93(1), 423-428.

Fontaine, X., and K. Yamada, (2012), "Economic Comparison and Group Identity: Lessons from India," Working Paper hal-00711212.

Frank, R. H., (1999), Luxury Fever: Why Money Fails to Satisfy In An Era of Excess, The Free Press, Simon and Schuster: New York.

Frank, R. H., (2012), ”The Easterlin Paradox Revisited” Emotion 12(6): 1188-1191.

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," in M. J. Artis and A. R. Nobay, eds., Essays in Economic Analysis: Proceedings of the 1975 AUTE Confetence, Sheffield (Cambridge: Cambridge University Press), 211-243.

Hansen, L. P., (1982), "Large Sample Properties of Generalized Method of Moments Estimators," Econometrica, 50(4), 1029–1054.

Hu, Y. and 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., and. R. S. Winer, (1995), "Empirical generalizations from reference price research," Marketing Science, 14(3 supplement), G161-G169.

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

Lee, L.-F., (2007), "Identification and estimation of econometric models with group interactions, contextual factors and fixed effects," Journal of Econometrics 140(2), 333-374.

Lewbel, A. and K. Pendakur, (2009), "Tricks with Hicks: The EASI Demand System," American Economic Review, 99(3): 827-63.

Luttmer, E. F. P., (2005), "Neighbors as Negatives: Relative Earnings and Well-Being," Quarterly Journal of Economics, 120(3), 963-1002.

Maurer, J., and A. 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, C. F., (2000), "Economic analysis of social interactions," Journal of Economic Perspectives, 14, 115-136.

Pendakur, K, (1999), "Semiparametric Estimation and Testing of Base-Independent Equivalence Scales," Journal of Econometrics, 88(1), 1-40.

Pendakur, K., (2005), "Semiparametric Estimation of Lifetime Equivalence Scales," Journal of Appied Econometrics, 20(4), 487–507.

Rabin, M., (1998), "Psychology and economics," Journal of Economic Literature, 36(1), 11-46.

Ravallion, M., and M. Lokshin, (2010), "Who cares about relative deprivation?" Journal of Economic Behavior and Organization, 73(2), 171-185.

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

Roy, R., (1947). "La Distribution du Revenu Entre Les Divers Biens," Econometrica, 15(3), 205–225.

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., and M. Sato, (2013), "Another Avenue for Anatomy of Income Comparisons: Evidence from Hypothetical Choice Experiments," Journal of Economic Behavior and Organization, 89, 35–57.

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 $\overline{y}_g = E\left(y_i \mid i \in g\right)$, $\widehat{y}_{g,-ii'} = \sum_{l \in g, l \neq i,i'} y_l/(n_g - 2)$, 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 $\overline{\mathbf{x}}_g = E\left(\mathbf{x}_i \mid i \in g\right)$, $\overline{\mathbf{x}\mathbf{x}'_g} = E\left(\mathbf{x}_i\mathbf{x}'_i \mid i \in g\right)$, 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

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

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 (25) 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 (25).

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 (25). $\mathbf{x}_i$ is a K-dimensional vector of covariates. For each $k \in \{1, ..., K\}$, for each group g with $i \in g$ and $i' \in g$, $\Pr(\mathbf{x}_{ik} = \mathbf{x}_{i'k}) > 0$. 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 (25) 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

$$\overline{y}_g = \frac{1 - a(2\mathbf{b}'\overline{\mathbf{x}}_g d + 1) \pm \sqrt{[1 - a(2\mathbf{b}'\overline{\mathbf{x}}_g d + 1)]^2 - 4a^2 d[d\mathbf{b}'\overline{\mathbf{x}\mathbf{x}'_g}\mathbf{b} + \mathbf{b}'\overline{\mathbf{x}}_g + v_g]}}{2a^2 d} \tag{26}$$

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 (26) 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 (25), which we do by subtracting off another individual in the same group. For each $(i, i') \in g$, consider pairwise difference

$$\begin{aligned} y_i - y_{i'} &= 2ad\overline{y}_g \mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'}) + d\mathbf{b}'(\mathbf{x}_i\mathbf{x}_i' - \mathbf{x}_{i'}\mathbf{x}_{i'}')\mathbf{b} + \mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'}) + u_i - u_{i'} \\ &= 2ad\widehat{y}_{g,-ii'}\mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'}) + d\mathbf{b}'(\mathbf{x}_i\mathbf{x}_i' - \mathbf{x}_{i'}\mathbf{x}_{i'}')\mathbf{b} + \mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'}) \\ &+ u_i - u_{i'} - 2ad\varepsilon_{yg,-ii'}\mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'}), \end{aligned} \tag{27}$$

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 (27) 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

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

where

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

Substituting this expression into equation (27) 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 $\varepsilon_{\mathbf{x}g,-ii'}$ and $\varepsilon_{\mathbf{xx}g,-ii'}$, 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 $E\left((\mathbf{x}_i - \overline{\mathbf{x}}_g) \mid i \in g, \overline{\mathbf{x}}_g, \overline{\mathbf{x}\mathbf{x}'}_g, v_g, \mathbf{r}_g\right) = 0$, $E\left((\mathbf{x}_i\mathbf{x}_i' - \overline{\mathbf{x}\mathbf{x}'}_g) \mid i \in g, \mathbf{r}_g\right) = 0$, and that $\mathbf{x}_i - \overline{\mathbf{x}}_g$ and $\mathbf{x}_i\mathbf{x}_i' - \overline{\mathbf{x}\mathbf{x}'}_g$ 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 $\varepsilon_{ix} = \mathbf{x}_i - \overline{\mathbf{x}}_g$ be independent across individuals, and assume that $E(\varepsilon_{ix} \mid \overline{\mathbf{x}}_g, \overline{\mathbf{x}\mathbf{x}'}_g, v_g, \mathbf{r}_g \text{ for } i \in g) = 0$ and $E\left(\varepsilon_{ix}\varepsilon_{ix}' \mid \overline{\mathbf{x}}_g, \mathbf{r}_g \text{ for } i \in g\right) = E\left(\varepsilon_{ix}\varepsilon_{ix}' \mid i \in g\right)$. To see this, we have

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

A simpler but stronger sufficient condition would just be that ε_{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 $E\left(\varepsilon_{yg,-ii'} \mid \mathbf{r}_g\right) = 0$, and the independence of measurement errors across individuals is used to show $E\left(\varepsilon_{yg,-ii'}(\mathbf{x}_i - \mathbf{x}_{i'}) \mid \mathbf{r}_g, \mathbf{x}_i, \mathbf{x}_{i'}\right) = (\mathbf{x}_i - \mathbf{x}_{i'})E\left(\varepsilon_{yg,-ii'} \mid \mathbf{r}_g\right) = 0$, so that $\mathbf{x}_i$ and $\mathbf{x}_{i'}$ are valid instruments. Given Assumptions A1 and A4, one can directly verify that

$$E\left[y_i - y_{i'} - (2ad\widehat{y}_{g,-ii'}\mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'}) + d\mathbf{b}'(\mathbf{x}_i\mathbf{x}_i' - \mathbf{x}_{i'}\mathbf{x}_{i'}')\mathbf{b} + \mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'})) \mid \mathbf{r}_g, \mathbf{x}_i, \mathbf{x}_{i'}\right] = 0. \tag{28}$$

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 (26) 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 $\mathbf{x}_i = \varepsilon_{ix} + \overline{\mathbf{x}}_g$.

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 (28) that

$$E\left[\left(y_i - y_{i'} - (1+2ad\widehat{y}_{g,-ii'})\textstyle\sum_{k=1}^{K} b_k(x_{ki}-x_{ki'}) - d\sum_{k=1}^{K}\sum_{k'=1}^{K} b_k b_{k'}(x_{ki}x_{k'i} - x_{ki'}x_{k'i'})\right)\mathbf{r}_{gii'}\right] = 0. \tag{29}$$

Let

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

Equation (29) 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 (29) over all $(i,i') \in \Gamma_g$ gives the unconditional group level moment vector

$$E\left(\mathbf{Y}_{1g} - \textstyle\sum_{k=1}^{K} b_k\mathbf{Y}_{2kg} - 2ad\sum_{k=1}^{K} b_k\mathbf{Y}_{3kg} - d\sum_{k=1}^{K}\sum_{k'=1}^{K} b_k b_{k'}\mathbf{Y}_{4kk'g}\right) = 0. \tag{30}$$

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 (30).

