Consumption peer effects and utility needs in India

Authors: Arthur Lewbel, Samuel Norris, Krishna Pendakur, Xi Qu

Boston College Working Papers in Economics, 2018

Consumption Peer Effects and Utility Needs in India*

Arthur Lewbel, Samuel Norris, Krishna Pendakur, and Xi Qu
Boston College, U. of Chicago, Simon Fraser U., and Shanghai Jiao Tong U.

original July 2014, revised August 2018

Abstract

We construct a peer effects model of consumption where mean expenditures of consumers in one's peer group affects one's utility through perceived consumption needs. We show model identification with standard household-level consumer expenditure survey microdata, even when most members of each peer group are not observed. We find that in India, each additional rupee spent by one's peers increases one's perceived needs, thereby reducing money metric utility, by 0.5 rupees. One implication is that welfare gains of hundreds of billions of rupees per year might be possible by replacing private government transfers with the provision of public goods.

*We thank Chris Muris and Niharika Singh for comments on this work. We also thank seminar participants at Harvard, Boston University, University of Pennsylvania, Vanderbilt, and Rice. Norris and Pendakur are grateful for financial support from the Social Sciences and Humanities Research Council of Canada through its Doctoral Fellowship Awards and Insight Research Grants Programs, respectively.

I Introduction

It is well established that there are substantial peer effects in income and consumption. People's evaluation of their own income depends on the income of their peers (Kahneman 1992; Luttmer 2004; Clark Frijters, and Shields 2008). Their consumption choices also depend on those of their peers (Boneva 2013; de Giorgi, Frederiksen, and Pistaferri 2016), their evaluation of those consumption choices depends on the consumption of those in their peer groups (Gali 1994, Maurer and Meier 2008), and the perceived value of individual goods or brands depends on the consumption of those goods in relevent reference groups (Rabin 1998, Kalyanaram and Winer 1998, Chao and Schor 1998).

Despite the strong evidence showing the existence of peer effects in consumption, there has been much less work evaluating their economic costs in lost welfare. In this paper we study the effect of changes in peer mean expenditures on money metric utility, asking how much one's own expenditure would have to increase to compensate for a one unit change in peer expenditures. The results have very large implications for redistribution policies, e.g., we find with data from India that welfare gains on the order of billions of US dollars per year might be possible by modifying just one existing India government transfer program.

One way to measure the economic costs of peer effects would be to directly regress an observed utility measure (i.e., stated well being data) on own and peer expenditures, as in Luttmer (2005). This has the drawback of relying on coarse self-reports of well-being which may suffer from framing biases, measurement errors and problems of interpretation.

Most empirical consumption peer effects studies instead directly model individual consumption as a function of average peer group consumption and other covariates. See, e.g., Chao and Schor (1998) or Boneva (2013). However, while such regressions can reveal behavioral responses to peer expenditures, without a structural model they say nothing about the utility and welfare implications of these peer effects.

We propose a structural model that uses revealed preference methods to recover the utility implications of peer expenditures on consumption behavior. As in classical demand analysis, this model relates observable consumption decisions to underlying money metric utility, but additionally allows peer expenditure to affect welfare. By studying how consumption decisions vary across and within peer groups with different mean expenditures, the model backs out an estimate of the money-metric cost of peer consumption.

Much progress has been made in overcoming the endogeneity of peer effects by the use of detailed social network information. For example, de Giorgi, Frederiksen and Pistaferri (2016) instrument for peer consumption with information on friend-of-friend consumption. However, in our application we use only standard repeated cross section consumer expenditure survey data, of the type that is commonly collected by many governments all over the world. As a result, we cannot make use of detailed network information like variation in peer group sizes (as in Lee 2007) to obtain identification. Indeed, most members of each group are not observed in our data. This gives rise to some unusual econometric issues that must be overcome.

In our keeping-up-with-the-Joneses type model, one's perceived required expenditures, or "needs," depend on, among other things, the average expenditures of one's peer group. The higher are these perceived needs, the more one must spend to attain the same level of utility. Consistent with other empirical evidence (e.g., Luttmer 2005, Ravina 2008, and Clark and Senik 2010), our model implies that consumers lose utility from feeling poorer when their peers get richer. One feels that one needs more when one's peers have more.

We estimate the model using consumption data from India. Our main finding is that an average decrease in spending by one's peers of two rupees has the same effect on one's utility as an increase in one's own expenditures of about one rupee.

This result has enormous implications for tax and redistribution policy. It suggests that consumption or income taxes may be far less costly in terms of social welfare and utility than is implied by standard demand model estimates that ignore peer effects (see, e.g., Boskin and Shoshenski 1978). To illustrate, suppose you experience a two rupee tax increase. If your peers also have their taxes increase by the same amount, then your loss in utility will be equivalent to that of just a one rupee tax increase. If the utility associated with public goods are not subject to these peer effects, then government can increase welfare at far lower cost by using taxes to increase expenditures on public goods instead of by redistributing via transfers of money or private goods.

To assess the magnitude of these effects, we perform a rough calculation which shows that replacing India's "Public Distribution System" food subsidy program with more generous provision of public goods, such as education or cleaner air and water, could increase money metric welfare by over 300 billion rupees (4.4 billion US dollars) per year at no additional cost.

We also find some evidence that these peer effects may be smaller for lower socio-economic status groups. If so, then transfers from higher to lower status groups can also increase total welfare, by reducing peer effect externalities. The usual argument for transfers of money from rich to poor (and more generally for progressive tax rates) is the belief that the poor have a higher marginal utility of money, but that is hard to verify and quantify. Our results suggest potentially large gains from such transfers and from progressive taxes, even if all consumers have the same marginal utility of money, and even if social welfare is inequality-neutral.

The remainder of this introduction lays out our peer effects model more explicitly, and describes the econometric obstacles to identification and estimation that our model must address.

I.A Modeling Needs, Consumption Decisions and Utility

Consider a model where i indexes consumers, g indexes groups of consumers, overbars indicate true within-group means, and hats indicate sample averages. Let $\mathbf{q}_i$ be the vector of quantities of goods that consumer i consumes (in continuous quantities). There is a long history going back to Samuelson (1947) of modeling needs in utility functions as analogous to fixed costs or overheads in production,

$$U_i = U(\mathbf{q}_i - \mathbf{f}_i) \tag{1}$$

where U_i is the attained utility level of consumer i, U is a utility function, and $\mathbf{f}_i$ is a vector of needs. The vector $\mathbf{f}_i$ is a quantity vector, equal to the minimum quantity consumer i must consume of each good before he or she starts to get any utility from consumption. For now, let the utility function U be common to all consumers, though in our actual model we will introduce both observed and unobserved heterogeneity at both the individual and group levels.

In the context of a linear model, Samuelson (1947) defines the quantity vector $\mathbf{f}_i$ as the "necessary set" of goods. The Stone (1954) and Geary (1949) linear expenditure system is just Cobb-Douglas utility function U with constant needs $\mathbf{f}$. Gorman (1976) introduced the general model of equation (1) for arbitrary utility functions U, letting $\mathbf{f}_i$ depend on a vector of demographic variables or other taste shifters $\mathbf{z}_i$. In our model, we let needs $\mathbf{f}_i$ also depend on $\overline{\mathbf{q}}_g$, the mean value of the expenditure vector $\mathbf{q}$ among the members of consumer i's peer group g. The model therefore has[1]

$$\mathbf{f}_i = \mathbf{f}\left(\mathbf{z}_i, \overline{\mathbf{q}}_g\right)$$

for a given needs function $\mathbf{f}$. Let $\mathbf{p}$ be the price vector corresponding to $\mathbf{q}_i$, and let x_i be consumer i's budget (total expenditures). Assuming consumer i chooses the vector $\mathbf{q}_i$ to maximize his or her utility function $U\left(\mathbf{q}_i - \mathbf{f}\left(z_i, \overline{\mathbf{q}}_g\right)\right)$ under the linear budget constraint $\mathbf{p}'\mathbf{q}_i \leq x_i$, we can calculate the consumer's resulting demand functions, expressing $\mathbf{q}_i$ as a function of $\mathbf{p}$, x_i, and $\mathbf{f}\left(z_i, \overline{\mathbf{q}}_g\right)$.

Given these demand functions, we can answer the question: If peer spending $\overline{\mathbf{q}}_g$ increases, how much poorer does consumer i feel? More precisely, how much more would consumer i

[1] It is of course possible that peer group expenditures matter in other ways than just though group means $\overline{\mathbf{q}}_g$. We only consider group means here because of data limitations and other econometric issues discussed later.

need to spend (how much would his or her budget x_i need to increase) to give that consumer the same level of utility she had before $\overline{\mathbf{q}}_g$ increased? This conception of how peer consumption affects individual utility generalizes the existing literature, which typically models utility as being affected by mean peer expenditure $\mathbf{p}'\overline{\mathbf{q}}_g$ (e.g., Luttmer 2005). In our empirical implementation we find that this parsimonious specification does a good job of capturing how peer consumption affects utility, and so we focus our attention on this model.

I.B Econometric Obstacles

Starting from the utility function with needs defined by equation (1), and introducing suitable unobserved heterogeneity across consumers, by revealed preference we derive demand equations to estimate of the general form

$$\mathbf{q}_i = \mathbf{h}\left(\mathbf{p}, x_i - \mathbf{p}'\mathbf{f}\left(z_i, \overline{\mathbf{q}}_g\right)\right) + \mathbf{f}\left(z_i, \overline{\mathbf{q}}_g\right) + \mathbf{v}_g + \mathbf{u}_i. \quad (2)$$

Here $\mathbf{h}$ is a vector valued function (quadratic in our empirical application) that is based on the utility function U, $\mathbf{v}_g$ is a vector of group level fixed effects or random effects, and $\mathbf{u}_i$ is a vector of idiosyncratic errors. This is an example of a social interactions model, since it includes the group mean $\overline{\mathbf{q}}_g$ as a vector of regressors.

Our model differs from standard social interactions models (e.g., Manski 1993, 2000, Brock and Durlauf 2001, Lee 2007, and Blume, Brock, Durlauf, and Ioannides 2010), in a variety of ways. First, our model is nonlinear and vector-valued while most such models are linear and scalar-valued. This nonlinearity helps to overcome the reflection problem in peer effect models like ours. However, while this nonlinearity is a necessary consequence of utility maximization with empirically plausible demand functions, it exacerbates and complicates the effect of measurement error on coefficient estimates.

The issue of measurement error interacts with a second, larger, difference between our work and the previous literature. We estimate our model from standard consumer expenditure survey data, which is the type of data that many countries collect for constructing consumer price indices. Since such surveys do not contain social network data, we define peer groups based on demographic characteristics. And as a result of this being survey data, we can only observe a small number of the members in each peer group.

Most analyses of social interactions models use network information to help identify the model. Examples include the use of exogenous variation in group composition or size (e.g., Lee 2007, Carrell, Fullerton and West 2009, and Duflo, Dupas and Kremer 2011), or the use of detailed network structure data like intransitive triads, where essentially data on friends of friends provides instruments for identification (e.g., Bramoullé, Djebbari and Fortin 2009;

de Giorgi, Frederiksen, and Pistaferri 2016).

In contrast, we don't observe group sizes and don't observe network information like friends of friends, and so we cannot make use of these existing methods of identifying and estimating the model. We also do not observe most members of each group, and so do not come close to observing $\overline{\mathbf{q}}_g$. We can at best construct an estimate $\widehat{\mathbf{q}}_g$, by averaging across the small number of members that we do observe in each group. This greatly complicates identification and estimation of our model, because replacing $\overline{\mathbf{q}}_g$ with $\widehat{\mathbf{q}}_g$ introduces group level measurement error into the model, and this measurement error $\widehat{\mathbf{q}}_g - \overline{\mathbf{q}}_g$ is endogenous and correlated with other components of the model. The measurement error is further exacerbated by potential nonlinearity of $\mathbf{h}$, resulting in errors that contain interaction terms like $\left(\widehat{\mathbf{q}}_g - \overline{\mathbf{q}}_g\right) x_i$.