Given the identification in Theorem 1, based on equation (30) we can immediately construct a corresponding group level GMM estimator

$$\left(\widehat{a},\widehat{b}_1,...\widehat{b}_K,\widehat{d}\right)=\arg\min\left[\frac{1}{G}\sum_{g=1}^{G}\left(\mathbf{Y}_{1g}-\textstyle\sum_{k=1}^{K}b_k\mathbf{Y}_{2kg}-2ad\sum_{k=1}^{K}b_k\mathbf{Y}_{3kg}-d\sum_{k=1}^{K}\sum_{k'=1}^{K}b_kb_{k'}\mathbf{Y}_{4kk'g}\right)\right]'$$
$$\cdot\widehat{\Omega}\left[\frac{1}{G}\sum_{g=1}^{G}\left(\mathbf{Y}_{1g}-\textstyle\sum_{k=1}^{K}b_k\mathbf{Y}_{2kg}-2ad\sum_{k=1}^{K}b_k\mathbf{Y}_{3kg}-d\sum_{k=1}^{K}\sum_{k'=1}^{K}b_kb_{k'}\mathbf{Y}_{4kk'g}\right)\right] \tag{31}$$

for some positive definite moment weighting matrix $\widehat{\Omega}$. In equation (31), 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 (29), 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,

$$\left(\widehat{a},\widehat{b}_1,...\widehat{b}_K,\widehat{d}\right)=\arg\min\left(\frac{\sum_{g=1}^{G}\sum_{(i,i')\in\Gamma_g}\mathbf{m}_{gii'}}{\sum_{g=1}^{G}\sum_{(i,i')\in\Gamma_g}1}\right)'\widehat{\Omega}\left(\frac{\sum_{g=1}^{G}\sum_{(i,i')\in\Gamma_g}\mathbf{m}_{gii'}}{\sum_{g=1}^{G}\sum_{(i,i')\in\Gamma_g}1}\right), \tag{32}$$

where

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

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 (29), $\mathbf{r}_{gii'}$ 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 (26) 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 (26) 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 (25), 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 (25), we have

$$y_i = \overline{y}_g^2 da^2 + a(2d\mathbf{x}_i'\mathbf{b} + 1)\overline{y}_g + \mathbf{b}'\mathbf{x}_i\mathbf{x}_i'\mathbf{b}d + \mathbf{x}_i'\mathbf{b} + v_g + u_i \tag{33}$$

Taking the within group expected value of this expression gives

$$\overline{y}_g = \overline{y}_g^2 da^2 + a(2d\mathbf{b}'\overline{\mathbf{x}}_g + 1)\overline{y}_g + d\mathbf{b}'\overline{\mathbf{x}\mathbf{x}'_g}\mathbf{b} + \mathbf{b}'\overline{\mathbf{x}}_g + v_g. \tag{34}$$

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 (34) 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 (33), 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 (34) then gives the measurement error

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

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

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

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

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

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

$$
E\left(\mathbf{Y}_{1g} - \textstyle\sum_{k=1}^{K} b_k \mathbf{Y}_{2kg} - 2ad\sum_{k=1}^{K} b_k \mathbf{Y}_{3kg} - d\sum_{k=1}^{K}\sum_{k'=1}^{K} b_k b_{k'} \mathbf{Y}_{4kk'g}\right) = 0. \qquad (35)
$$

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,

$$
y_{ji} = (\overline{\mathbf{y}}_g'\mathbf{a}_j + \mathbf{x}_i'\mathbf{b}_j)^2 d_j + \left(\overline{\mathbf{y}}_g'\mathbf{a}_j + \mathbf{x}_i'\mathbf{b}_j\right) + v_{jg} + u_{ji}, \qquad (36)
$$

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

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

Using the same analysis as before,

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

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

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

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

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

$$- \begin{pmatrix} d_1 \sum_{k=1}^{K}\sum_{k'=1}^{K} b_{1k} b_{1k'} L_{4kk'gii'}\mathbf{r}_{gii'} \\ \vdots \\ d_J \sum_{k=1}^{K}\sum_{k'=1}^{K} b_{Jk} b_{Jk'} L_{4kk'gii'}\mathbf{r}_{gii'} \end{pmatrix}.$$

8.4 Multiple Equation Generic Model With Random Effects

Here we provide the derivation of equation (12), 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 (25).

$$y_i = \overline{y}_g^2 a^2 d + a\left(1 + 2\mathbf{b}'\mathbf{x}_i d\right)\overline{y}_g + \mathbf{b}'\mathbf{x}_i + \mathbf{b}'\mathbf{x}_i\mathbf{x}_i'\mathbf{b}d + v_g + u_i, \tag{37}$$

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

$$\overline{y}_g = \overline{y}_g^2 da^2 + a(2d\mathbf{b}'\overline{\mathbf{x}}_g + 1)\overline{y}_g + d\mathbf{b}'\overline{\mathbf{x}\mathbf{x}'_g}\mathbf{b} + \mathbf{b}'\overline{\mathbf{x}}_g + \mu, \tag{38}$$

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

$$y_i = \widehat{y}_{g,-ii'} y_{i'} a^2 d + a\left(1 + 2\mathbf{b}'\mathbf{x}_i d\right)\widehat{y}_{g,-ii'} + \mathbf{b}'\mathbf{x}_i + \mathbf{b}'\mathbf{x}_i\mathbf{x}_i'\mathbf{b}d + v_g + u_i + \widetilde{\varepsilon}_{gii'}. \tag{39}$$

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

$$E(v_g + u_i + \widetilde{\varepsilon}_{gii'} \mid \overline{\mathbf{x}}_g, \overline{\mathbf{x}\mathbf{x}'_g}, \mathbf{x}_i) = \mu - da^2\sigma^2 = v_0. \tag{40}$$

By construction $v_g + u_i + \widetilde{\varepsilon}_{gii'}$ is also independent of $\mathbf{r}_g$. Given this, equation (12) then follows from equations (39) (40).

8.5 Identification of the Demand System With Fixed Effects

Here we outline how the parameters of the demand system are identified. This is followed by the formal proof of identification, based on the corresponding moments we construct for estimation. As with the generic model, equation (17) entails the complications associated with nonlinearity, and the issues that the fixed effects $\mathbf{v}_g$ correlate with regressors, and that $\overline{\mathbf{q}}_g$ is not observed. As before, 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.

We show identification of the parameters of the demand system (17) in two steps. The first step identifies some of the model parameters by closely following the identification strategy of our simpler generic model, holding prices fixed. The second step then identifies the remaining parameters based on varying prices. We summarize these steps here, then provide formal assumptions and proof of the identification in the next section.

For the first step, consider data just from a single time period and region, so there is no price variation and $\mathbf{p}$ can be treated as a vector of constants. Let $\alpha = \mathbf{A}'\mathbf{p}$, $\beta = \mathbf{p}^{1/2\prime}\mathbf{R}\mathbf{p}^{1/2}$, $\widetilde{\gamma} = \widetilde{\mathbf{C}}'\mathbf{p}$, $\kappa = \mathbf{D}'\mathbf{p}$, $\delta = \mathbf{b}/\mathbf{p}$, $r_j = r_{jj} + 2\sum_{k>j} r_{jk}p_j^{-1/2}p_k^{1/2}$, 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 (17) reduces to the system of Engel curves

$$\begin{aligned}\mathbf{q}_i = &\left(x_i - \beta - \alpha'\overline{\mathbf{q}}_g - \widetilde{\gamma}'\widetilde{\mathbf{z}}_i - \kappa'\widetilde{\mathbf{z}}_g\right)^2 \mathbf{m} + \left(x_i - \beta - \alpha'\overline{\mathbf{q}}_g - \widetilde{\gamma}'\widetilde{\mathbf{z}}_i - \kappa'\widetilde{\mathbf{z}}_g\right)\delta \\ &+ \mathbf{r} + \mathbf{A}\overline{\mathbf{q}}_{\mathbf{g}} + \widetilde{\mathbf{C}}\widetilde{\mathbf{z}}_i + \mathbf{D}\widetilde{\mathbf{z}}_g + \mathbf{v}_g + \mathbf{u}_i,\end{aligned} \tag{41}$$

This has a very similar structure to the generic multiple equation system of equations (36), and we proceed similarly.

Define $\widetilde{\mathbf{v}}_g = \left(\alpha'\overline{\mathbf{q}}_g + \beta + \kappa'\widetilde{\mathbf{z}}_g\right)^2 \mathbf{m} - \left(\alpha'\overline{\mathbf{q}}_g + \beta + \kappa'\widetilde{\mathbf{z}}_g\right)\delta + \mathbf{r} + \mathbf{A}\overline{\mathbf{q}}_{\mathbf{g}} + \mathbf{D}\widetilde{\mathbf{z}}_{\mathbf{g}} + \mathbf{v}_g$. Then equation (41) can be rewritten more simply as

$$\mathbf{q}_i = \left(x_i - \widetilde{\gamma}'\widetilde{\mathbf{z}}_i\right)^2 \mathbf{m} - 2\left(x_i - \widetilde{\gamma}'\widetilde{\mathbf{z}}_i\right)\left(\alpha'\overline{\mathbf{q}}_g + \beta + \kappa'\widetilde{\mathbf{z}}_g\right)\mathbf{m} + \left(x_i - \widetilde{\gamma}'\widetilde{\mathbf{z}}_i\right)\delta + \widetilde{\mathbf{C}}\widetilde{\mathbf{z}}_i + \widetilde{\mathbf{v}}_g + \mathbf{u}_i, \tag{42}$$

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 (42) gives, for each good j,

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

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

Define group level instruments $\mathbf{r}_g$ as in the generic model. In particular, $\mathbf{r}_g$ can include $\widetilde{\mathbf{z}}_g$, 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 $\mathbf{r}_g$, $x_i, \widetilde{\mathbf{z}}_i, x_{i'}, \widetilde{\mathbf{z}}_{i'}$, and squares and cross products of these elements. We then, analogous to the generic model, obtain unconditional moments

$$\begin{aligned} 0 = E\{[(q_{ji} - q_{ji'}) - \left((x_i - \widetilde{\gamma}'\widetilde{\mathbf{z}}_i)^2 - (x_{i'} - \widetilde{\gamma}'\widetilde{\mathbf{z}}_{i'})^2\right) m_j - \widetilde{\mathbf{c}}_j'(\widetilde{\mathbf{z}}_i - \widetilde{\mathbf{z}}_{i'}) \\ - (\delta_j - 2m_j(\alpha'\widehat{\mathbf{q}}_{g,-ii'} + \beta + \kappa'\widetilde{\mathbf{z}}_g)) \left((x_i - \widetilde{\gamma}'\widetilde{\mathbf{z}}_i) - (x_{i'} - \widetilde{\gamma}'\widetilde{\mathbf{z}}_{i'})\right)]\mathbf{r}_{gii'}\}. \end{aligned} \quad (43)$$

Combining common terms, we have

$$\begin{aligned} 0 = E\{[(q_{ji} - q_{ji'}) - (x_i^2 - x_{i'}^2)m_j + 2\left(x_i\widetilde{\mathbf{z}}_i - x_{i'}\widetilde{\mathbf{z}}_{i'}\right)'\widetilde{\gamma}m_j - \widetilde{\gamma}'(\widetilde{\mathbf{z}}_i\widetilde{\mathbf{z}}_i' - \widetilde{\mathbf{z}}_{i'}\widetilde{\mathbf{z}}_{i'}')\widetilde{\gamma}m_j \\ - \left(\widetilde{\mathbf{c}}_j' - (\delta_j - 2m_j\beta)\widetilde{\gamma}'\right)(\widetilde{\mathbf{z}}_i - \widetilde{\mathbf{z}}_{i'}) - (\delta_j - 2m_j\beta)(x_i - x_{i'}) \\ + 2m_j\left(\alpha'\widehat{\mathbf{q}}_{g,-ii'} + \kappa'\widetilde{\mathbf{z}}_g\right)(x_i - x_{i'}) - 2\left(\widetilde{\mathbf{z}}_i - \widetilde{\mathbf{z}}_{i'}\right)'\widetilde{\gamma}m_j\left(\alpha'\widehat{\mathbf{q}}_{g,-ii'} + \kappa'\widetilde{\mathbf{z}}_g\right)]\mathbf{r}_{gii'}\}. \end{aligned} \quad (44)$$

From the above equation, for each $j = 1,...,J-1$, m_j can be identified from the variation in $(x_i^2 - x_{i'}^2)$, $\widetilde{\gamma}m_j$ can be identified from the variation in $x_i\left(\widetilde{\mathbf{z}}_{i'} - \widetilde{\mathbf{z}}_i\right)$, $\delta_j - 2m_j\beta$ and $\widetilde{\mathbf{c}}_j' - (\delta_j - 2m_j\beta)\widetilde{\gamma}'$ can be identified from the variation in $x_i - x_{i'}$ and $\widetilde{\mathbf{z}}_i - \widetilde{\mathbf{z}}_{i'}$; $m_j\alpha$ and $m_j\kappa$ are identified from the variation in $\widehat{\mathbf{q}}_{g,-ii'}\left(x_i - x_{i'}\right)$ and $\widetilde{\mathbf{z}}_g\left(x_i - x_{i'}\right)$. To summarize, $\widetilde{\gamma}$, α, κ m_j, $\delta_j - 2m_j\beta$, and $\widetilde{\mathbf{c}}_j'$ are identified for each $j = 1,...,J-1$, given sufficient variation in the covariates and instruments. Let $\eta = \delta - 2\mathbf{m}\beta$. As $\sum_{j=1}^J m_jp_j = \left(e^{-\mathbf{b}'\ln\mathbf{p}}\right)\sum_{j=1}^J d_j = 0$ and $\sum_{j=1}^J \eta_jp_j = \sum_{j=1}^J b_j = 1$, $\mathbf{m}$ and η are identified. Also $\widetilde{\mathbf{c}}_J$ can be identified from $\widetilde{\mathbf{c}}_J = \left(\widetilde{\gamma} - \sum_{j=1}^{J-1}\widetilde{\mathbf{c}}_jp_j\right)/p_J$ and hence $\widetilde{\mathbf{C}}$, $\widetilde{\gamma}$, α, κ, $\mathbf{m}$, and $\eta = \delta - 2\mathbf{m}\beta$ are identified. We now employ price variation to identify the remaining parameters.

Assume we observe data from T different price regimes. Let $\mathbf{P}$ be the matrix consisting of columns $\mathbf{p}_t$ for $t = 1,...,T$. The above Engel curve identification can be applied separately in each price regime t, so the Engel curve parameters that are functions of $\mathbf{p}_t$ are now given t subscripts.

Denote the parameters to be identified in $\mathbf{R}$ as $(r_{11},...,r_{JJ},r_{12},...,r_{J-1,J})$ and $\mathbf{b}$ as $(b_1,...,b_{J-1})$. This is a total of $[J-1+J(J+1)/2]$ parameters. Given T price regimes, we have $(J-1)T$ equations for these parameters: $\delta_{jt} = b_j/p_{jt}$, $m_{jt} = \left(e^{-\mathbf{b}'\ln\mathbf{p}_t}\right)d_j/p_{jt}$ and $\beta_t = \mathbf{p}_t^{1/2\prime}\mathbf{R}\mathbf{p}_t^{1/2}$ for each j and T, since m_{jt} and $\delta_{jt} - 2m_{jt}\beta_t$ are already identified. So for large enough T, that is, $T \geq 1 + \frac{J(J+1)}{2(J-1)}$, we get more equations than unknowns, allowing $\mathbf{R}$ and $\mathbf{b}$ to be identified given a suitable rank condition. Once $\mathbf{b}$ is identified, d_j is then identified from $d_j = p_j m_j e^{\mathbf{b}' \ln \mathbf{p}}$ for $j = 1, ..., J-1$ and $d_J = -\sum_{j=1}^{J-1} d_j$. In our data, prices vary by time and region, yielding T much higher than necessary.

We now formalize the above steps, starting from the Engel curve model without price variation. This Engel curve model is

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

from which we can construct

$$\begin{aligned}\overline{\mathbf{q}}_g &= \overline{x_g^2} \mathbf{m} + \left(\widetilde{\gamma}' \overline{\mathbf{z}\mathbf{z}'}_g \widetilde{\gamma}\right)\mathbf{m} + \mathbf{m}\left(\alpha' \overline{\mathbf{q}}_g + \kappa' \widetilde{\mathbf{z}}_g + \beta\right)^2 - 2\mathbf{m}\left(\alpha' \overline{\mathbf{q}}_\mathbf{g} + \kappa' \widetilde{\mathbf{z}}_\mathbf{g} + \beta\right)(\overline{x}_g - \widetilde{\gamma}' \overline{\mathbf{z}}_g) \\ &\quad - 2\mathbf{m}\widetilde{\gamma}' \overline{x\mathbf{z}}_g + \left(\overline{x}_g - \beta - \alpha' \overline{\mathbf{q}}_g - \widetilde{\gamma}' \overline{\mathbf{z}}_g - \kappa' \widetilde{\mathbf{z}}_g\right)\delta + \mathbf{r} + \mathbf{A}\overline{\mathbf{q}}_\mathbf{g} + \widetilde{\mathbf{C}}\overline{\mathbf{z}}_g + \mathbf{D}\widetilde{\mathbf{z}}_g + \mathbf{v}_g;\end{aligned}$$

$$\begin{aligned}\widehat{\mathbf{q}}_{g,-ii'} &= \widehat{x^2}_{g,-ii'} \mathbf{m} + (\widetilde{\gamma}' \widehat{\mathbf{z}\mathbf{z}'}_{g,-ii'} \widetilde{\gamma})\mathbf{m} + \mathbf{m}(\alpha' \overline{\mathbf{q}}_g + \kappa' \widetilde{\mathbf{z}}_g + \beta)^2 - 2\mathbf{m}(\alpha' \overline{\mathbf{q}}_g + \kappa' \widetilde{\mathbf{z}}_g + \beta)(\widehat{x}_{g,-ii'} - \widetilde{\gamma}' \widehat{\mathbf{z}}_{g,-ii'}) \\ &\quad -2\mathbf{m}\widetilde{\gamma}' \widehat{\mathbf{z}x}_{g,-ii'} + (\widehat{x}_{g,-ii'} - \beta - \alpha' \overline{\mathbf{q}}_g - \widetilde{\gamma}' \widehat{\mathbf{z}}_{g,-ii'} - \kappa' \widetilde{\mathbf{z}}_g)\delta + \mathbf{r} + \mathbf{A}\overline{\mathbf{q}}_\mathbf{g} + \widetilde{\mathbf{C}}\widehat{\mathbf{z}}_{g,-ii'} + \mathbf{v}_g + \widehat{\mathbf{u}}_{g,-ii'}.\end{aligned}$$

Hence,

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

Pairwise differencing gives

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

where the composite error is

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

Make the following assumptions.

Assumption B1: Each individual i in group g satisfies equation (41). 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.

Assumption B2: The coefficients $\mathbf{A}, \mathbf{R}, \mathbf{C} = (\widetilde{\mathbf{C}}, \mathbf{D}), \mathbf{b}, \mathbf{d}$ are unknown constants satisfying $\mathbf{b}'\mathbf{1} = 1$, $\mathbf{d}'\mathbf{1} = 0$, $\mathbf{d} \neq 0$. There exist values of $\overline{\mathbf{q}}_g$ that satisfy

$$\begin{aligned}\overline{\mathbf{q}}_g &= \overline{x_g^2}\mathbf{m} + \left(\widetilde{\gamma}'\overline{\mathbf{z}\mathbf{z}'}_g\widetilde{\gamma}\right)\mathbf{m} + \mathbf{m}\left(\alpha'\overline{\mathbf{q}}_g + \kappa'\widetilde{\mathbf{z}}_g + \beta\right)^2 - 2\mathbf{m}\left(\alpha'\overline{\mathbf{q}}_g + \kappa'\widetilde{\mathbf{z}}_g + \beta\right)\left(\overline{x}_g - \widetilde{\gamma}'\overline{\mathbf{z}}_g\right) \\ &\quad - 2\mathbf{m}\widetilde{\gamma}'\overline{x\mathbf{z}}_g + \left(\overline{x}_g - \beta - \alpha'\overline{\mathbf{q}}_g - \widetilde{\gamma}'\overline{\mathbf{z}}_g - \kappa'\widetilde{\mathbf{z}}_g\right)\delta + \mathbf{r} + \mathbf{A}\overline{\mathbf{q}}_{\mathbf{g}} + \widetilde{\mathbf{C}}\overline{\mathbf{z}}_g + \mathbf{D}\widetilde{\mathbf{z}}_g + \mathbf{v}_g. \qquad (45)\end{aligned}$$