Our main econometric contribution in this paper is to show how these identification issues can be solved, and how consistent estimates of peer effects can be extracted from ordinary consumer expenditure data. We find that these group mean measurement error issues are so large that ignoring them results in underestimating the true peer effect by up to 70% in some specifications.

The remainder of the paper proceeds as follows. In Section II we expand on the structural model of utility, demand and peer effects introduced in Section I.A. Section III illustrates our general procedure for dealing with the above econometric issues in the main text using a simple quadratic model. This procedure should be of independent interest to others wishing to estimate peer effects using survey data. We prove in the appendix that the procedure also works for our more general demand functions. Results are presented in Section IV, and policy implications in Section V. Section VI concludes.

II Utility and Demand With Peer Effects in 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 (2007) 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 (Kahneman and Tversky 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 this literature include Kahneman (1992) and Clark, Frijters, and Shields (2008).

Taken together, this literature suggests that the utility of consumer i should depend on both $\mathbf{q}_i$ and $\overline{\mathbf{q}}_g$, and that utility is increasing in $\mathbf{q}_i$ and decreasing in $\overline{\mathbf{q}}_g$.[2] If we could observe utility, we could directly test this. Luttmer (2005) estimates an approximation of this relationship, by regressing a crude measure of utility (reported life satisfaction on a coarse ordinal scale) not on $\mathbf{q}_i$ and $\overline{\mathbf{q}}_g$, but on x_i and $\overline{x}_g$. In a preliminary data analysis we estimate a similar regression using data from India and groups that are roughly comparable to those in our main empirical analysis. Our preliminary analyses agree with Luttmer (2005) in support of our model's underlying assumption that increases in peer expenditures decrease rather than increase utility. Our main model will not depend on crude utility measures, but will instead identify comparable structural parameters obtained from utility-derived demand functions via revealed preference.

A number of papers relate consumption choices to peer consumption levels, although these analyses are essentially nonstructural (Chao and Schor 1998, Boneva 2013, de Giorgi, Frederiksen and Pistaferri, 2016). All these papers suggest that the magnitudes of peer effects in consumption choices are large. In our notation, these papers are analogous to regressing $\mathbf{q}_i$ on x_i and $\overline{\mathbf{q}}_g$. However, establishing how much consumption $\mathbf{q}_i$ changes when peer consumption $\overline{\mathbf{q}}_g$ changes does not answer the welfare question of how $\overline{\mathbf{q}}_g$ affects utility, and hence how much x_i would need to increase to compensate for the loss of utility from an increase in $\overline{\mathbf{q}}_g$. Answering this type of welfare question requires linking expenditures to utility, which is what our structural model does.

II.A The Model

Ignoring unobserved preference and demand heterogeneity for now, we begin with utility given by

$$U_i = U\left(\mathbf{q}_i - \mathbf{f}_i\right) \quad \text{where} \quad \mathbf{f}_i = \mathbf{f}\left(\mathbf{z}_i, \overline{\mathbf{q}}_g\right). \tag{3}$$

[2] One could imagine alternative models where utility is increasing in $\overline{\mathbf{q}}_g$, such as being happy for the success of your peers. But the empirical evidence, including Luttmer (2005) and the results we present below, is that the effect is negative.

One can equivalently represent preferences using an indirect utility function, defined as the maximum utility attainable with a given budget x_i when facing prices $\mathbf{p}$. Gorman (1976) shows that for a utility function in the form of equation (3), given any U function that satisfies standard regularity conditions of utility maximization,[3] there exists a corresponding indirect utility function V such that

$$U_i = V\left(\mathbf{p}, x_i - \mathbf{p}'\mathbf{f}(\mathbf{z}_i, \overline{\mathbf{q}}_g)\right) \tag{4}$$

Blackorby and Donaldson (1994) show that indirect utility functions of the form $U_i = V(\mathbf{p}, x_i - \mathbf{p}'\mathbf{f}_i)$ have a desirable property for social welfare calculations known as Absolute Equivalence Scale Exactness (AESE). AESE structures interpersonal comparisons of utilities. Equation (4) invokes equality, rather than just ordinal equality, of indirect utility across differing values of the arguments of $\mathbf{f}$. For preferences that satisfy AESE, Blackorby and Donaldson (1994) define equivalent income as $x_i - \mathbf{p}'\mathbf{f}_i$ and show that the sum of equivalent income across consumers is a valid money-metric based social welfare function. Gorman (1976) and Blackorby and Donaldson (1994) obtained their results without the presence of $\overline{\mathbf{q}}_g$ in $\mathbf{f}_i$, but one can immediately verify that their same derivations go through with $\overline{\mathbf{q}}_g$ included along with $\mathbf{z}_i$.

Blackorby and Donaldson (1994) and Donaldson and Pendakur (2006) show that the function $\mathbf{f}$ without $\overline{\mathbf{q}}_g$ is uniquely identified up to location from consumer behaviour. The responses of $\mathbf{f}$ to changes in $\mathbf{z}_i$ can therefore be identified from consumer behavior. We show below that we can also identify, from consumer demand data, how $\mathbf{f}$ responds to changes in $\overline{\mathbf{q}}_g$. We can therefore show how money metric social welfare responds to changes in average peer group expenditure levels.

Luttmer (2005) directly estimates a simplified version of equation (4) where each U_i is a self-reported measure of happiness. His model sets self reported U_i equal to a function of a linear index in x_i, $\boldsymbol{z}_i$ and $\overline{x}_g$ (the within-group average income). Under the assumption that these reported happiness measures are ordinally fully comparable, which is a strong form of interpersonal comparability, his method recovers the welfare cost of group expenditures, scaled against the welfare gain of own expenditures.[4]

Luttmer's model specification implies that, in equation (4), $\mathbf{f}(\mathbf{z}_i, \overline{\mathbf{q}}_g) = a\overline{\mathbf{q}}_g + \mathbf{C}\mathbf{z}_i$ where a is a scalar (since $\overline{x}_g = \boldsymbol{p}'\overline{\mathbf{q}}_g$). In our work, we will allow for the possibility that the needs function takes this form. Luttmer finds that a is 0.76, meaning that a 100 dollar increase in

[3]See Deaton and Muellbauer (1980) for a summary of the regularity conditions associated with direct and indirect utility functions.

[4]Our welfare calculations, which do not assume u_i can be observed, instead rely on a weaker notion of comparability, ratio scale comparability. See Blackorby and Donaldson (1994).

group-average income has the same effect on utility as a 76 dollar reduction in own-income, motivating his title "Neighbors as Negatives." We estimate an object analogous to Luttmer's a coefficient, but instead of assuming that U_i equals an observed happiness measure, we let U_i be unobserved. We derive demand equations from the associated utility function and recover the implied peer effects on utility using revealed preference methods.

The demand functions that result from maximizing the utility function in equation (3) can be obtained by applying Roy's (1947) identity to the indirect utility function of equation (4). The resulting demand functions have the form

$$\mathbf{q}_i = \mathbf{h}(\mathbf{p}, x - \mathbf{p}'\mathbf{f}_i) + \mathbf{f}_i,$$

where $\mathbf{f}_i = \mathbf{f}(\mathbf{z}_i, \overline{\mathbf{q}}_g)$ and the vector valued function $\mathbf{h}$ is defined by $\mathbf{h}(\mathbf{p}, x) = -\frac{\nabla_{\mathbf{p}} V(\mathbf{p},x)}{\nabla_x V(\mathbf{p},x)}$. This structure is called demographic translation (Pollak and Wales 1981) or quantity shape invariance (Pendakur 2005).

We take the function $\mathbf{f}$ to be linear, so

$$\mathbf{f}_i = \mathbf{A}\overline{\mathbf{q}}_g + \mathbf{C}\mathbf{z}_i \tag{5}$$

for some matrices of parameters $\mathbf{A}$ and $\mathbf{C}$. Linearity of $\mathbf{f}_i$ in $\mathbf{z}_i$ is commonly assumed in empirical demand analysis; we extend that linearity to the additional variables $\overline{\mathbf{q}}_g$. If $\boldsymbol{A}$ is restricted to be diagonal with all elements equal to the scalar α, we then get demand equations analogous to those that would come from Luttmer's (2005) utility model.

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 $l \in g$.

The terms $\mathbf{v}_g + \mathbf{u}_i$ can be interpreted either as departures from utility maximization by individuals, or as unobserved preference heterogeneity. Assuming that the price weighted sum $\mathbf{p}'(\mathbf{v}_g + \mathbf{u}_i)$ equals 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 parameters. In our analysis we do not take a stand on whether $\mathbf{v}_g + \mathbf{u}_i$ represents preference heterogeneity or departures from utility maximization.

In both the fixed and random effects specifications, $\mathbf{v}_g$ is allowed to vary by time as well as by group. In the fixed effects model, the group level fixed effect $\mathbf{v}_g$ is permitted to correlate with other regressors like $\mathbf{p}$ and $\overline{\mathbf{q}}_g$. As is familiar from other contexts, the random effects model is much more efficient at the cost of the additional restriction that $\mathbf{v}_g$ is independent of regressors.

The above derivations yield demand functions of the general form

$$\mathbf{q}_i = \mathbf{h}(\mathbf{p}, x_i - \mathbf{p}'\mathbf{A}\overline{\mathbf{q}}_g - \mathbf{p}'\mathbf{C}\mathbf{z}_i) + \mathbf{A}\overline{\mathbf{q}}_g + \mathbf{C}\mathbf{z}_i + \mathbf{v}_g + \mathbf{u}_i. \tag{6}$$

What remains in the demand specification is to choose the indirect utility function V, which then determines the vector-valued function $\mathbf{h}$.

A long empirical literature on commodity demands finds that observed demand functions are close to polynomial (Lewbel 1991, Banks, Blundell, and Lewbel 1997). Gorman (1981) shows that any polynomial demand system must have a maximum rank of three. Lewbel (1989) provides the tractable classes of indirect utility functions that yield rank three polynomials. The most commonly assumed rank three models in empirical practice are quadratic (see the above references and the Quadratic Expenditure System of Pollak and Wales 1978). The resulting class of indirect utility functions that yield rank three, quadratic in x demand functions have the form

$$V(\mathbf{p}, x) = -(x - R(\mathbf{p}))^{-1} B(\mathbf{p}) - D(\mathbf{p}) \tag{7}$$

for some differentiable functions R, B and D. Applying Roy's identity to obtain the function $\boldsymbol{h}$ and equation (6), we obtain demand equations

$$\begin{aligned} \mathbf{q}_i = {} & \left(x_i - R(\mathbf{p}) - \mathbf{p}'(\mathbf{A}\overline{\mathbf{q}}_g + \mathbf{C}\mathbf{z}_i)\right)^2 \frac{\nabla D(\mathbf{p})}{B(\mathbf{p})} \\ & + \left(x_i - R(\mathbf{p}) - \mathbf{p}'(\mathbf{A}\overline{\mathbf{q}}_g + \mathbf{C}\mathbf{z}_i)\right) \frac{\nabla B(\mathbf{p})}{B(\mathbf{p})} + \nabla R(\mathbf{p}) + \mathbf{A}\overline{\mathbf{q}}_g + \mathbf{C}\mathbf{z}_i + \mathbf{v}_g + \mathbf{u}_i. \end{aligned} \tag{8}$$

We assume homogeneity—the absence of money illusion— which is a necessary condition for rationality of preferences. This requires that $R(\mathbf{p})$ and $B(\mathbf{p})$ be homogeneous of degree 1 in $\mathbf{p}$ and that $D(\mathbf{p})$ be homogeneous of degree 0 in $\mathbf{p}$. Specifications of the price functions that satisfy these restrictions and yield price flexible (in the sense of Diewert 1974) demand 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 $\mathbf{d}'\mathbf{1} = 0$. See Lewbel (1997) for a survey of these demand function properties.