Assumption B1 just defines the model. Assumption B2 ensures that an equilibrium exists for each group, thereby avoiding Tamer's (2003) potential incoherence problem. To see this, observe that if $A \neq 0$ then $\overline{\mathbf{q}}_g$ has the solution

$$\begin{aligned}\overline{q}_g &= \frac{1}{2m\,(Ap)^2}\{(2mAp(\overline{x}_g - \widetilde{\gamma}'\overline{\mathbf{z}}_g - \kappa'\widetilde{\mathbf{z}}_g - \beta) + 1 - A + pA\delta) \pm [(2mAp(\overline{x}_g - \widetilde{\gamma}'\overline{\mathbf{z}}_g - \kappa'\widetilde{\mathbf{z}}_g - \beta) \\ &\quad + 1 - A + pA\delta)^2 - 4m\,(Ap)^2\left(m\overline{x_g^2} + m\gamma'\overline{\mathbf{z}\mathbf{z}'}_g\gamma + m(\kappa'\widetilde{\mathbf{z}}_g + \beta)^2 - 2m(\kappa'\widetilde{\mathbf{z}}_g + \beta)(\overline{x}_g - \widetilde{\gamma}'\overline{\mathbf{z}}_g)\right. \\ &\quad \left.\left. -2m\widetilde{\gamma}'\overline{x\mathbf{z}}_g + (\overline{x}_g - \beta - \widetilde{\gamma}'\overline{\mathbf{z}}_g - \kappa'\widetilde{\mathbf{z}}_g)\,\delta + r + \widetilde{C}\overline{\mathbf{z}}_g + D\widetilde{\mathbf{z}}_g + v_g\right)\right)]^{1/2}\}, \qquad (46)\end{aligned}$$

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 (46), we can see $\overline{\mathbf{q}}_g$ is an implicit function of $\overline{x_g^2}, \overline{x}_g, \overline{\mathbf{z}}_g, \widetilde{\mathbf{z}}_g, \overline{\mathbf{z}\mathbf{z}'}_g, \overline{x\mathbf{z}}_g$, and $\mathbf{v}_g$. In the case of multiple equilibria, we do not take a stand on which root of equation (45) 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 $(x_i, \widetilde{\mathbf{z}}_i)$ is a random sample drawn from a distribution that has mean $(\overline{x}_g, \overline{\mathbf{z}}_g) = E\left((x_i, \widetilde{\mathbf{z}}_i) \mid i \in g\right)$ 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 $\sigma_{xg}^2 = Var(x_i \mid i \in g)$, $\sigma_{x\mathbf{z}g} = Cov(x_i, \widetilde{\mathbf{z}}_i \mid i \in g)$ and $\Sigma_{\mathbf{z}g} = Var(\widetilde{\mathbf{z}}_i \mid i \in g)$. Denote $\varepsilon_{ix} = x_i - \overline{x}_g$ and $\varepsilon_{iz} = \widetilde{\mathbf{z}}_i - \overline{\mathbf{z}}_g$. Assume $E\left((\varepsilon_{ix}, \varepsilon_{iz})|\overline{\mathbf{z}}_g, \widetilde{\mathbf{z}}_g, \overline{x\mathbf{z}}_g, \overline{\mathbf{z}\mathbf{z}'}_g, \overline{x}_g, \overline{x_g^2}, \mathbf{v}_g, \mathbf{r}_g\right) = 0$ 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 $(\varepsilon_{ix}, \varepsilon_{iz})$ from some distribution that is unrelated to the group level distribution, to then determine the individual level observables $x_i = \overline{x}_g + \varepsilon_{ix}$ and $\widetilde{\mathbf{z}}_i = \overline{\mathbf{z}}_g + \varepsilon_{iz}$. 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 $E(\varepsilon_{zg,-ii'} \mid x_i, \mathbf{z}_i, x_{i'}, \mathbf{z}_{i'}, \mathbf{r}_g) =$

0. With similar arguments in the generic model, Assumption B4 suffices to ensure that

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

Then we have the moment condition

$$E\{[\mathbf{q}_i - \mathbf{q}_{i'} + 2\mathbf{m}\left(\alpha'\widehat{\mathbf{q}}_{g,-ii'} + \kappa'\widetilde{\mathbf{z}}_g\right)[(x_i - x_{i'}) - \widetilde{\gamma}'(\widetilde{\mathbf{z}}_i - \widetilde{\mathbf{z}}_{i'})] - (x_i^2 - x_{i'}^2)\mathbf{m} - \widetilde{\gamma}'\left(\widetilde{\mathbf{z}}_i\widetilde{\mathbf{z}}_i' - \widetilde{\mathbf{z}}_{i'}\widetilde{\mathbf{z}}_{i'}'\right)\widetilde{\gamma}\mathbf{m} \tag{47}$$

$$+2\mathbf{m}\widetilde{\gamma}'(\widetilde{\mathbf{z}}_i x_i - \widetilde{\mathbf{z}}_{i'}x_{i'}) - \eta(x_i - x_{i'}) + (\eta\widetilde{\gamma}' - \widetilde{\mathbf{C}})(\widetilde{\mathbf{z}}_i - \widetilde{\mathbf{z}}_{i'})]|x_i, x_{i'}, \mathbf{z}_i, \mathbf{z}_{i'}, \mathbf{r}_g\} = 0$$

for the Engel curves, where $\eta = \delta - 2\mathbf{m}\beta$, and so

$$E\left[\left(\mathbf{q}_i - \mathbf{q}_{i'} + 2e^{-\mathbf{b}'\ln\mathbf{p}_t}\frac{\mathbf{d}}{\mathbf{p}_t}\left(\mathbf{p}_t'\mathbf{A}\widehat{\mathbf{q}}_{gt,-ii'} + \mathbf{p}_t'\mathbf{D}\widetilde{\mathbf{z}}_g\right)[(x_i - x_{i'}) - \mathbf{p}_t'\widetilde{\mathbf{C}}(\widetilde{\mathbf{z}}_i - \widetilde{\mathbf{z}}_{i'})] - e^{-\mathbf{b}'\ln\mathbf{p}_t}\frac{\mathbf{d}}{\mathbf{p}_t}\right.\right.$$
$$[(x_i^2 - x_{i'}^2) + \mathbf{p}_t'\widetilde{\mathbf{C}}\left(\widetilde{\mathbf{z}}_i\widetilde{\mathbf{z}}_i' - \widetilde{\mathbf{z}}_{i'}\widetilde{\mathbf{z}}_{i'}'\right)\widetilde{\mathbf{C}}'\mathbf{p}_t - 2\mathbf{p}_t'\widetilde{\mathbf{C}}(\mathbf{z}_i x_i - \mathbf{z}_{i'}x_{i'})] - \left(\frac{\mathbf{b}}{\mathbf{p}_t} - 2e^{-\mathbf{b}'\ln\mathbf{p}_t}\frac{\mathbf{d}}{\mathbf{p}_t}\mathbf{p}_t^{1/2\prime}\mathbf{R}\mathbf{p}_t^{1/2}\right)$$
$$\left.\cdot(x_i - x_{i'}) + [(\frac{\mathbf{b}}{\mathbf{p}_t} - 2e^{-\mathbf{b}'\ln\mathbf{p}_t}\frac{\mathbf{d}}{\mathbf{p}_t}\mathbf{p}_t^{1/2\prime}\mathbf{R}\mathbf{p}_t^{1/2})\widetilde{\mathbf{C}}'\mathbf{p}_t - \widetilde{\mathbf{C}}](\widetilde{\mathbf{z}}_i - \widetilde{\mathbf{z}}_{i'})|x_i, x_{i'}, \mathbf{z}_i, \mathbf{z}_{i'}, \mathbf{r}_g\right] = 0. \tag{48}$$

for the full demand system.

We define the instrument vector $\mathbf{r}_{gii'}$ to be linear and quadratic functions of $\mathbf{r}_g$, $(x_i, \mathbf{z}_i')'$, and $(x_{i'}, \mathbf{z}_{i'}')'$. Denote

$$\begin{aligned}
&L_{1jgii'} = (q_{ji} - q_{ji'}),\ \ L_{2jgii'} = \widehat{q}_{jg,-ii'}(x_i - x_{i'}),\ \ L_{3jkgii'} = \widehat{q}_{jgt,-ii'}(\widetilde{z}_{ki} - \widetilde{z}_{ki'}),\\
&L_{4k_2gii'} = \widetilde{z}_{k_2g}(x_i - x_{i'}),\ \ \ L_{5kk_2gii'} = \widetilde{z}_{k_2g}(\widetilde{z}_{ki} - \widetilde{z}_{ki'}),\ \ \ L_{6gii'} = x_i^2 - x_{i'}^2,\\
&L_{7kk'gii'} = \widetilde{z}_{ki}\widetilde{z}_{k'i} - \widetilde{z}_{ki'}\widetilde{z}_{k'i'},\ \ L_{8kgii'} = \widetilde{z}_{ki}x_i - \widetilde{z}_{ki'}x_{i'},\ \ L_{9gii'} = x_i - x_{i'},\ \ \ L_{10kgii'} = \widetilde{z}_{ki} - \widetilde{z}_{ki'},
\end{aligned} \tag{49}$$

For $\ell \in \{1j, 2j, 3jk, 4k_2, 5kk_2, 6, 7kk', 8k, 9, 10k \mid j = 1, ..., J; k, k' = 1, ..., K, k_2 = 1, ..., K_2\}$, 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

$$E\left(\mathbf{Q}_{1jg} + 2m_j\sum_{j'=1}^{J}\alpha_{j'}\mathbf{Q}_{2j'g} - 2m_j\sum_{j'=1}^{J}\sum_{k=1}^{K}\alpha_{j'}\widetilde{\gamma}_k\mathbf{Q}_{3j'kg} + 2m_j\sum_{k_2=1}^{K_2}\kappa_{k_2}\mathbf{Q}_{4k_2g} - 2m_j\sum_{k=1}^{K}\sum_{k_2=1}^{K_2}\widetilde{\gamma}_k\kappa_{k_2}\mathbf{Q}_{5kk_2g}\right.$$
$$\left.-m_j\mathbf{Q}_{6g} - m_j\sum_{k=1}^{K}\sum_{k'=1}^{K}\widetilde{\gamma}_k\widetilde{\gamma}_{k'}\mathbf{Q}_{7gkk'} + 2m_j\sum_{k=1}^{K}\widetilde{\gamma}_k\mathbf{Q}_{8kg} - \eta_j\mathbf{Q}_{9g} + \sum_{k=1}^{K}(\eta_j\widetilde{\gamma}_k - \widetilde{c}_{jk})\mathbf{Q}_{10kg}\right) = 0,$$

where $\widetilde{\gamma}_k$ is the kth element of $\widetilde{\gamma} = \widetilde{\mathbf{C}}'\mathbf{p}$, κ_{k_2} is the k_2th element of $\kappa = \mathbf{D}'\mathbf{p}$, and $\widetilde{c}_{jk}$ is the

(j,k)th element of $\widetilde{\mathbf{C}}$.

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

$$\mathbf{Q}_g = (\mathbf{Q}_{21g}, ..., \mathbf{Q}_{2Jg}, \mathbf{Q}_{311g}, ..., \mathbf{Q}_{3JKg}, \mathbf{Q}_{41g}, ..., \mathbf{Q}_{4K_2g}, \mathbf{Q}_{511g}, ..., \mathbf{Q}_{5KK_2g}, \\ \mathbf{Q}_{6g}, \mathbf{Q}_{711g}, ..., \mathbf{Q}_{7KKg}, \mathbf{Q}_{81g}, ..., \mathbf{Q}_{8Kg}, \mathbf{Q}_{9g}, \mathbf{Q}_{101g}, ..., \mathbf{Q}_{10Kg}).$$

Under Assumption B5, we can identify

$$(-2m_j\alpha', 2m_j\alpha_1\widetilde{\gamma}', ..., 2m_j\alpha_J\widetilde{\gamma}', -2m_j\kappa', 2m_j\kappa_1\widetilde{\gamma}', ..., 2m_j\kappa_{K_2}\widetilde{\gamma}', m_j, m_j\widetilde{\gamma}_1\widetilde{\gamma}', ..., m_j\widetilde{\gamma}_K\widetilde{\gamma}', \\ -2m_j\widetilde{\gamma}', \eta_j, \mathbf{c}_j' - \eta_j\widetilde{\gamma}')' = \left[E\left(\mathbf{Q}_g'\right)E\left(\mathbf{Q}_g\right)\right]^{-1}E\left(\mathbf{Q}_g'\right)E\left(\mathbf{Q}_{1jg}\right)$$

for each $j = 1, ..., J-1$. From this, α, κ, $\widetilde{\gamma}$, $\widetilde{\mathbf{C}}$, $\mathbf{m}$, and $\eta = \delta - 2\mathbf{m}\beta$ are identified. To identify the full demand system, let $\mathbf{p}_t$ denote the vector of prices in a single price regime t. Let

$$\mathbf{P} = \begin{pmatrix} \mathbf{p}_1' \\ \vdots \\ \mathbf{p}_T' \end{pmatrix}, \text{ and } \mathbf{\Lambda} = \begin{pmatrix} \Lambda_1 \\ \Lambda_2 \\ \vdots \\ \Lambda_T \end{pmatrix}$$

with the $(J-1) \times [J-1+J(J+1)/2]$ matrix

$$\Lambda_t = \begin{pmatrix} \frac{1}{p_{1t}} & 0 & \cdots & 0 & -2m_{1t}\mathbf{p}_t' & -4m_{1t}p_{1t}^{1/2}p_{2t}^{1/2} & \cdots & -4m_{1t}p_{J-1,t}^{1/2}p_{Jt}^{1/2} \\ 0 & \frac{1}{p_{2t}} & \cdots & 0 & -2m_{2t}\mathbf{p}_t' & -4m_{2t}p_{1t}^{1/2}p_{2t}^{1/2} & \cdots & -4m_{2t}p_{J-1,t}^{1/2}p_{Jt}^{1/2} \\ & & \ddots & & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & \frac{1}{p_{J-1,t}} & -2m_{J-1,t}\mathbf{p}_t' & -4m_{J-1,t}p_{1t}^{1/2}p_{2t}^{1/2} & \cdots & -4m_{J-1,t}p_{J-1,t}^{1/2}p_{Jt}^{1/2} \end{pmatrix}.$$

Then we have

$$\mathbf{PA} = \begin{pmatrix} \alpha_1' \\ \vdots \\ \alpha_T' \end{pmatrix}, \ \mathbf{PD} = \begin{pmatrix} \kappa_1' \\ \vdots \\ \kappa_T' \end{pmatrix}, \text{ and } \mathbf{\Lambda}\left(b_1, ...b_{J-1}, r_{11}, ..., r_{JJ}, r_{12}, ..., r_{J-1,J}\right)' = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_T \end{pmatrix},$$

where $\eta_t = (\eta_{1t}, ..., \eta_{J-1,t})'$. Hence, we need the $T \times J$ matrix $\mathbf{P}$ has full column rank to further identify parameters in $\mathbf{A}$ and $\mathbf{D}$; need the $(J-1)T \times [J-1+J(J+1)/2]$ matrix $\mathbf{\Lambda}$ has full column rank to identify $\mathbf{b}$ and $\mathbf{R}$. Once $\mathbf{b}$ is identified, we can identify $\mathbf{d}$. Using the groups that are observed facing this set of prices, from above we can identity all parameters in $\mathbf{A}$, $\widetilde{\mathbf{C}}$, $\mathbf{D}$, $\mathbf{b}$, $\mathbf{d}$, and $\mathbf{R}$.

Assumption B6: Data are observed in T price regimes $\mathbf{p}_1$, ..., $\mathbf{p}_T$ such that the $T \times J$

matrix $\mathbf{P} = (\mathbf{p}_1, ..., \mathbf{p}_T)'$ and the $(J-1)T \times [J-1+J(J+1)/2]$ matrix Λ both have full column rank.

Given Assumption B6, $\mathbf{A}$ and $\mathbf{D}$ are identified by

$$\mathbf{A} = (\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}' \begin{pmatrix} \alpha_1' \\ \vdots \\ \alpha_T' \end{pmatrix} \text{ and } \mathbf{D} = (\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}' \begin{pmatrix} \kappa_1' \\ \vdots \\ \kappa_T' \end{pmatrix};$$

$\mathbf{R}$ and $\mathbf{b}$ are identified by

$$(b_1, ...b_{J-1}, r_{11}, ..., r_{JJ}, r_{12}, ..., r_{J-1,J})' = (\mathbf{\Lambda}'\mathbf{\Lambda})^{-1}\mathbf{\Lambda}' \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_T \end{pmatrix};$$

$\mathbf{d}$ is identified by $d_j = p_{jt} m_{jt} e^{\mathbf{b}' \ln \mathbf{p}_t}$ for $j = 1, ..., J$ and $d_J = -\sum_{j=1}^{J-1} d_j$.

To illustrate, in the two goods system, i.e., $J = 2$, this means that we can identify $\mathbf{A}$ and $\mathbf{D}$ if the $T \times 2$ matrix

$$\mathbf{P} = \begin{pmatrix} p_{11}, p_{21} \\ \vdots \\ p_{1T}, p_{2T} \end{pmatrix}$$

has rank 2 and the $T \times 4$ matrix

$$\mathbf{\Lambda} = \begin{pmatrix} \frac{1}{p_{11}}, & -2e^{-\mathbf{b}' \ln \mathbf{p}_1} \frac{d_1}{p_{11}} p_{11}, & -2e^{-\mathbf{b}' \ln \mathbf{p}_1} \frac{d_1}{p_{11}} p_{21}, & -4e^{-\mathbf{b}' \ln \mathbf{p}_1} \frac{d_1}{p_{11}} p_{11}^{1/2} p_{21}^{1/2} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{1}{p_{1T}}, & -2e^{-\mathbf{b}' \ln \mathbf{p}_T} \frac{d_1}{p_{1T}} p_{1T}, & -2e^{-\mathbf{b}' \ln \mathbf{p}_T} \frac{d_1}{p_{1T}} p_{2T}, & -4e^{-\mathbf{b}' \ln \mathbf{p}_T} \frac{d_1}{p_{1T}} p_{1T}^{1/2} p_{2T}^{1/2} \end{pmatrix}$$

has rank 4.

The above derivation proves the following theorem:

Theorem 2. Given Assumptions B1-B5, the parameters $\widetilde{\mathbf{C}}$, α, $\widetilde{\gamma}$, κ, $\mathbf{m}$, and $\eta = \delta - 2\mathbf{m}\beta$ in the Engel curve system (41) are identified. If Assumption B6 also holds, all the parameters $\mathbf{A}$, $\mathbf{b}$, $\mathbf{R}$, $\mathbf{d}$, $\widetilde{\mathbf{C}}$ and $\mathbf{D}$ in the full demand system (17) are identified.