For each good j, the resulting demand model is

$$q_{ji} = Q_j\left(\mathbf{p}, x_i, \overline{\mathbf{q}}_g, \mathbf{z}_i\right) + v_{jg} + u_{ji}, \tag{9}$$

where each Q_j function is given by

$$Q_j\left(\mathbf{p}, x_i, \overline{\mathbf{q}}_g, \mathbf{z}_i\right) = \left(x_i - \mathbf{p}^{1/2\prime}\mathbf{R}\mathbf{p}^{1/2} - \mathbf{p}'\mathbf{A}\overline{\mathbf{q}}_g - \mathbf{p}'\mathbf{C}\mathbf{z}_i\right)^2 e^{-\mathbf{b}'\ln\mathbf{p}}\frac{d_j}{p_j}$$
$$+\left(x_i - \mathbf{p}^{1/2\prime}\mathbf{R}\mathbf{p}^{1/2} - \mathbf{p}'\mathbf{A}\overline{\mathbf{q}}_g - \mathbf{p}'\mathbf{C}\mathbf{z}_i\right)\frac{b_j}{p_j} + R_{jj} + \sum_{k\neq j} R_{jk}\sqrt{p_k/p_j} + \mathbf{A}_j'\overline{\mathbf{q}}_g + \mathbf{C}_j'\mathbf{z}_i \tag{10}$$

Here $\mathbf{A}_j'$ is row j of $\mathbf{A}$ and $\mathbf{C}_j'$ is row j of $\mathbf{C}$. These quantity demand functions are quadratic in the budget x_i.

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

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.[5]

In all of the above, different consumers can be observed in different time periods (no consumer is observed more than once). 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 $t \in \{1, 2, ..., T\}$, so we add a t subscript to every group level variable above.

While we report some results using $J = 4$ goods, most of our analyses will be based on $J = 2$, with the two goods being luxuries and necessities. In this case, we only need to estimate the demand equation for good 1 (luxuries). Much of our analyses will also assume $\mathbf{A}$ and $\mathbf{R}$ are both diagonal.[6] In fact, our baseline specification assumes that all diagonal elements of $\mathbf{A}$ are equal, and the non-diagonal elements are zero. This specification parsimoniously captures the effect of total peer group expenditure on welfare, following

[5]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 (education dummies) 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.

[6]Like most modern continuous demand models (e.g., Banks, Blundell, and Lewbel, 1997; 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 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.

previous work on the topic.

With these simplifications, equation (10) reduces to a single equation[7]:

$$Q_1\left(\mathbf{p}, x_i, \overline{\mathbf{q}}_g, \mathbf{z}_i\right) = X_i^2 e^{-(b_1 \ln p_{1t} + (1-b_1)\ln p_{2t})} d_1/p_{1t} + X_i b_1/p_{1t} + R_{11} + A_{11}\overline{q}_{g1t} + \mathbf{C}_1'\mathbf{z}_i,$$

where

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

As is common in empirical work in demand analysis, we recast quantity demand equations as spending equations by multiplying by price. Substituting the above into (9) and multiplying by p_{1t} yields our primary estimation model:

$$p_{1t}q_{1i} = X_i^2 e^{-(b_1 \ln p_{1t} + (1-b_1)\ln p_{2t})} d_1 + X_i b_1 + R_{11}p_{1t} + A_{11}p_{1t}\overline{q}_{g1t} + \mathbf{C}_1' p_{1t}\mathbf{z}_i + p_{1t}v_{1gt} + p_{1t}u_{1i}. \tag{12}$$

The goal will be estimation of the set of parameters $\{\mathbf{A}, \mathbf{C}, \mathbf{R}, \mathbf{d}, \mathbf{b}\}$. In particular, the matrix $\mathbf{A}$ embodies the impact of peer effects on needs, and hence on social welfare.

III Identification and Estimation: Econometric Issues

There are many obstacles to identifying and estimating our model. These issues stem from: 1) model nonlinearity (which arises from utility maximization); 2) the presence of fixed or random effects $\mathbf{v}_g$ without panel data; 3) the possible absence of an equilibrium among group members; 4) the fact that $\overline{\mathbf{q}}_g$ is endogenous (as in the Manski 1993 reflection problem); and, 5) $\overline{\mathbf{q}}_g$ cannot be directly observed nor consistently estimated, because the data only contain a small number of members of each group.

To illustrate how we overcome these econometric issues, we first consider a very simple model that suffers from all these same problems. We show how we can identify and estimate this simple, generic model. Formal proofs of the identification and estimator consistency of this simple generic model are provided in the Appendix, along with the extension of these methods to the identification and estimation of our full consumer demand model given by (12).

We have repeated cross section data. In each of a small number of time periods, we observe a sample of individuals. Each individual i is only observed once, so different individuals are observed in each time period. To save on notation, we drop the time subscript for now. Assume each individual i is in a peer group $g \in \{1, ...G\}$. The number of peer groups

[7] Since $J = 2$, Q_2 does not need to be estimated because its parameters are given by $Q_2 = (x_i - p_1 Q_1)/p_2$,

G is large, so we assume $G \to \infty$. In our data we will only observe a small number n_g of the individuals who are actually in each peer group g, so asymptotics assuming $n_g \to \infty$ are inapplicable. We therefore assume n_g is fixed and so does not grow with the sample size.

Our simple illustrative, generic model relates a scalar outcome y_i for person i in group g to $\overline{y}_g$, where $\overline{y}_g = E(y_j \mid j \in g)$, so $\overline{y}_g$ is the population mean value of y_j over all people j in person i's peer group g. For simplicity, the generic model assumes we have a single scalar covariate x_i that affects y_i. A typical peer specification with such data would be linear, e.g.,

$$y_i = \overline{y}_g a + x_i b + u_i \tag{13}$$

where u_i is an error term uncorrelated with x_i, and the pair of constants (a, b) are parameters to estimate (see, e.g., Manski 1993, 2000 and Brock and Durlauf 2001). However, to account for nonlinearity and heterogeneity issues associated with our demand model, consider the more general specification

$$y_i = \left(\overline{y}_g a + x_i b\right)^2 d + \left(\overline{y}_g a + x_i b\right) + v_g + u_i \tag{14}$$

where the term v_g is a group level fixed or random effect, and the constants (a, b, d) are the parameters to identify and estimate.

We are *not* claiming that the functional form of equation (14) is in some way fundamental or essential (though it does include the standard linear model as a special case). Rather, it is just a simple nonlinear specification that can be used to demonstrate all the issues (and solutions) associated with identification and estimation of our utility derived demand model. Because of its simplicity, it may also be useful to other researchers working on peer effects models in other contexts.

Equation (14) differs from equation (13) in two important ways. First, it allows for nonlinearity. In our application, potential nonlinearity is an unavoidable consequence of utility maximization with empirically plausible demand functions. Our generic model has this nonlinearity in the form of a squared linear index because that is a simple specification that resembles the nonlinearity in our real demand equation (12).

The second way equation (14) differs from equation (13) is with the inclusion of a group-level fixed or random effect v_g. In social interaction models, typical ways of obtaining identification in the presence of such effects is to exploit specialized data that includes observable network structures like "intransitive triads" (Bramoullé, Djebbari, and Fortin 2009, Jochmans and Weidner 2016, and de Giorgi, Frederiksen, and Pistaferri 2016). We do not have access to such network information in our data. Alternatively, one might obtain identification using common panel data methods, such as by differencing out fixed effects over time. However, we only have repeated cross section rather than true panel data, and it is important to allow the fixed or random effects v_g to vary by time, because many demand determinants vary by time.

Next, 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 an estimate $\widehat{y}_g$ making equation (14) equal to

$$y_i = (\widehat{y}_g a + x_i b)^2 d + (\widehat{y}_g a + x_i b) + v_g + u_i + \varepsilon_{gi} \tag{15}$$

where error term ε_{gi} is an additional error term. By construction, ε_{gi} is given by

$$\varepsilon_{gi} = \left(\overline{y}_g^2 - \widehat{y}_g^2\right) a^2 d + 2\left(\overline{y}_g - \widehat{y}_g\right) x_i abd + \left(\overline{y}_g - \widehat{y}_g\right) a. \tag{16}$$

Note that this ε_{gi} error depends on $\overline{y}_g$, $\widehat{y}_g$ and on x_i, which creates endogeneity issues.

Inspection of equations (15) and (16) shows many of the obstacles to identifying and estimating the model parameters a, b, and d. First, v_g can be correlated with $\overline{y}_g$ and $\widehat{y}_g$. 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 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.

The primary obstacles to identification and estimation will be dealing with the above correlations between observed covariates like $\widehat{y}_g$ and x_i, and the model unobservables like v_g and ε_{gi}. In contrast, two additional problems that are common in social interactions and network models will be more readily overcome. One is the Manski (1993) reflection problem, which does not arise here primarily because the group mean of x_i does not appear in the model.[8] Another possible problem is that the model might not have an equilibrium. For example, it could be that some members increasing their spending by one dollar causes others to spend more by two dollars, making the original members feel the need to increase further to three dollars, etc. In the Appendix we show that a single inequality ensures existence of an equilibrium. Roughly, an equilibrium exists as long as the peer effects are not too large.

We employ two somewhat different methods for identifying and estimating this model, depending on whether each v_g is assumed to be a fixed effect or a random effect. For each

[8]The group mean $\overline{x}_g$ does not appear in our model because our underlying utility theory of revealed preference with needs only gives rise to group quantities (corresponding to $\overline{y}_g$ in the generic model) in the model. When v_g is a fixed effect the reflection problem could still arise, in that v_g could be correlated with $\overline{x}_g$, but in that case we exploit the nonlinear structure of the model to overcome this issue. See the Appendix for details.

case, we construct a set of moment conditions that suffice to identify the coefficients, and are used for estimation via GMM.

III.A Generic Model With Group-Time Fixed Effects

In the fixed effects model, we make no assumptions about how v_g may correlate with other covariates, or about how v_g might vary over time. Identification and estimation will therefore require removing these fixed effects in some way. As a result, identification will depend on specific features of our functional form, and so for example will require that $d \neq 0$. In contrast, our later random effects model will make additional assumptions regarding v_g, but will be applicable to any linear or quadratic specification.

We cannot difference across time to remove v_g, so we begin by looking at the difference between the outcomes of two people i and i' observed in the same time period and in the same group g. In addition, to remove remaining correlation issues, we 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 $\widehat{y}_{g,-ii'}$ is just the sample average of y for everyone who is observed in group g in the given time period, except for the individuals i and i'. Let $\widehat{y}_g$ from equations (15) equal the estimator $\widehat{y}_{g,-ii'}$. Then differencing equations (15) and (16) between the individuals i and i' gives

$$y_i - y_{i'} = 2\widehat{y}_{g,-ii'} \left(x_i - x_{i'}\right) abd + \left(x_i^2 - x_{i'}^2\right) b^2 d + \left(x_i - x_{i'}\right) b + u_i - u_{i'} + \varepsilon_{gi} - \varepsilon_{gi'}. \quad (17)$$

where

$$\varepsilon_{gi} - \varepsilon_{gi'} = 2\left(\overline{y}_g - \widehat{y}_{g,-ii'}\right)\left(x_i - x_{i'}\right) abd. \quad (18)$$

We can then show that (see Theorem 1 in the Appendix), with these definitions along with some standard regression assumptions,

$$E\left(u_i - u_{i'} + \varepsilon_{gi} - \varepsilon_{gi'} \mid x_i, x_{i'}\right) = 0, \quad (19)$$

which we can then use to construct moments for estimation of equation (17).

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, $\varepsilon_{gi} - \varepsilon_{gi'}$ is linear in $x_i - x_{i'}$, with a coefficient that, given some exogeneity assumptions can be shown to be conditionally mean zero.

Equation (17) contains functions of $\widehat{y}_{g,-ii'}$, x_i, and $x_{i'}$ as regressors, and equation (19) shows that we can use functions of x_i and $x_{i'}$ as instruments. 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. Fortunately, we have repeated cross section data. No particular consumer is sampled more than once, but we have observations of other consumers in the same group from different time periods. These consumers may or may not have different fixed effects v_g and different mean expenditures $\overline{y}_g$, but all we need to assume about them is the exogeneity assumption that each x_i is independent of the idiosyncratic error $u_{i'}$ of every person i' in person $i's$ group, and that the average group value $\widehat{x}_g$ is autocorrelated over time (see the derivation of Theorem 1 in the appendix for details). We take (functions of) these observations of $\widehat{x}_g$ from other time periods to be the instruments for (functions of) $\widehat{y}_{g,-ii'}$ that we require.

Note that even if our survey data only came from a single cross section, other data sets might also provide these group level data. For example, if x_i is a demographic variable, then census data could provide the needed group level data. Similarly, if x_i is a consumption budget as in our application, then average group level income data from wage or income surveys could suffice.

Let $\mathbf{r}_g$ denote the vector of observations of $\widehat{x}_g$ for every time period in our data other than the time period of the cross section under consideration (or include group level variables that correlate with $\widehat{x}_g$ from other data sources). Let $\mathbf{r}_{gii'}$ denote the vector of $x_i, x_{i'}, \mathbf{r}_g$, and squares and cross products of these variables. We then obtain the unconditional moments

$$E\left[\left(y_i - y_{i'} - 2\widehat{y}_{g,-ii'}\left(x_i - x_{i'}\right)abd - \left(x_i^2 - x_{i'}^2\right)b^2 d - \left(x_i - x_{i'}\right)b\right)\mathbf{r}_{gii'}\right] = 0. \tag{20}$$

Based on equation (20), the parameters a, b, and d, can now be estimated using Hansen's (1982) GMM estimator. Each observation consists of a pair of individuals observed in a given group in a given time period, so our data consists of all such pairs i and i'. The estimator is equivalent to regressing each $y_i - y_{i'}$ on the variables $\widehat{y}_{g,-ii'}\left(x_i - x_{i'}\right)$, $(x_i^2 - x_{i'}^2)$, and $(x_i - x_{i'})$, using GMM with instruments $\mathbf{r}_{gii'}$, and then recovering the parameters a, b, and d, from the estimated coefficients. By construction, the errors in this model are correlated across all pairs of individuals within each group, so we must use clustered standard errors, clustered at the group level, to obtain proper inference.

Theorem 1 in Appendix A.2 describes these results formally and extends this model to a vector $\mathbf{x}_i$, provides formal conditions for proving that an equilibrium exists, and shows that the parameters of the model are identified by GMM using these moments. We then further extend this result in Appendix A.3 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. Theorem 2 in the Appendix A.5 then gives a final extension of these results, showing identification and consistent estimation of our utility derived demand model, given by equations (9) and (10) for each good j. This demand model has a more complicated functional form than the generic model and, e.g., includes prices that vary by time but not necessarily by group, but the method for obtaining identification and constructing the associated GMM estimator is the same.

III.B Generic Model With Group-Time 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 add the additional assumptions that v_g is homoskedastic and independent of x_i, and provide moments for a GMM estimator that does not entail differencing.

Another drawback of the fixed effects model is that it depends on specific features of the functional form for identification, e.g, it requires the nonlinearity of $d \neq 0$. In contrast, our random effects model identification and estimation works with linear models (i.e., with $d = 0$), and can also be shown to hold for a general quadratic model, though for simplicity we will stick with the specification of equation (14) here, since our utility derived demand model has a similar structure.

To describe the random effects estimator it will be convenient to rewrite equation (14) as

$$y_i = \overline{y}_g^2 a^2 d + (a + 2x_i abd)\,\overline{y}_g + \left(x_i b + x_i^2 b^2 d\right) + v_g + u_i. \tag{21}$$

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 $\overline{y}_g^2$ 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 construct $\widehat{y}_{g,-ii'}$ as before. Now suppose we replace $\overline{y}_g$ with $\widehat{y}_{g,-ii'}$ as before, again introducing the additional error ε_{gi} as in the previous section. The problem now is that the term $\widehat{y}_{g,-ii'}^2 - \overline{y}_g^2$ in ε_{gi} is not differenced out, and this term 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 now 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 use the GMM estimator to essentially average the resulting moments over all individuals i' in g, thereby once again exploiting all of the information in the group. With this replacement, equation (21) 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 by construction the error $\widetilde{\varepsilon}_{gii'}$ has the form

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

In Appendix A.4 we show that $E(\widetilde{\varepsilon}_{gii'} | x_i, \mathbf{r}_g) = -da^2 Var\left(v_g\right)$ and so equals a constant. 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 these derivations that

$$E\left[y_i - \widehat{y}_{g,-ii'} y_{i'} a^2 d - (a + 2x_i abd)\, \widehat{y}_{g,-ii'} - \left(x_i b + x_i^2 b^2 d\right) - v_0 \mid x_i, \mathbf{r}_g\right] = 0 \tag{22}$$

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

$$E\left[\left(y_i - \widehat{y}_{g,-ii'} y_{i'} a^2 d - (a + 2x_i abd)\, \widehat{y}_{g,-ii'} - \left(x_i b + x_i^2 b^2 d\right) - v_0\right) \mathbf{r}_{gi}\right] = 0 \tag{23}$$

which we can estimate using GMM exactly as before, treating as observations every pair of individuals in every group and time period, and using group level clustered standard errors.

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. Appendix A.4 provides the formal proof of identification and associated GMM estimation for the random effects generic model as discussed above (and for the extension to multiple equations), and Appendix A.6 proves that this identification and estimation extends to our full utility derived demand model with random effects.

IV Empirical Results

IV.A Preliminary Analysis: Well-being, Consumption, and Luxuries

Our model makes several assumptions about how peer consumption affects welfare. First, we assume that the effects of peer expenditures on utility have observable implications in the corresponding demand functions (via Roy's identity). This could be violated if, e.g., utility were additively separable in $\overline{\mathbf{q}}_g$ and $\mathbf{q}$. Second, we assume that an increase in peer expenditures causes a decrease in utility. Before proceeding with our main structural results, we implement two preliminary data analyses to examine these key assumptions, and also to check other modeling assumptions (e.g., that $\mathbf{q}_i$ is quadratic in x_i). Details of the data construction and empirical results of these preliminary data analyses are given in Appendix B. Here we just briefly summarize our main findings from these empirical analyses.

Our first preliminary analysis uses all observations from the 61$^{\text{st}}$ round (conducted in 2004 and 2005) of the National Sample Survey (NSS) of India, a nationally representative survey on, among other things, household spending patterns (more description of these data are below). We estimate the generic model fixed-effects model given by equation (17), letting y_i equal expenditures on luxuries and letting x_i equal total household expenditure. Groups are defined by education level and district (analogous to counties in the USA). The main results of this analysis, in Appendix Table B2, confirm that precise coefficient estimates can be obtained using the generic model, that peer-average luxury expenditures significantly affect demands for luxury demands, that both linear and quadratic terms in the budget x_i are statistically significant, and that a linear index structure in peer effects and the budget (as is implied by the structural model assumption that needs are linear in mean peer expenditures) appears to adequately capture the effects of both. The fixed effects results anticipate the main structural results; in column 8 of Appendix Table B2 we find that a one rupee increase in peer expenditure makes individuals behave as if they were 0.59 rupees poorer. We discuss these preliminary results in full in Appendix Section B.1.

Our second preliminary analysis addresses head-on the question of whether higher peer consumption reduces own utility. As we discuss in Section B.2, the structural finding that higher peer expenditures makes consumers behave as if they were poorer could alternatively be consistent with some positive network effects (e.g., from cellphones) so this is key for correctly interpreting the welfare consequences of our model. We use a completely different data set from India, the 5th (2006) and 6th (2014) waves of the World Values Surveys (WVS). The WVS asks respondents about their subjective well-being with the question "All things considered, how satisfied are you with your life as a whole these days?" and codes the response on a five-point scale. It also includes information on household income bins. Putting these together, we directly test whether this crude, self-reported measure of utility is decreased by higher peer expenditure.

We regress self-reported well-being (both linearly and by ordered logit) on income based approximations of the budget and peer expenditures in Appendix Table B3. We find that the resulting coefficient estimates have signs that are consistent with our theory (own expenditures increase utility while peer expenditures decrease utility). We would expect a marginal rate of substitution between the two to have a value between zero and one. Our point estimate of 1.45 (see column 2 of Table B3) is outside this range, but with a standard error of 0.85, so we cannot reject any value between zero and one. Luttmer (2005) performs a similar exercise with an American data set, with the same results regarding coefficient signs. He finds a marginal rate of substitution of 0.76, meaning that an increase in peer expenditures of 100 dollars has the same effect on utility as a decrease in one's consumption budget of 76 dollars. Finally, we include an interaction term (the product of the budget and peer expenditures) in the regression in columns 3 and 6, and find its coefficient to be insignificantly different from zero, which is consistent with our linear index modeling assumption.

IV.B Data

For our main empirical analyses, we use household consumption data from the 59th through 62nd rounds of the NSS, which were conducted between 2003 and 2007. The NSS are large annual surveys, with roughly 30,000 to 100,000 observations of household-level data in each round. They collect data on household demographics and household spending patterns. The latter data are used to compute, among other things, the consumer price indices that are commonly used in India.

In our baseline empirical work, we include only non-urban Hindu households to minimize within-group heterogeneity.[9] We further exclude scheduled caste/scheduled tribe (SC/ST)

[9] Urban households are typically immigrants from many different areas and varying sub-castes, and so even Hindus of similar caste living near to each other in the same city may not be similar enough to treat as a peer group.

households (even if Hindu).[10] We consider only households that are between the 1st and 99th percentiles of household expenditure in each state/year. We use only non-urban households whose state-identifier is not masked and with 12 or fewer members whose head is aged 20 or more.

Our peer groups are 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), allowing each group to be observed up to four times (once in each of the four NSS rounds). 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 dropped with this restriction.[11] Even with this restriction we are still left with relatively few observations per group. The average number of members observed in each peer group is 24 households, and the median is 15. These small group sizes illustrate the importance of showing identification and consistent estimation without assuming that the number of observed members per group goes to infinity, and without assuming that most or all of the members of each group are observed.

For our main sample of non-urban Hindu non-SC/ST households, we have 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. Our resulting dataset has 56,516 distinct households. Our estimators use all unique household-pairs within each period-group, and we have a total of 2,055,776 such pairs.

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. Note that these are poor households, so typical luxuries here are goods and services like sweets, ghee, processed foods, transportation, shampoo, and toothpaste.

We let t index price regimes. Our observed prices vary by 4 time periods and by 33 states, so t ranges from 1 to $T = 4 \times 33 = 132$. Each individual household i is only observed once, in one price regime and belonging to one group. We construct prices of our demand

[10] Below, we report additional results for samples of non-urban non-Hindu households and samples of non-urban SC/ST households, which we find have some significant behavioral differences from our main sample.

[11] Our theorems show that identification is possible with as few as three observed members per group, but when very few group members are actually observed, estimates of group means become extremely noisy, resulting in a substantial decrease in estimation precision.

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.

We condition on 7 demographic variables $\mathbf{z}$. 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; 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 1 shows summary statistics for the our NSS sample. We provide summary statistics at the level of the household, and at the level of the household-pairs used for estimation. Table 1 also reports summary statistics for prices and quantities of visible and invisible subcomponents of luxuries and necessities, using the categorization of Roth (2014) to classify goods as visible versus invisible. 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.

Total expenditures, and its components of luxury and necessity spending, are expressed in units of average household expenditure in round 59, so the average total expenditure of 1.12 reported in Table 1 shows that household spending was 12% higher in our overall sample than in the first round of the data. Roughly one-quarter of household spending is classified as luxury spending (0.31/1.12). Scanning Table 1 reveals that in our non-urban Hindu sample, only 6 per cent of households have high school education (high education) and almost all households have married household heads. Roughly one-quarter of households have ration cards entitling them to subsidized basic foods.