8.6 Estimation of the Demand System with Fixed Effects

For the full demand system, the GMM estimation builds on the above, treating each value of gt as a different group, so the total number of relevant 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 linear and quadratic functions of $\mathbf{r}_{gt}$, $(x_i, \mathbf{z}_i')'$, and $(x_{i'}, \mathbf{z}_{i'}')'$. The GMM estimator, using group level clustered standard errors, is then

$$\left(\widehat{\mathbf{A}}_1', ..., \widehat{\mathbf{A}}_J', \widehat{b}_1, ..., \widehat{b}_{J-1}, \widehat{d}_1, ..., \widehat{d}_{J-1}, \widehat{\widetilde{\mathbf{c}}}_1', ...\widehat{\widetilde{\mathbf{c}}}_J', , \widehat{\mathbf{D}}_1', ...\widehat{\mathbf{D}}_J', r_{11}, ...r_{JJ}, r_{12}, ..., r_{J-1J}\right)'$$
$$= \arg\min \left(\frac{\sum_{t=1}^T \sum_{g=1}^G \sum_{(i,i')\in\Gamma_{gt}} \mathbf{m}_{gtii'}}{\sum_{t=1}^T \sum_{g=1}^G \sum_{(i,i')\in\Gamma_{gt}} 1}\right)' \widehat{\Omega} \left(\frac{\sum_{t=1}^T \sum_{g=1}^G \sum_{(i,i')\in\Gamma_{gt}} \mathbf{m}_{gtii'}}{\sum_{t=1}^T \sum_{g=1}^G \sum_{(i,i')\in\Gamma_{gt}} 1}\right),$$

where the expression of $\mathbf{m}_{gtii'} = (\mathbf{m}_{1gtii'}', ..., \mathbf{m}_{J-1,gtii'}')$ is

$$\mathbf{m}_{jgtit'} = [(q_{ji} - q_{ji'}) - \left((x_i - \widetilde{\gamma}_t'\widetilde{\mathbf{z}}_i)^2 - (x_{i'} - \widetilde{\gamma}_t'\widetilde{\mathbf{z}}_{i'})^2\right) m_{jt} - \widetilde{\mathbf{c}}_j'(\widetilde{\mathbf{z}}_i - \widetilde{\mathbf{z}}_{i'})$$
$$- (\delta_{jt} - 2m_{jt}(\alpha_t'\widehat{\mathbf{q}}_{g,-ii'} + \beta_t + \kappa_t'\widetilde{\mathbf{z}}_{gt}))\left((x_i - \widetilde{\gamma}_t'\widetilde{\mathbf{z}}_i) - (x_{i'} - \widetilde{\gamma}_t'\widetilde{\mathbf{z}}_{i'})\right)]\mathbf{r}_{gtii'}$$

with

$$m_{jt} = e^{-\mathbf{b}'\ln \mathbf{p}_t}\frac{d_j}{p_{jt}},\ \alpha_t = \mathbf{A}'\mathbf{p}_t,\ \widetilde{\gamma}_t = \widetilde{\mathbf{C}}'\mathbf{p}_t,\ \kappa_t = \mathbf{D}'\mathbf{p}_t,\ \beta_t = \mathbf{p}_t^{1/2\prime}\mathbf{R}\mathbf{p}_t^{1/2},\ \delta_{jt} = \frac{b_j}{p_{jt}}.$$

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 earlier 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 $(\varepsilon_{ix}, \varepsilon_{iz})$ 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

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

because

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

and

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

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 $\mathbf{r}_{gii'} = \mathbf{r}_g \otimes (\mathbf{x}_i^* - \mathbf{x}_{i'}^*, \mathbf{x}_i^*\mathbf{x}_i^{*\prime} - \mathbf{x}_{i'}^*\mathbf{x}_{i'}^{*\prime})$, 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 $\mathbf{r}_{gtii'} = \mathbf{r}_{gt} \otimes (\mathbf{x}_i^* - \mathbf{x}_{i'}^*, \mathbf{x}_i^*\mathbf{x}_i^{*\prime} - \mathbf{x}_{i'}^*\mathbf{x}_{i'}^{*\prime})$, 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 Identification and Estimation of the Demand System with Random Effects

The Engel curve model with random effects is

$$\mathbf{q}_i = x_i^2\mathbf{m} + (\widetilde{\gamma}'\widetilde{\mathbf{z}}_i\widetilde{\mathbf{z}}_i'\widetilde{\gamma})\,\mathbf{m} - 2\mathbf{m}\widetilde{\gamma}'\widetilde{\mathbf{z}}_i x_i + \mathbf{m}\left(\alpha'\overline{\mathbf{q}}_g + \kappa'\widetilde{\mathbf{z}}_g + \beta\right)^2 - 2\mathbf{m}\left(\alpha'\overline{\mathbf{q}}_g + \kappa'\widetilde{\mathbf{z}}_g + \beta\right)(x_i - \widetilde{\gamma}'\widetilde{\mathbf{z}}_i)$$
$$+ \left(x_i - \beta - \alpha'\overline{\mathbf{q}}_g - \widetilde{\gamma}'\widetilde{\mathbf{z}}_i - \kappa'\widetilde{\mathbf{z}}_g\right)\delta + \mathbf{r} + \mathbf{A}\overline{\mathbf{q}}_g + \widetilde{\mathbf{C}}\widetilde{\mathbf{z}}_i + \mathbf{D}\widetilde{\mathbf{z}}_g + \mathbf{v}_g + \mathbf{u}_i,$$

Therefore,

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

By rewriting q_{ji} as

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

where

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

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

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

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

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

where $v_{j0} = \mu_j - m_j\alpha'\mathbf{\Sigma}_v\alpha$ 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

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

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 $i \in g$,

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

Denote

$$L_{1jgi} = q_{ji},\ L_{2jj'gi} = \frac{1}{n_g - 1} \sum_{i' \in g, i' \neq i} \widehat{q}_{jg,-ii'} q_{j'i'},\ L_{3gi} = x_i^2,\ L_{4kk'gi} = \widetilde{z}_{ki}\widetilde{z}_{k'i},\ L_{5k_2k_2'gi} = \widetilde{z}_{k_2g}\widetilde{z}_{k_2'g},$$
$$L_{6kgi} = \widetilde{z}_{ki}x_i,\quad L_{7k_2gi} = \widetilde{z}_{k_2g}x_i,\quad L_{8jgi} = \widehat{q}_{jg,-i}x_i,\quad L_{9jkgi} = \widehat{q}_{jg,-i}\widetilde{z}_{ki},\ L_{10jk_2gi} = \widehat{q}_{jg,-i}\widetilde{z}_{k_2g},$$
$$L_{11kk_2gi} = \widetilde{z}_{ki}\widetilde{z}_{k_2g},\ L_{12jgi} = \widehat{q}_{jg,-i},\quad L_{13gi} = x_i,\quad L_{14kgi} = \widetilde{z}_{ki},\quad L_{15k_2gi} = \widetilde{z}_{k_2g},\quad L_{16gi} = 1.$$

For $\ell \in \{1j, 2jj', 3, 4kk', 5k_2k_2', 6k, 7k_2, 8j, 9jk, 10jk_2, 11kk_2, 12j, 13, 14k, 15k_2, 16 \mid j, j' = 1, ..., J; k, k' = 1, ..., K; k_2, k_2' = 1, ..., K_2\}$, 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

$$E\Bigg(\mathbf{H}_{1jg} - m_j \sum_{j=1}^{J}\sum_{j'=1}^{J} \alpha_{j'}\alpha_j \mathbf{H}_{2jj'g} - m_j \mathbf{H}_{3g} - m_j \sum_{k=1}^{K}\sum_{k'=1}^{K} \widetilde{\gamma}_k\widetilde{\gamma}_{k'} \mathbf{H}_{4kk'g} - m_j \sum_{k_2=1}^{K_2}\sum_{k_2'=1}^{K_2} \kappa_{k_2}\kappa_{k_2'} \mathbf{H}_{5k_2k_2'g}$$
$$+ 2m_j \sum_{k=1}^{K} \widetilde{\gamma}_k \mathbf{H}_{6kg} + 2m_j \sum_{k_2=1}^{K_2} \kappa_{k_2} \mathbf{H}_{7k_2g} + 2m_j \sum_{j'=1}^{J} \alpha_{j'} \mathbf{H}_{8j'g} - 2m_j \sum_{j'=1}^{J}\sum_{k=1}^{K} a_{j'}\widetilde{\gamma}_k \mathbf{H}_{9j'kg}$$
$$- 2m_j \sum_{j'=1}^{J}\sum_{k_2=1}^{K_2} a_{j'}\kappa_{k_2} \mathbf{H}_{10j'k_2g} - 2m_j \sum_{k=1}^{K}\sum_{k_2=1}^{K_2} \widetilde{\gamma}_k\kappa_{k_2} \mathbf{H}_{11kk_2g} + \sum_{j'=1}^{J} [(\delta_j - 2m_j\beta)\alpha_{j'} - A_{jj'}] \mathbf{H}_{12j'g}$$
$$+ (2m_j\beta - \delta_j)\mathbf{H}_{13g} + \sum_{k=1}^{K} [(\delta_j - 2m_j\beta)\widetilde{\gamma}_k - c_{jk}] \mathbf{H}_{14kg} + \sum_{k_2=1}^{K_2} [(\delta_j - 2m_j\beta)\kappa_{k_2} - D_{jk_2}] \mathbf{H}_{15k_2g} - \xi_j \mathbf{H}_{16g}\Bigg) = 0,$$

where $\xi_j = m_j\beta^2 - \delta_j\beta + r_j + v_{j0}$.