IV.C Baseline Structural Model Estimates

We estimate all models by GMM. For the 2-good system (luxuries and necessities), we use the model or equation (12) and employ the associated moment conditions (20) and (23) for the fixed- and random-effects specifications, respectively. Both models use pairwise data formed of all unique pairs of observations within each group, and are clustered at the group-year level to obtain valid inference.

Our fixed-effects approach involves substituting the leave-two-out within-group sample average quantity $\widehat{q}_{gjt,-ii'}$ for the within-group mean $\overline{q}_{gjt}$, and differencing across people within groups. Thus, we substitute $\widehat{q}_{gjt,-ii'}$ for $\overline{q}_{gjt}$ in the definition of X_i (eq. (11)) to create $\widehat{X}_i$:

$$\widehat{X}_i = x_i - R_{11}p_{1t} - R_{22}p_{2t} - \left(A_{11}\widehat{q}_{g1t,-ii'} + \mathbf{C}_1'\mathbf{z}_i\right)p_{1t} - \left(A_{22}\widehat{q}_{g2t,-ii'} + \mathbf{C}_2'\mathbf{z}_i\right)p_{2t},$$

and substitute $\widehat{q}_{gjt,-ii'}$ and $\widehat{X}_i$ into the demand equation (12). Then, we difference the demand equation across individuals within groups to generate a moment condition analogous to (20):

$$E[(p_{1ti}q_{1ti} - p_{1ti'}q_{1ti'} - (\widehat{X}_i^2 - \widehat{X}_{i'}^2)e^{-(b_1 \ln p_{1t} + (1-b_1)\ln p_{2t})}d_1 - (\widehat{X}_i - \widehat{X}_{i'})b_1 + \mathbf{C}_1' p_{1t}\left(\mathbf{z}_i - \mathbf{z}_{i'}\right))\mathbf{r}_{gtii'}] = 0.$$

Notice that, as in the generic model, many group-varying terms, including $A_{11}p_{1t}\overline{q}_{g1t}$, drop out as a result of this differencing. Further, since $\left(\widehat{X}_i - \widehat{X}_{i'}\right) = x_i - x_{i'} - \mathbf{C}_1'\left(\mathbf{z}_i - \mathbf{z}_{i'}\right)p_{1t} - \mathbf{C}_2'\left(\mathbf{z}_i - \mathbf{z}_{i'}\right)p_{2t}$, these terms are present only in the quadratic term $\left(\widehat{X}_i^2 - \widehat{X}_{i'}^2\right)$ via interactions between group-average quantities $\overline{q}_{g1t}$ and other elements of $\widehat{X}_i$ (e.g., x_i). The formal derivation of these moments for GMM estimation is given in Appendix A.5.

Our random-effects approach, derived in Appendix A.6, involves substituting the within-group sample average quantity and another group member's quantity for the within-group means. We use the above definition of $\widehat{X}_i$ for the linear term in the demand equation (12) and compute a new variable $\widetilde{X}_{ii'}$ for the squared term as follows:

$$\widetilde{X}_{ii'} = \widehat{X}_i[x_i - R_{11}p_{1t} - R_{22}p_{2t} - \left(A_{11}q_{g1ti'} + \mathbf{C}_1'\mathbf{z}_i\right)p_{1t} - \left(A_{22}q_{g2ti'} + \mathbf{C}_2'\mathbf{z}_i\right)p_{2t}].$$

Finally, we substitute $\widehat{q}_{g1t,-ii'}$, $\widehat{X}_i$ and $\widetilde{X}_{ii'}$ into the demand equation (12) to generate a moment condition analogous to (23):

$$E[(p_{1t}q_{1t} - \widetilde{X}_{ii'}e^{-(b_1 \ln p_{1t} + (1-b_1)\ln p_{2t})}d_1 - \widehat{X}_i b_1 - R_{11}p_{1t} - A_{11}p_{1t}\widehat{q}_{g1t,-ii'} - \mathbf{C}_1'p_{1t}\mathbf{z}_i - p_{1t}v_0)\mathbf{r}_{gti}] = 0.$$

These moments use pair-specific instruments which differ between our fixed- and random effects models. To instrument for $\widehat{q}_{gjt}$, we use group-averages from other time periods. Let the subscript $-t$ indicate averages from all other time periods. For both fixed- and random-effects instruments, we create a group-level instrument $\check{q}_{gjt}$ equal to the OLS predicted value of $\widehat{q}_{gjt}$ conditional on $\widehat{x}_{g,-t}, \widehat{x^2}_{g,-t}, \sqrt{\widehat{x}_{g,-t}}, \widehat{x}^2_{g,-t}, \widehat{\mathbf{z}}_{g,-t}$.[12] Additionally, let $\widetilde{\mathbf{z}}_{it}$ and $\widetilde{\mathbf{z}}_{gt}$ be the

[12] This is similar to including these values as instruments for $\mathbf{q_{gt}}$, but reduces the dimensionality of the instrument vector. This dimensionality reduction is quite significant because $\check{\mathbf{q}}_{gt}$ is multiplied by the demographic controls to generate the final instrument vector.

individually-varying and group-level, respectively, subvectors of $\mathbf{z}_i$. In our baseline model, $\widetilde{\mathbf{z}}_{it}$ includes 5 household-level variables, and $\widetilde{\mathbf{z}}_{gt}$ includes just the remaining 2 variables: dummy variables for primary- and secondary-school education levels. Letting $\cdot$ denote element-wise multiplication, our instrument list for the fixed-effects model is:

$$\mathbf{r}_{gtii'} = \left(x_{it}^2 - x_{i't}^2\right), (x_{it} - x_{i't})\cdot(1, \mathbf{p}_t \cdot \check{\mathbf{q}}_{gt}, \mathbf{p}_t \cdot \widetilde{\mathbf{z}}_{gt}), \mathbf{p}_t\cdot(\widetilde{\mathbf{z}}_{it} - \widetilde{\mathbf{z}}_{i't})\cdot(1, \mathbf{p}_t \cdot \check{\mathbf{q}}_{gt}), x_{it}\mathbf{p}_t\cdot(\widetilde{\mathbf{z}}_{it} - \widetilde{\mathbf{z}}_{i't}).$$

Our instrument list for the random-effects model is:

$$\mathbf{r}_{gti} = (1, \mathbf{p}_t, \mathbf{p}_t \cdot \check{\mathbf{q}}_{gt}, \mathbf{p}_t \cdot \mathbf{z}_{it}) \ , \ x_{it} \cdot (1, \mathbf{p}_t, x_{it}, \mathbf{p}_t \cdot \check{\mathbf{q}}_{gt}, \mathbf{p}_t \cdot \mathbf{z}_{it}) \ , \ \mathbf{p}_t \cdot \mathbf{p}_t.$$

The last term provides instruments for v_0.

Our primary focus is estimation of peer effects given by elements of the matrix $\mathbf{A}$, but first we consider the general reasonableness of our coefficients in the context of demand system estimation. Complete estimates for our baseline models (corresponding to Tables 2 and 3) are given in Appendix Tables B4 and B5. Our estimated quantity demand for luxuries has positive curvature. All four baseline specifications have d_1, the coefficient on the squared budget term, being statistically significantly positive and of large magnitude, as expected for luxuries. Regarding demographic covariates, it is reasonable to expect that needs would rise with household size. In all four baseline specifications we find that this is indeed the case, specifically, the parameters $C_{j,hhsize}$, are statistically significant and positive. Additionally, their magnitudes are reasonable: an additional household member increases the needs for luxuries by roughly 0.06 and the needs for necessities by roughly 0.15, where the units are normalized to equal 1 for the average income in 2009.

Table 2 gives estimates of the spillover (peer effects) parameter matrix $\mathbf{A}$ using the fixed effects estimator 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 2, labeled "A diagonal" gives estimates for the case where $\mathbf{A}$ is a general diagonal matrix. Later we consider cross-effects, allowing $\mathbf{A}$ to have non-zero off diagonal elements.

In the "A same" case, spending on needs is given by $\mathbf{p}'\boldsymbol{F}_i = a\overline{x}_{gt} + \mathbf{p}'\mathbf{C}\mathbf{z}_i$, so a gives the response of spending on needs to a change in the average total expenditure in the group. This baseline model is a revealed preference derived demand function analog to regressing measured utility on total peer expenditures, as in Luttmer (2005) and our preliminary analysis that used WVS data. This is also the most parsimonious version of our structural model. 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 estimate roughly comparable to Luttmer's (2005) estimate of 0.76 using stated well-being data.

In the bottom of Table 2 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 2), we obtain estimates that lie far outside the plausible range of $[0, 1]$. These estimates are also very imprecise, with standard errors that are roughly triple those in the left panel. The explanation for this imprecision and corresponding wildness of the estimates in this $\mathbf{A}$ varying case is multicollinearity. More preciesely, in the FE model, varying parameters in $\mathbf{A}$ are identified from the $(x_i - x_{i'})\overline{\mathbf{q}}_{gt}$ interaction terms (recall here and below that the actual $\overline{\mathbf{q}}_{gt}$ elements are unobserved and are replaced by estimates $\widehat{\mathbf{q}}_{gt,-ii'}$). In our data, the elements of our estimate of $\overline{\mathbf{q}}_{gt}$ are highly correlated with each other across groups and time, with a correlation coefficient of 0.85, resulting in a large degree of multicollinearity.[13] The result is that the estimated first element of $\mathbf{A}$ is implausibly low, offset by the second diagonal element of $\mathbf{A}$ that is implausibly high by a similar magnitude. We take this as evidence that the data are only rich enough to support the fixed effects implementation of the "A same" specification.

This problem of multicollinearity is considerably reduced in the random effects 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, with both terms helping to identify $\mathbf{A}$ in the RE model. At the bottom of Table 2, 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" baseline specification, but are rejected in the more general "A diagonal" specification.

The estimates of $\mathbf{A}$ in the RE model are reported in Table 3. 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 $\mathbf{A}$ in the RE model are equal.

The interpretation of these separate coefficients is slightly more complicated than in

[13]A possible reason for this multicollinearity is that, for this population, the differences between luxuries and necessities is not large (see the examples of luxuries given earlier). Perhaps in a wealthier country these differences would be larger. This may also explain the relatively small empirical difference between visible and invisible goods that we report later.

the "A same" case. To compare these estimates to the scalar a, suppose group average expenditures $\overline{x}_{gt}$ increased by 100 rupees. Then group average luxury quantities, $\overline{q}_{1gt}$ would increase by about $30/p_1$ (since, in Table 1, luxuries are about 30% of total spending), and so spending on needs in the luxury category ($p_1 A_{11} \overline{q}_{1gt}$) would increase by about 14 rupees ($p_1 \times 0.46 \times 30/p_1$). Similarly, spending on needs in the necessities category would increase by about 40 rupees ($p_2 \times 0.57 \times 70/p_2$). This yields a total increase in spending on perceived needs of 54 rupees, which is very close to the estimate one gets with a scalar a (50 rupees in the FE model or 54 rupees in the RE model).

In the rightmost panel of Table 3, 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 highly correlated 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 independent variation in group quantities and within group total expenditures to overcome these multicollinearity issues and obtain reliable estimates with an unrestricted $\mathbf{A}$ matrix.

IV.D Small Group Sizes

One of our main model innovations is obtaining consistent estimates of peer effects using standard survey data. This required dealing with the measurement error in group means that results from only observing a relatively small number of the members of each peer group. Both our FE and RE estimators account for the fact that most group members are unobserved, and that the number of group members sampled in each group is small and fixed (does not tend to infinity). This results in observed group mean-expenditures in each group that are endogenous, and have correlated measurement errors. Our estimators dealt with this problem by a combination of pairing observations, using a leave-two-out estimator of the group means, and appropriate construction of instruments.

Table 4 considers whether or not these measurement error corrections matter empirically. To see why accounting for these errors might be important, it is helpful to return to the simple generic model moment equations for the FE and RE models given in equations (20) and (23), respectively. Consider first the moment conditions for the random-effects model given by equation (23). If we didn't correct for measurement error, we would instrument for $\widehat{y}_{g,-ii'}$ with contemporaneous group-level averages of regressors (rather than group-level averages from other time periods), as is common in linear social-interactions models. But, this instrument would be polluted with correlated measurement errors leading to bias in the estimated parameters. This moment equation is linear in the variables (though nonlinear in the parameters), so we can think of measurement error in $\widehat{y}_{g,-ii'}$ through two channels: 1) standard attenuation bias on the reduced form effects of $\widehat{y}_{g,-ii'}$ and its interactions; and 2) bias induced by the fact that the measurement error gets squared and interacted with other variables. For the RE model, the first order effect comes through the attenuation bias on a multiplying $\widehat{y}_{g,-ii'}$. This should shrink the estimated a towards zero when we don't correct for the small group size measurement error.

The impacts of squared and interacted measurement errors are more complex, running through parameters multiplying the interaction terms $\widehat{y}_{g,-ii'}y_{i'}$ and $\widehat{y}_{g,-ii'}x_i$. Attenuation of these coefficients would multiply products of a and other parameters, so the direction of bias induced on the estimated a is uncertain. Similarly, the bias induced on a from the fact that measurement error itself gets squared and interacted with other variables has an uncertain effect.

Turning to the moment condition for the fixed-effects model given by equation (20), we have similar biases from the interaction terms $\widehat{y}_{g,-ii'}(x_i - x_{i'})$, but no first order attenuation bias because the $\widehat{y}_{g,-ii'}$ term is differenced out in the fixed effects moment equation. The biases induced from squared and interacted measurement errors are also present in the fixed-effects moment equation, but are again of uncertain direction. Thus, we expect a first-order attenuation bias plus smaller unsigned bias in the random effects model, and an unsigned bias of uncertain magnitude in the fixed effects model.

Columns 1 and 3 of Table 4 give estimates of the FE and RE models do not correct for these measurement errors. These are estimates that would be consistent if the within-group observed sample sizes went to infinity. The model here is the "Same A" specification that instruments for $\widehat{q}_{gjt,-ii'}$ using $\check{q}_{gjt}$ estimated conditional on contemporaneous $\widehat{x}_{g,t}, \widehat{x^2}_{g,t}, \sqrt{\widehat{x}_{g,t}}, \widehat{x}^2_{g,t}, \widehat{\mathbf{z}}_{g,t}$ (rather than their $-t$ analogs). Here, we see that the estimated value of a for the FE model is equal to 0.79 which is somewhat larger than the measurement-corrected estimate given in column 4. This suggests that the combined effects of attenuation bias in the interaction terms involving $\widehat{q}_{gjt,-ii'}$ and the biases from the squared and interacted measurement errors, are to bias a away from zero. In contrast, the estimate of a in the RE model is 0.17. This estimate additionally includes first-order attenuation bias from the level term $\widehat{q}_{g1t,-ii'}$, suggesting that this first-order attenuation bias dominates the estimated coefficient, and drives it close to zero, as expected.

To summarize these results, we found earlier that measurement-error corrected fixed and random effects models both yield estimated values of a near 0.5 for our baseline "Same A"specification. This is no longer the case when we fail to account for the measurement errors induced by only observing a small number of members of each peer group. Correcting for these errors is empirically important. This is especially true for the random-effects estimator, where failure to correct results in an estimate of a that is severely attenuated.

IV.E Alternative Structural Model Estimates

In Table 5, we turn to the question of whether consumption externalities vary depending on whether or not goods are visible or invisible to one's peers, according to the characterisation of Roth (2014). The idea is that peer effects may be larger for visible goods, both because they are more conspicuous, and because of potential Veblen (1899) effects. By this theory 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. Dividing both luxuries and necessities into visible and invisible components yields a demand system with $J = 4$ goods (of which we estimate 3 equations, with the fourth equation coefficients being obtained by the adding up constraint).

The first and second columns of Table 5 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 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 rightmost column of Table 5 give the RE estimates of the "diagonal A" model. As before, we find that luxuries have somewhat smaller externalities than necessities. However, the estimated element of $\mathbf{A}$ for visible luxuries is smaller than that for invisible luxuries, while the estimated value for visible necessities is larger than for invisible necessities. So the Veblen or conspicuous consumption story for visible goods is supported for necessities, but not for luxuries. Since necessities make up about 70% of total spending, this implies that the overall peer effect is larger for visible than invisible consumption.

In Table 6, 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 three different subsamples of the non-urban population: Hindu non-SC/ST households, SC/ST households, and non-Hindu SC/ST households. There are roughly one-quarter as many Hindu SC/ST households as Hindu households, and roughly one-fifth as many non-Hindu households as Hindu households, so separate regressions are feasible. 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 closer to zero, though with larger standard errors. This means that peer effects may vary by caste and religion. Interestingly, the peer effects are largest among the largest and most dominant social group, suggesting that some density of the peer group may be necessary to detect peer effects.

The right hand panel of Table 6 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. Further, only 6 per cent of households have high education, so separate estimation for this group is not tenable due to sample size limitations. For efficiency 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 peer effect for the illiterate/barely literate groups. 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. We take this to mean that the very poorest households in India are close enough to subsistence that it is more costly to engage in status competitions.

The results in Table 6 are striking in that they show that poorer demographics, SC/ST and illiterate/barely literate, have much smaller peer effects than others. This finding is similar to Akay and Martinsson's (2011) finding for very poor Ethiopians. In Table 7, we further investigate this possibility by splitting the baseline (Hindu non-SC/ST) sample into households whose real income is below the district-year median real income and households whose real income is above the district-year median real income. The implicit assumption of this specification is that the poorer and richer halves of each education group within each district correspond to different peer groups. We present fixed-effects estimates for the model with a scalar a, and random-effects estimates for the model with scalar a and diagonal A, since these were the most precisely estimated models in our baseline specifications.

The fixed-effects estimates of a for poorer and richer households are 0.26 and 0.59, respectively. This difference is marginally statistically significant (z-stat of 1.86) but large in magnitude, implying peer effects that are almost twice as large among the rich groups as among the poor groups. We find a slightly larger difference in the random-effects estimates of a, which are 0.32 and 0.78 for poor and rich households, respectively. The random-effects estimates pass a Hausman test against the fixed-effects alternative for both poor and rich households. Finally, turning to the random-effects estimates with diagonal A matrices, we again find estimated spillovers that are much smaller for poor than for rich households. Interestingly, for poor households we find consumption externalities that are a bit larger on luxuries vs necessities, which is the opposite of what we found in other specifications, where necessities spillovers were a little larger.

IV.F Summary of Empirical Results

We draw the following lessons our empirical work. First, we find that overall, peer effects are of similar magnitudes for luxuries and necessities, suggesting that the matrix **A** can be reasonably approximated by a scalar a times the identity matrix (the "A same" specifications). This implies that the consumption externalities component of needs is close to equaling the scalar a times group-average total expenditures. This was not *a priori* obvious, but means that a parsimonious model relating peer group expenditure to utility captures the most important aspects of peer consumption effects.[14]

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

Third, the measurement error corrections we propose to account for only observing a small number of the members of each peer group are empirically important. These corrections are what allow us to analyze peer effects with standard survey data, instead of data that includes most or all members of each peer group. Both fixed and random effects estimators show bias without this correction, and in particular, our otherwise more efficient random effects estimates show severe attenuation bias when we do not account for this measurement error in our estimation method.

Fourth, our baseline estimates of the scalar a are at or a little above 0.5. However, alternative model specifications, and nonstructural estimates based on reported life satisfaction, suggest potentially higher peer effects of up to around 0.7. We also find evidence that particular subgroups, especially poorer subgroups, have peer effects lower than 0.5.

V Implications for Tax and Transfers Policy

Our peer effects finding that needs rise with group-average consumption (with a coefficient of 0.5 or more in most groups) has significant implications for policies regarding redistribution, transfer systems, public goods provision, and economic growth. Like consumption rat race models and "keeping up with the Jones" models, our model is one where consumption has

[14] This result may be due to the relative similarity of luxuries vs necessities in this population, as discussed earlier. It seems likely that in a wealthier population, variation in peer effects across different types of goods would be larger and more significant.

negative externalities, in our case, increasing perceived needs and thereby reducing the utilities of peers. Boskin and Shoshenski (1978) consider optimal redistribution in models with general consumption externalities. They show that distortions due to negative externalities from consumption onto utility can generally be corrected by optimal taxation. In particular, their results imply that negative consumption externalities make the marginal cost of public funds lower than it would otherwise be, so the optimal amount of redistribution is greater than it would otherwise be. Here we apply the Boskin and Shoshenski (1978) logic to our estimated consumption peer effects, and in particular show how potentially large free lunch gains are possible.

As discussed in Section II, the sum (over households) of income less the sum of spending on needs as we define them is a valid money-metric social welfare index. This means that if needs go down, ceteris paribus, social welfare goes up. Consider the money metric costs in lost utility of, say, an across-the-board tax increase. This tax increase lowers average expenditures by households, which in turn lowers perceived needs, thereby offsetting some of the utility that was lost by having to pay the tax. For example, suppose you experience a two rupee tax increase. If your peers also have their taxes increase by the same amount, then (with $a = 0.5$) your loss in utility will only be equivalent to that of a one rupee tax increase.

However, we must also consider the potential peer effects in how the government uses the additional tax revenue. If the money is transferred to other groups of consumers who also have peer effect spillovers of $a = 0.5$, then the welfare gains from reduced expenditures on needs by the taxed consumers will be exactly offset by the welfare losses associated with increased perceived needs by the recipients of those transfers.

There are two ways we can reduce or eliminate these offsetting welfare losses, thereby exploiting the potential free lunch associated with the reduced perceived needs from taxing peers. One way is to transfer the funds to groups that have smaller peer effect spillovers, and the other is to spend the additional tax revenue on public goods.

We found that the size of the peer effect spillovers may be smaller for poorer and the least educated groups than for other consumers. If so, then transfers from richer groups to poorer ones will lead to an overall increase in social welfare, by reducing the total negative consumption externalities of the peer effects. This is true even with an inequality-neutral social welfare function. Similarly, our estimates suggest social welfare gains to progressive vs flat taxes, even if, ceteris paribus, the marginal utility of money were the same for all consumers.

An alternative way to exploit the potential free lunch associated with the reduced perceived needs from taxing peers is to spend the resulting tax revenues on public goods. To the extent that jealousy or envy are the underlying cause of the peer externalities we identify, public goods would not invoke those effects (or should at least induce smaller effects), because by definition public goods are consumed by all members of the group. This along with the Boskin and Shoshenski theorems suggests that public goods may be a partially free lunch.

To illustrate the magnitude of these potential welfare gains, we consider just one existing transfer program in India. This is the Public Distribution System (PDS), as currently funded by the National Food Security Act of 2013. This program is estimated to cost roughly 1.35% of GDP when fully implemented (Puri 2017; Ministry of Consumer Affairs 2018). The PDS aims to provide subsidized cereals to roughly 75 per cent of Indian households at roughly 1/3 of market price, and so, in our framework aims to increase the consumption of necessities. Our estimates imply that the resulting increased consumption would result in increased perceived needs, and so would not raise utility as much as an alternative policy that did not induce these negative externalities. Such alternatives could be provision of public goods, i.e., policies that provide resources to the poor but are equally available to all households. Such public goods might include clean water, public sanitation, better air quality, or better schools.