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

$$\mathbf{H}_g = (\mathbf{H}_{211g}, ..., \mathbf{H}_{2JJg}, \mathbf{H}_{3g}, \mathbf{H}_{411g}, ..., \mathbf{H}_{4KKg}, \mathbf{H}_{511g}, ..., \mathbf{H}_{5K_2K_2g}, \mathbf{H}_{61g}, ..., \mathbf{H}_{6Kg},$$
$$\mathbf{H}_{71g}, ..., \mathbf{H}_{7K_2g}, \mathbf{H}_{81g}, ..., \mathbf{H}_{8Jg}, \mathbf{H}_{911g}, ..., \mathbf{H}_{9JKg}, \mathbf{H}_{1011g}, ..., \mathbf{H}_{10JK_2g}, \mathbf{H}_{1111g}, ..., \mathbf{H}_{11KK_2g},$$
$$\mathbf{H}_{121g}, ..., \mathbf{H}_{12Jg}, \mathbf{H}_{13g}, \mathbf{H}_{141g}, ..., \mathbf{H}_{14Kg}, \mathbf{H}_{151g}, ..., \mathbf{H}_{15K_2g}, \mathbf{H}_{16g}).$$

Under Assumptions B1-B4 and Assumption C, we can identify

$$(m_j\alpha_1\alpha', ..., m_j\alpha_J\alpha', m_j, m_j\widetilde{\gamma}_1\widetilde{\gamma}', ..., m_j\widetilde{\gamma}_K\widetilde{\gamma}', m_j\kappa_1\kappa', ..., m_j\kappa_{K_2}\kappa', -2m_j\widetilde{\gamma}', -2m_j\kappa', -2m_j\alpha',$$
$$2m_j\widetilde{\gamma}_1\alpha', ..., 2m_j\widetilde{\gamma}_K\alpha', 2m_j\kappa_1\alpha', ..., 2m_j\kappa_{K_2}\alpha', 2m_j\kappa_1\widetilde{\gamma}', ..., 2m_j\kappa_{K_2}\widetilde{\gamma}', \mathbf{A}'_j - (\delta_j - 2m_j\beta)\alpha', \delta_j - 2m_j\beta,$$
$$\mathbf{c}_j - (\delta_j - 2m_j\beta)\widetilde{\gamma}, \mathbf{D}_j - (\delta_j - 2m_j\beta)\kappa, m_j\beta^2 - \delta_j\beta + r_j + v_{j0})' = \left[E\left(\mathbf{H}'_g\right)E\left(\mathbf{H}_g\right)\right]^{-1} E\left(\mathbf{H}'_g\right)E\left(\mathbf{H}_{1jg}\right).$$

for each $j = 1, ..., J-1$. From this, $\widetilde{\gamma}$, κ, α, $\mathbf{m}$, $\eta = \delta - 2\mathbf{m}\beta$, $\mathbf{A}_j$, $\widetilde{\mathbf{c}}_j$, $\mathbf{D}_j$, and $m_j\beta^2 - \delta_j\beta + r_j + v_{j0}$ for $j = 1, ..., J-1$ are all identified. Then, $\mathbf{A}_J = \left(\alpha - \sum_{j=1}^{J-1}\mathbf{A}_j p_j\right)/p_J$, $\widetilde{\mathbf{c}}_J = (\widetilde{\gamma} - \sum_{j=1}^{J-1}\widetilde{\mathbf{c}}_j p_j)/p_J$, and $\mathbf{D}_J = (\kappa - \sum_{j=1}^{J-1}\mathbf{D}_j p_j)/p_J$ are identified. Here without price variation, we can identify $\mathbf{A}$ and $\mathbf{D}$. 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, and only $\widetilde{\mathbf{C}}$ can be identified from the cross product of $\overline{\mathbf{q}}_{\mathbf{g}}$ and $(x_i, \widetilde{\mathbf{z}}_i)$. Furthermore, to identify the structural parameters $\mathbf{b}$, $\mathbf{d}$, and $\mathbf{R}$, we need the rank condition in Assumption B6(2).

With our data spanning multiple time regimes t, we estimate the full demand system model simultaneously over all values of t, instead of as Engel curves separately in each t as above. To do so, in the above moments we replace the Engel curve coefficients α, β, $\widetilde{\gamma}$, κ, δ, r_j, and $\mathbf{m}$ with their corresponding full demand system expressions, i.e., $\alpha = \mathbf{A}'\mathbf{p}$, $\beta = \mathbf{p}^{1/2\prime}\mathbf{R}\mathbf{p}^{1/2}$, etc, and add t subscripts wherever relevant. The resulting GMM estimator based on these moments (and estimated using group level clustered standard errors), is then

$$(\widehat{\mathbf{A}}'_1, ..., \widehat{\mathbf{A}}'_J, \widehat{b}_1, ..., \widehat{b}_{J-1}, \widehat{d}_1, ..., \widehat{d}_{J-1}, \widehat{\widetilde{\mathbf{c}}}'_1, ...\widehat{\widetilde{\mathbf{c}}}'_J, , \widehat{\mathbf{D}}'_1, ...\widehat{\mathbf{D}}'_J, \widehat{R}_{11}, ...\widehat{R}_{JJ}, \widehat{R}_{12}, ..., \widehat{R}_{J-1J},$$
$$\widehat{\mu}, \widehat{\mathbf{\Sigma}}_{v,11}, ..., \widehat{\mathbf{\Sigma}}_{v,JJ}, \widehat{\mathbf{\Sigma}}_{v,12}, ..., \widehat{\mathbf{\Sigma}}_{v,J-1,J},)'$$
$$= \arg\min\left(\frac{\sum_{t=1}^{T}\sum_{g=1}^{G}\sum_{i\in\Gamma_{gt}}\mathbf{m}_{gti}}{\sum_{t=1}^{T}\sum_{g=1}^{G}\sum_{i\in\Gamma_{gt}}1}\right)'\widehat{\Omega}\left(\frac{\sum_{t=1}^{T}\sum_{g=1}^{G}\sum_{i\in\Gamma_{gt}}\mathbf{m}_{gti}}{\sum_{t=1}^{T}\sum_{g=1}^{G}\sum_{i\in\Gamma_{gt}}1}\right),$$

where the expression of $\mathbf{m}_{gti} = (\mathbf{m}'_{1gti}, ..., \mathbf{m}'_{J-1,gti})$ is

$$\mathbf{m}_{jgti} = \{q_{ji} - m_{jt}\alpha'_t\widehat{\mathbf{q}}_{gt,-ii'}\alpha'_t\mathbf{q}_{i'} - m_{jt}(x_i - \widetilde{\gamma}'_t\widetilde{\mathbf{z}}_i)^2 - m_{jt}(\kappa'_t\widetilde{\mathbf{z}}_{gt} + \beta_t)^2$$
$$+ [(2m_{jt}(x_i - \widetilde{\gamma}'_t\widetilde{\mathbf{z}}_i - \kappa'_t\widetilde{\mathbf{z}}_{gt} - \beta_t) + \delta_{jt})\alpha'_t - \mathbf{A}'_j]\widehat{\mathbf{q}}_{gt,-ii'} + 2m_{jt}(\kappa'_t\widetilde{\mathbf{z}}_g + \beta_t)(x_i - \widetilde{\gamma}'_t\widetilde{\mathbf{z}}_i)$$
$$- \delta_{jt}(x_i - \beta_t - \widetilde{\gamma}'_t\widetilde{\mathbf{z}}_i - \kappa'_t\widetilde{\mathbf{z}}_{gt}) - r_{jt} - \widetilde{\mathbf{c}}'_j\widetilde{\mathbf{z}}_i - \mathbf{D}'_j\widetilde{\mathbf{z}}_g - v_{jt0}\}\mathbf{r}_{gti}$$

with

$$m_{jt} = e^{-\mathbf{b}'\ln\mathbf{p}_t}\frac{d_j}{p_{jt}},\ \alpha_t = \mathbf{A}'\mathbf{p}_t,\ \widetilde{\gamma}_t = \widetilde{\mathbf{C}}'\mathbf{p}_t,\ \kappa_t = \mathbf{D}'\mathbf{p}_t,\ \beta_t = \mathbf{p}_t^{1/2\prime}\mathbf{R}\mathbf{p}_t^{1/2},$$

$$\eta_{jt} = \frac{b_j}{p_{jt}} - 2m_{jt}\mathbf{p}_t^{1/2\prime}\mathbf{R}\mathbf{p}_t^{1/2},\ \delta_{jt} = \frac{b_j}{p_{jt}},\ r_{jt} = R_{jj} + 2\sum_{k>j} R_{jk}\sqrt{p_{kt}/p_{jt}},$$

$$v_{jt0} = \mu_{jt} - e^{-\mathbf{b}'\ln\mathbf{p}_t}\frac{d_j}{p_{jt}}\sum_{j_1=1}^{J}\sum_{j_2=1}^{J}\sum_{j=1}^{J}\sum_{j'=1}^{J} A_{j_1j}p_{j_1t}A_{j_2j'}p_{j_2t}\mathbf{\Sigma}_{vt,jj'}.$$

Note that v_{jt0} are constants for each value of j and t, that must be estimated along with the other parameters. In our data T is large (since prices vary both by time and district). To reduce the number of required parameters and thereby increase efficiency, assume that $\mu = E(\mathbf{v}_{gt})$ and $\mathbf{\Sigma}_v = Var(\mathbf{v}_{gt})$ do not vary by t. Then we can replace v_{jt0} with

$$v_{jt0} = \mu_j - e^{-\mathbf{b}'\ln\mathbf{p}_t}\frac{d_j}{p_{jt}}\sum_{j_1=1}^{J}\sum_{j_2=1}^{J}\sum_{j=1}^{J}\sum_{j'=1}^{J} A_{j_1j}p_{j_1t}A_{j_2j'}p_{j_2t}\mathbf{\Sigma}_{v,jj'}$$

Moreover, since $\mathbf{v}_{gt}$ represents deviations from the utilty derive demand functions, it may be reasonable to assume that $\mu = 0$. With these substitutions we only need to estimate the parameters $\mathbf{\Sigma}_v$ instead of all the separate v_{jt0} constants.