A rough back-of-the-envelope calculation of the magnitude of these potential gains proceeds as follows. The entitlement of rice under the PDS is up to 5 kg (kilograms) per month per person at 3 rupees per kg. Suppose the market price of rice is 15 rupees per kg (as it was in 2016). Thus, the public cost of providing this rice subsidy is about 12 rupees per kg, or 60 rupees per month per person. We can bound each consumer's behavioral response to the subsidy by noting that necessities consumption could rise by as much as 60 rupees per month per person, or at the other extreme, the consumer could choose to keep their rice consumption unchanged and spend the 60 rupees per month on luxuries. The actual response would likely be somewhere in between.[15]

For simplicity in constructing bounds, suppose that within each peer group either everyone or nobody qualifies for (or takes up) the PDS entitlement. At one extreme, suppose every consumer who gets the entitlement increases their necessities spending by 60 rupees per person per month. Then, taking our baseline random-effects estimate of the spillover from necessities of 0.57, we would have that the needs of every group member rises by 34 (0.57 times 60) rupees, resulting in an increase of only 60 - 34 = 26 rupees per consumer per month in their money-metric utility. At the other extreme, if all consumers who get the

[15] We could use our model to estimate what portion would be spent on luxuries vs necessities based on the distribution of prices and budgets in the population, but that complication turns out to not be necessary for our rough calculations. Note that a portion of the gain could also be saved, but that just implies spending it on luxuries or necessities at some future date.

entitlement use all of the extra resources provided to buy luxuries, then the corresponding spillover estimate is 0.46, which by a comparable calculation results in a 28 rupees increase in needs and therefore a 60 - 28 = 32 rupees gain in utility per consumer. Thus the government's expenditures of 60 rupees only increases money metric utility by 26 to 32 rupees per person per month. This is in contrast to a full benefit of 60 rupees per person per month that might be obtained by provision of public goods.[16] The PDS program targets roughly 1 billion people, yielding potential money-metric welfare gains (of switching from rice subsidies to a public goods program) of roughly 336 billion to 408 billion rupees per year.

Note that this calculation used our estimates of 0.57 and 0.42 for the peer effects of necessities and luxuries. Since it is poorer households that receive the PDS ration cards, it may be that the more appropriate estimate of peer effects to use is 0.26 or 0.32, the estimates we obtained for just poorer households (albeit with larger standard errors). In that case the benefits of switching to public goods we calculated above may be halved, but that still corresponds to money metric savings of over 160 billion rupees per year. Also, if the difference in peer effects between rich and poor is that large, then the NFSA program itself is much less expensive in money metric terms than it appears. This is because, as noted above, the money metric utility loss due to peer effects among the program's recipients would be smaller than the corresponding gains among the richer groups who pay most of the taxes that fund the program.

VI Conclusion

We show identification and GMM estimation of peer effects in a generic quadratic model, using ordinary survey (not panel) data where most members of each group might not be observed. The model allows for peer group level fixed or random effects, and allows the number of observed individuals in each peer group to be fixed asymptotically. This means 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 cross section survey data where the vast majority of members of each peer group are not observed, and detailed network structure is not provided.

We provide a utility derived consumer demand model, where one's perceived needs for

[16] An important caveat is that the benefits of this alternative might be reduced to the extent that some households derive less utility from the public good than others, but may also be increased to the extent that people in groups that did not qualify for or take up the rice entitlement might also benefit from the public good. The relative benefits could also be reduced or increased if peer expenditures have positive or negative externalities that we are not measuring. Examples could include positive network effects from increased cell phone ownership, or negative congestion effects from increased use of public roads.

each commodity depend 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 total expenditure increases are spent on "keeping up with the Joneses" type peer effects. This demand model, in which peer expenditures affect perceived 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 roughly 50 rupees or more in most peer groups. 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 models (that ignore these peer effects) would imply.

These results also suggest that income or consumption taxes have far lower 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. We show this is particularly true to the extent that taxes are used to provide public goods (that are less likely to induce peer effects) rather than transfers.

We provide some calculations showing that the magnitudes of these peer effects on social welfare calculations, which are ignored by standard models of government tax and spending policies, can be very large. For example, we find potential free lunch welfare gains of hundreds of billions of rupees may be available in just a single existing government transfer program in India. We find similarly that the welfare gains in transfers from richer to poorer households (and more generally from progressive vs flat taxes) may be much larger than previously thought, if those poorer households do indeed have smaller peer effect spillovers.

VII References

Angrist, Joshua D, (2014), “The perils of peer effects,” Labour Economics, 30, 98-108.

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(4), 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) 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.

Blundell, Richard, Xiaohong Chen, and Dennis Kristensen, (2007), “Semi-Nonparametric IV Estimation of Shape? Invariant Engel Curves,” Econometrica, 75(6), 1613-1669.

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

Boskin, M.J. and Sheshinski, E., (1978), “Optimal redistributive taxation when individual welfare depends upon relative income,” The Quarterly Journal of Economics, 92(4), 589-601.

Bramoullé, Yann, Habiba Djebbari, and Bernard Fortin, (2009), “Identification of peer effects through social networks.” Journal of Econometrics, 150(1), 41-55.

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,” 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.

Davezies, Laurent, Xavier d’Haultfoeuille, and Denis Fougère, (2009), “Identification of peer effects using group size variation.” The Econometrics Journal, 12(3), 397-413.

De Giorgi, Giacomo, Anders Frederiksen, and Luigi Pistaferri, (2016), “Consumption network effects.” No. w22357. National Bureau of Economic Research.

Deaton, Angus, and John Muellbauer, (1980), “An almost ideal demand system,” American Economic Review, 70(3), 312-326.

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(3), 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,” American Economic Review, 93(1), 423-428.

Easterlin, Richard A., (1974), “Does economic growth improve the human lot? Some empirical evidence,"Nations and Households in Economic Growth, 89, 89-125.

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(1), 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.

Gorman, W.M., (1981), “Comments on James Quirk’s Article,” Computer-Assisted Analysis and Model Simplification, 175-189.

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

Hausman, Jerry A., Hausman, J. A., Newey, W. K., Ichimura, H., and Powell, J. L., (1991), “Identification and estimation of polynomial errors-in-variables models,” Journal of Econometrics, 50(3), 273-295.

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.

Jochmans, Koen, and Martin Weidner., (2016), “Fixed-effect regressions on network data.” arXiv preprint 1608.01532.

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

Kahneman, Daniel, and Amos Tversky, (1979), “Prospect theory: An analysis of decision under risk,” Econometrica, 47(2), 263-291.

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 Reviewk 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.

Manski, C. F., (1993), “Identification of Endogenous Social Effects: The Reflection Problem,” Review of Economic Studies, 60(3), 531-542.

Manski, C. F., (2000), “Economic analysis of social interactions,” Journal of Economic Perspectives, 14(3), 115-136.

Maurer, J., and A. Meier, A., (2008), "Smooth it Like the ‘Joneses’? Estimating Peer-Group Effects in Intertemporal Consumption Choice,” Economic Journal, 118(527), 454-476.

Ministry of Consumer Affairs, Department of Food and Public Distribution, Government of India, (2018), “Salient Features of the National Food Security Act”, http://dfpd.nic.in/Salient-features-National-Food-Security-Act.htm.

Moffitt, Robert A., (2001), “Policy interventions, low-level equilibria, and social interactions,” Social dynamics, 4, 6-17.

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 Applied Econometrics, 20(4), 487–507.

Pollak, R.A. and Wales, T.J., (1978), “Estimation of complete demand systems from household budget data: the linear and quadratic expenditure systems,” American Economic Review, 68(3), 348-359.

Pollak, R.A. and Wales, T.J.,. (1981), “Demographic variables in demand analysis,” Econometrica, 49(6), 1533-1551.

Puri, Raghav., (2017), “India’s National Food Security Act (NFSA): Early Experiences”, LANSA, https://assets.publishing.service.gov.uk/media/5964831e40f0b60a44000154/NFSA-LWP.pdf.

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,” Review of Economic Studies, 15(2), 88-90.

Sojourner, Aaron J., (2009), "Inference on peer effects with missing peer data: Evidence from project STAR," Economic Journal, 123(569), 574–605.

Stevenson, Betsey, and Justin Wolfers, (2008), "Economic growth and subjective well-being: Reassessing the Easterlin paradox," No. w14282. National Bureau of Economic Research.

Stone, R., (1954), "Linear Expenditure System and Demand Analysis: An Application to the Pattern of British Demand," Economic Journal, 64(255), 511-527.

Tamer, Elie, (2003), "Incomplete simultaneous discrete response model with multiple equilibria," Review of Economic Studies, 70(1), 147-165.

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

Wikipedia, National Food Security Act, https://en.wikipedia.org/wiki/National_Food_Security_Act,_20

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.

VIII Tables

Table 1: Summary statistics for NSS consumption data

	Observations (N=56,516)				Pairs (N=2,055,776)			
	Mean	SD	Min	Max	Mean	SD	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

Summary statistics for estimation sample. Includes all 2354 group-rounds with 10 or more obs of Hindu non-SC/ST households. Groups defined as the cross of education (less than primary, primary, secondary or more) and district.

Table 2: Structural demand model, fixed effects estimates

	A Same	A Diagonal
A (own luxuries)	0.50	-2.63
	(0.11)	(0.40)
A (own necessities)	0.50	2.99
	(0.11)	(0.28)
χ^2 A same	80	
P-value	*[0.00]*	
Hausman test (A luxuries)	-0.31	-7.8
P-value	*[0.76]*	*[0.00]*
Hausman test (A necessities)		8.8
P-value		*[0.00]*

Selected estimates for structural demand model. Table displays effect of group consumption on needs. χ^2 A same tests whether the diagonal A coefficients in the second column are the same. Hausman tests are for the FE coefficient against the RE coefficient.

Table 3: Structural demand model, random effects effects estimates

	A Same	A Diagonal	A Full
A (own luxuries)	0.55	0.46	0.20
	(0.02)	(0.02)	(0.09)
A (own necessities)	0.55	0.57	1.09
	(0.02)	(0.02)	(0.10)
A (cross luxuries)			0.42
			(0.08)
A (cross necessities)			-0.33
			(0.11)
χ^2 A same	43		
P-val	*[0.00]*		

Selected estimates for structural demand model. Table displays effect of group consumption on needs. χ^2 statistic tests the first-column model that constrains the diagonal elements of A to be the same for necessitities and luxuries against the diagonal A in the second column.

Table 4: Estimated peer effects by measurement error correction

	RE		FE	
	(1)	(2)	(3)	(4)
	Naive	Baseline	Naive	Baseline
A (own consumption)	0.17	0.55	0.79	0.50
	(0.081)	(0.016)	(0.14)	(0.11)
Observations	2,055,776	2,055,776	2,055,776	2,055,776

Selected estimates for structural demand model, with and without correction for measurement error in group averages.

Table 5: Structural demand model, four consumption categories

	Fixed effects	Random effects	
	A same	A same	A diagonal
A (visible luxuries)	0.71	0.65	0.54
	(0.05)	(0.01)	(0.01)
A (invisible luxuries)	0.71	0.65	0.62
	(0.05)	(0.01)	(0.01)
A (visible necessities)	0.71	0.65	0.761
	(0.05)	(0.01)	(0.01)
A (invisible necessities)	0.71	0.65	0.66
	(0.05)	(0.01)	(0.01)
Hausman test RE	1.26		
	[0.21]		
χ^2 A same			658
			[0.00]

Selected estimates for structural demand model. Table displays effect of group consumption on needs.

Table 6: Structural demand model, fixed effects estimates

	Religion	Education
A (Hindu, non-SC/ST)	0.50	
	(0.11)	
A (SC/ST)	0.13	
	(0.18)	
A (non-Hindu)	-0.06	
	(0.23)	
A (less than primary)		0.08
		(0.15)
A (primary)		0.56
		(0.12)
A (secondary)		0.37
		(0.22)

Selected estimates for structural demand model. Religion models are estimated separately by demographic subgroup. Table displays effect of group consumption on needs.