Table 1 Descriptive Statistics: Indian Consumption Data from NSS Round 61

3259 groups	Observations (N=70,217)				Pairs (N=3,009,614)			
	mean	std dev	min	max	mean	std dev	min	max
x_i	1.15	0.72	0.07	8.58	1.15	0.71	0.07	8.58
q_i luxuries	0.30	0.38	0.00	7.70	0.29	0.38	0.00	7.70
q_i neccessities	0.84	0.42	0.02	4.18	0.83	0.41	0.02	4.18
$\widehat{q}_{g,-ii'}$ luxuries					0.26	0.18	0.00	3.46
$\widehat{q}_{g,-ii'}$ neccessities					0.79	0.21	0.13	3.51
p luxuries	1.00	0.08	0.87	1.29	1.01	0.08	0.87	1.29
p neccessities	1.01	0.08	0.90	1.25	1.03	0.08	0.90	1.25
Dalit/ST	0.24	0.43	0.00	1.00	0.18	0.38	0.00	1.00
Muslim/Other	0.23	0.42	0.00	1.00	0.27	0.45	0.00	1.00
Educ med	0.42	0.49	0.00	1.00	0.44	0.50	0.00	1.00
Educ high	0.09	0.29	0.00	1.00	0.04	0.19	0.00	1.00

Table 2: Subjective well-being summary statistics

	Mean	SD	Min	Max
Life satisfaction	3.07	1.22	1.00	5.00
Imputed expenditure, CPI deflated	2.20	1.44	0.70	9.51
Group expenditure, CPI deflated	3.86	1.30	1.70	10.60
Household size	4.06	1.85	1.00	10.00
Age	40.81	14.53	18.00	93.00
Married (=1)	0.84	0.37	0.00	1.00
Non-Hindu (=1)	0.24	0.42	0.00	1.00
Primary education (=1)	0.10	0.29	0.00	1.00
Secondary education (=1)	0.14	0.35	0.00	1.00
Observations	3236			

Life satisfaction variable from World Values Survey. Participants asked "All things considered, how satisfied are you with your life as a whole these days?", and asked to point to a position on a ladder. Coded as 1-5 in 2006, and 1-10 in 2014. We collapsed to a 1-5 scale in 2014. Income measured in thousands of Rs/month. Excluded categories are less than primary education, and Hindu religion.

Table 3: Luxury spending as a function of group spending, generic model estimates

	RE			FE		
	(1)	(2)	(3)	(4)	(5)	(6)
a						
Constant	-0.246*	-0.239**	-0.186*	-0.233**	-0.663***	-0.602***
	(0.134)	(0.115)	(0.111)	(0.116)	(0.141)	(0.149)
b						
Constant	0.478***	0.453***	0.464***	0.426***	0.438***	0.438***
	(0.017)	(0.014)	(0.014)	(0.010)	(0.012)	(0.028)
d						
Constant	0.290***	0.319***	0.294***	0.307***	0.341***	0.327***
	(0.030)	(0.035)	(0.032)	(0.034)	(0.040)	(0.054)
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
Group controls	No	Yes	Yes	No	Yes	Yes
Price controls	No	No	Yes	No	No	Yes
Hausman for β				-0.04	26.75	17.47
P-value				1.00	0.00	0.00
Number of groups	3259	3259	3259	3259	3259	3259
Number of within-group pairs	3009614	3009614	3009614	3009614	3009614	3009614

Dependent variable is household luxury spending. Individual controls include household size, age, marital status and amount of land owned. Group controls include religion indicators and education indicators. Price controls are laspeyres indices for luxury and nonluxury spending. Standard errors in parentheses and clustered at the group level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Satisfaction on household and peer income

	OLS (SDs)		Ordered logit	
	(1)	(2)	(3)	(4)
main				
Imputed expenditure	0.068***		0.179***	
	(0.013)		(0.031)	
Group expenditure	-0.100**		-0.203*	
	(0.049)		(0.115)	
Imputed expenditure, CPI deflated		0.131***		0.335***
		(0.025)		(0.058)
Group expenditure, deflated		-0.190*		-0.424*
		(0.107)		(0.256)
Year FEs	Yes	Yes	Yes	Yes
P(Own + group = 0)	0.528	0.588	0.848	0.734
Dependent mean	0.00	0.00	3.07	3.07
Dependent SD	1.00	1.00	1.22	1.22
Observations	3236	3236	3236	3236

Dependent variable as noted in column header, in SD. Subjective well being data from World Values Survey, imputations from NSS. Peer groups defined as intersection of education (below primary, primary or partial secondary, secondary+) and religion (Hindu and non-Hindu). All columns include controls for household size, age, sex, marital status and education. Standard errors in parentheses and clustered at the group level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Summary Statistics for Indian NSS Data 2354 group-rounds with 10 or more obs of Hindu non-SC/ST households								
	Observations (N=56,516)				Pairs (N=2,055,776)			
	mean	std dev	min	max	mean	std dev	min	max
x_i	1.12	0.66	0.10	8.75	1.08	0.64	0.10	8.75
q_i luxuries	0.31	0.37	0.00	7.96	0.30	0.36	0.00	7.96
q_i necessities	0.83	0.40	0.03	4.32	0.79	0.38	0.03	4.32
$\widehat{q}_{g,-ii'}$ luxuries					0.26	0.15	0.02	1.78
$\widehat{q}_{g,-ii'}$ neccessities					0.74	0.17	0.26	1.83
p luxuries	0.98	0.08	0.81	1.29	0.99	0.08	0.81	1.29
p neccessities	0.99	0.07	0.86	1.34	1.00	0.07	0.86	1.34
Educ med	0.48	0.50	0.00	1.00	0.50	0.50	0.00	1.00
Educ high	0.06	0.24	0.00	1.00	0.03	0.17	0.00	1.00
(hhsize-1)/10	0.40	0.22	0.00	1.10	0.39	0.22	0.00	1.10
headage/120	0.40	0.11	0.17	0.94	0.40	0.11	0.17	0.94
married	0.87	0.34	0.00	1.00	0.87	0.34	0.00	1.00
ln(land+1)	0.60	0.58	0.00	2.30	0.53	0.55	0.00	2.30
ration card	0.23	0.42	0.00	1.00	0.26	0.44	0.00	1.00
q_i vis luxuries	0.13	0.23	0.00	7.54	0.13	0.23	0.00	7.54
q_i invis luxuries	0.18	0.22	0.00	5.07	0.17	0.21	0.00	5.07
q_i vis necessities	0.13	0.09	0.00	2.37	0.12	0.08	0.00	2.37
q_i invis necessities	0.70	0.34	0.01	3.98	0.67	0.32	0.01	3.98
$\widehat{q}_{g,-ii'}$ vis luxuries					0.11	0.08	0.00	1.12
$\widehat{q}_{g,-ii'}$ inv luxuries					0.16	0.08	0.01	1.35
$\widehat{q}_{g,-ii'}$ vis necessities					0.11	0.04	0.02	0.49
$\widehat{q}_{g,-ii'}$ inv necessities					0.63	0.14	0.22	1.53
p vis luxuries	0.95	0.11	0.64	1.33	0.95	0.11	0.64	1.33
p invis luxuries	0.98	0.08	0.82	1.28	1.00	0.08	0.82	1.28
p vis necessities	0.98	0.14	0.70	1.50	1.01	0.15	0.70	1.50
p invis necessities	0.99	0.06	0.86	1.34	1.00	0.06	0.86	1.34

Table 6: 2 good system, Fixed Effects					
		Fixed Effects			
	Needs	A Same		A Diagonal	
	Response	est	*std err*	est	*std err*
luxuries	own	0.50	*0.11*	-2.63	*0.40*
neccessities	own	0.50	*0.11*	2.99	0.28
test A same	χ^2 stat, *[p-val]*			80	*[0.00]*
Hausman test RE	z stat, *[p-val]*	-0.31	*[0.76]*	-7.8	*[0.00]*
				8.8	*[0.00]*

Table 7: 2 good system, Random Effects							
		Random Effects					
	Needs	A Same		A Diagonal		A Full	
	Response	est	std err	est	std err	est	*std err*
luxuries	own	0.55	*0.02*	0.46	*0.02*	0.20	*0.09*
neccessities	own	0.55	*0.02*	0.57	*0.02*	1.09	*0.10*
luxuries	cross					0.42	*0.08*
neccessities	cross					-0.33	*0.11*
test A same	χ^2 stat, *[p-val]*			43	*[0.00]*		

Table 8: 4 Goods							
		Fixed Effects		Random Effects			
		FE: A Same		RE: A same		RE: A Diag	
		est	*std err*	est	*std err*	est	*std err*
luxuries	visible	0.71	*0.05*	0.65	*0.01*	0.54	0.01
	invisible	0.71	*0.05*	0.65	*0.01*	0.62	0.01
neccessities	visible	0.71	*0.05*	0.65	*0.01*	0.76	0.01
	invisible	0.71	*0.05*	0.65	*0.01*	0.66	0.01
Hausman test RE (Asame)	z stat, *[p-val]*	1.26	*[0.21]*				
test Asame for RE	χ^2 stat, *[p-val]*					658	*[0.00]*

Table 9: Fixed Effects, Same A, Subgroups				
	Religion separate regs		Education Hindus only	
	est	*std err*	est	*std err*
Hindu, non-SC/ST	0.50	*0.11*		
SC/ST	0.13	*0.18*		
non-Hindu	-0.06	*0.23*		
Illiterate/Barely Literate			0.08	*0.15*
Primary or some Secondary			0.56	*0.12*
Completed Secondary or more			0.37	*0.22*