Table 7: Structural demand model, by above/below median expenditure

	Fixed effects	Random effects	
	A same	A same	A diagonal
Panel A: Below median expenditure			
A (luxuries)	0.26	0.32	0.42
	(0.05)	(0.01)	(0.01)
A (necessities)	0.26	0.32	0.37
	(0.05)	(0.01)	(0.02)
Panel B: Above median expenditure			
A (luxuries)	0.59	0.78	0.65
	(0.17)	(0.03)	(0.04)
A (necessities)	0.59	0.78	0.86
	(0.17)	(0.03)	(0.04)

Selected estimates for structural demand model. Religion models are estimated separately by demographic subgroup. Table displays effect of group consumption on needs.

Appendix A: Derivations

A.1 Peer Effects as a Game

The interactions of peer group members may be interpreted as a game. 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. Estimation of complete games typically depend on having data on all members of each observed group. An example is Lee (2007). However, in our case 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. In either type of game there is 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 model 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.[17]

A.2 Generic Model Identification and Estimation With Fixed Effects

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

[17] A more difficult problem would be allowing for the possibility that group members may, like the econometrician, only observe group means with error. We do not attempt to tackle this issue. Doing so would require modeling how individuals estimate group means, how they incorporate uncertainty regarding group mean estimates into their purchasing decisions, and showing how all of that could be identified in the presence of the many other obstacles to identification that we face.

$\varepsilon_{\mathbf{xx}g,-ii'}$ 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

$$y_i = \left(\overline{y}_g a + \mathbf{x}_i'\mathbf{b}\right)^2 d + \left(\overline{y}_g a + \mathbf{x}_i'\mathbf{b}\right) + v_g + u_i \tag{24}$$

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

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

where h and k are additional constants to be estimated. However, 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 (24), 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 $y_i = \left(\overline{y}_g a + \mathbf{x}_i'\mathbf{b}\right)^2 d + \widetilde{v}_g + u_i$. Since this pure quadratic form 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 (24).

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 A5 below.

Assumption A1: Each individual i in group g satisfies equation (24). $\mathbf{x}_i$ is a K-dimensional vector of covariates. For each $k \in \{1, ..., K\}$, for each group g with $i \in g$ and $i' \in g$, $\Pr\left(\mathbf{x}_{ik} \neq \mathbf{x}_{i'k}\right) > 0$. Unobserved v_g are group level fixed effects. Unobserved errors

u_i are independent across groups g and have $E(u_i \; |\text{all } \mathbf{x}_{i'} \text{ having } i' \in g \text{ where } i \in g) = 0$. 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 $d \neq 0$ is needed to identify the parameter a in the fixed effects identification, because if $d = 0$ making the model linear, then after differencing, the parameter a would drop out of the model. This nonlinearity will not be required later for random effects model. Having $\mathbf{b} \neq 0$ is necessary since otherwise we would have nothing exogenous in the model.

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, plugging equation (24) for y_i into $\overline{y}_g = E\left(y_i \mid i \in g\right)$, we have

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

Taking the within group expected value of this expression gives

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

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

$$\overline{y}_g = \frac{1 - a(2\mathbf{b}'\overline{\mathbf{x}}_g d + 1) \pm \sqrt{[1 - a(2\mathbf{b}'\overline{\mathbf{x}}_g d + 1)]^2 - 4a^2 d[d\mathbf{b}'\overline{\mathbf{x}\mathbf{x}'_g}\mathbf{b} + \mathbf{b}'\overline{\mathbf{x}}_g + v_g]}}{2a^2 d}. \tag{27}$$

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.If the inequality in Assumption A2 is satisfied the this yields a quadratic in $\overline{y}_g$, which, if $a \neq 0$, has real solutions and having a solution means that an equilibrium exists. If a does equal zero, then the model will trivially have an equilibrium (and be identified) because in that case there aren't any peer effects. We do not take a stand on which root of equation (27) is chosen by consumers, we just make the following assumption.

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

The equilibrium of $\overline{y}_g$ therefore exists under Assumption A2 and is unique under Assumption A3.

For identification, we need to remove the fixed effect from equation (24), which we do by subtracting off another individual in the same group. For each $(i, i') \in g$, consider pairwise difference

$$\begin{aligned} y_i - y_{i'} &= 2ad\overline{y}_g \mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'}) + d\mathbf{b}'(\mathbf{x}_i\mathbf{x}_i' - \mathbf{x}_{i'}\mathbf{x}_{i'}')\mathbf{b} + \mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'}) + u_i - u_{i'} \\ &= 2ad\widehat{y}_{g,-ii'}\mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'}) + d\mathbf{b}'(\mathbf{x}_i\mathbf{x}_i' - \mathbf{x}_{i'}\mathbf{x}_{i'}')\mathbf{b} + \mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'}) + u_i - u_{i'} - 2ad\varepsilon_{yg,-ii'}\mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'}), \end{aligned} \tag{28}$$

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 (28) 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 that

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

where

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

Substituting this expression into equation (28) 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 $E\left((\mathbf{x}_i - \overline{\mathbf{x}}_g) \mid i \in g, \overline{\mathbf{x}}_g, \overline{\mathbf{xx'}}_g, v_g, \mathbf{r}_g\right) = 0$, $E\left(\left(\mathbf{x}_i\mathbf{x}_i' - \overline{\mathbf{xx'}}_g\right) \mid i \in g, \mathbf{r}_g\right) = 0$, and that $\mathbf{x}_i - \overline{\mathbf{x}}_g$ and $\mathbf{x}_i\mathbf{x}_i' - \overline{\mathbf{xx'}}_g$ are independent across individuals i.

Assumption A4 corresponds to (but is a little stronger than) standard instrument validity assumptions. A sufficient condition for the equalities in Assumption A4 to hold is to let $\varepsilon_{ix} = \mathbf{x}_i - \overline{\mathbf{x}}_g$ be independent across individuals, and assume that $E(\varepsilon_{ix} \mid \overline{\mathbf{x}}_g, \overline{\mathbf{xx'}}_g, v_g, \mathbf{r}_g$ 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

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

A simpler but stronger sufficient condition would just be that ε_{ix} are independent across individuals i and independent of group level variables $\overline{\mathbf{x}}_g, \overline{\mathbf{xx'}}_g, v_g, \mathbf{r}_g$. 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

$$E\left[y_i - y_{i'} - (2ad\widehat{y}_{g,-ii'}\mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'}) + d\mathbf{b}'(\mathbf{x}_i\mathbf{x}_i' - \mathbf{x}_{i'}\mathbf{x}_{i'}')\mathbf{b} + \mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'})) \mid \mathbf{r}_g, \mathbf{x}_i, \mathbf{x}_{i'}\right] = 0. \tag{29}$$

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 (27) 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. To simplify estimation, we construct unconditional rather than conditional moments for identification and 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 (29) that

$$E\left[\left(y_i - y_{i'} - (1 + 2ad\widehat{y}_{g,-ii'})\sum_{k=1}^{K} b_k(x_{ki} - x_{ki'}) - d\sum_{k=1}^{K}\sum_{k'=1}^{K} b_k b_{k'}(x_{ki}x_{k'i} - x_{ki'}x_{k'i'})\right)\mathbf{r}_{gii'}\right] = 0. \tag{30}$$

Let

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

Equation (30) 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 (30) over all $(i, i') \in \Gamma_g$ gives the unconditional group level moment vector

$$E\left(\mathbf{Y}_{1g} - \sum_{k=1}^{K} b_k\mathbf{Y}_{2kg} - 2ad\sum_{k=1}^{K} b_k\mathbf{Y}_{3kg} - d\sum_{k=1}^{K}\sum_{k'=1}^{K} b_k b_{k'}\mathbf{Y}_{4kk'g}\right) = 0. \tag{31}$$

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-A5, the coefficients a, $\mathbf{b}$, d are identified from

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

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 (31).

Given the identification above, based on equation (31) we can immediately construct a corresponding group level GMM estimator

$$\left(\widehat{a}, \widehat{b}_1, ...\widehat{b}_K, \widehat{d}\right) = \arg\min \left[\frac{1}{G}\sum_{g=1}^{G}\left(\mathbf{Y}_{1g} - \sum_{k=1}^{K} b_k \mathbf{Y}_{2kg} - 2ad\sum_{k=1}^{K} b_k \mathbf{Y}_{3kg} - d\sum_{k=1}^{K}\sum_{k'=1}^{K} b_k b_{k'} \mathbf{Y}_{4kk'g}\right)\right]' \cdot \widehat{\Omega}\left[\frac{1}{G}\sum_{g=1}^{G}\left(\mathbf{Y}_{1g} - \sum_{k=1}^{K} b_k \mathbf{Y}_{2kg} - 2ad\sum_{k=1}^{K} b_k \mathbf{Y}_{3kg} - d\sum_{k=1}^{K}\sum_{k'=1}^{K} b_k b_{k'} \mathbf{Y}_{4kk'g}\right)\right] \quad (32)$$

for some positive definite moment weighting matrix $\widehat{\Omega}$. In equation (32), 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 (30), 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,

$$\left(\widehat{a}, \widehat{b}_1, ...\widehat{b}_K, \widehat{d}\right) = \arg\min \left(\frac{\sum_{g=1}^{G}\sum_{(i,i')\in\Gamma_g}\mathbf{m}_{gii'}}{\sum_{g=1}^{G}\sum_{(i,i')\in\Gamma_g}1}\right)' \widehat{\Omega}\left(\frac{\sum_{g=1}^{G}\sum_{(i,i')\in\Gamma_g}\mathbf{m}_{gii'}}{\sum_{g=1}^{G}\sum_{(i,i')\in\Gamma_g}1}\right), \quad (33)$$

where

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

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 (30), $\mathbf{r}_{gii'}$ 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 (27) 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., $x_{ki}(x_{ki} - x_{ki'})$.

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 empirical application, what corresponds to the vector $\mathbf{x}_i$ here includes the total expenditures, age, and other characteristics 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 (27) shows that additional nonlinear functions of $\widehat{\mathbf{x}}_{gt\cdot}$, could also be valid and potentially useful additional instruments.

A.3 Multiple Equation Generic Model With Fixed Effects

Our actual demand application has a vector of J outcomes and a corresponding system of J equations. Extending the generic model to a multiple equation system introduces potential cross equation peer effects, resulting in more parameters to identify and estimate. Let

$\mathbf{y}_i = (y_{1i}, ..., y_{Ji})$ be a J-dimensional outcome vector, where y_{ji} denotes the j'th outcome for individual i. Then we extend the single equation generic model to the multi equation that for each good j,

$$y_{ji} = (\overline{\mathbf{y}}_g'\mathbf{a}_j + \mathbf{x}_i'\mathbf{b}_j)^2 d_j + \left(\overline{\mathbf{y}}_g'\mathbf{a}_j + \mathbf{x}_i'\mathbf{b}_j\right) + v_{jg} + u_{ji}, \tag{34}$$

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 in Appendix A.2,

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

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

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

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

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

and the identification comes from the group level unconditional moment equation

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

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

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

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

Then, $\mathbf{b}_j$, 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

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

A.4 Multiple Equation Generic Model With Random Effects

Here we provide the derivation of equation (22), 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 (24).

$$y_i = \overline{y}_g^2 a^2 d + a\left(1 + 2\mathbf{b}'\mathbf{x}_i d\right)\overline{y}_g + \mathbf{b}'\mathbf{x}_i + \mathbf{b}'\mathbf{x}_i\mathbf{x}_i'\mathbf{b}d + v_g + u_i, \tag{35}$$

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

$$\overline{y}_g = \overline{y}_g^2 da^2 + a(2d\mathbf{b}'\overline{\mathbf{x}}_g + 1)\overline{y}_g + d\mathbf{b}'\overline{\mathbf{x}\mathbf{x}'_g}\mathbf{b} + \mathbf{b}'\overline{\mathbf{x}}_g + \mu, \tag{36}$$

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

$$\varepsilon_{yi'} = 2ad\overline{y}_g\mathbf{b}'\varepsilon_{\mathbf{x}i'} + d\mathbf{b}'\varepsilon_{\mathbf{xx}i'}\mathbf{b} + \mathbf{b}'\varepsilon_{\mathbf{x}i'} + u_{i'} + v_g - \mu$$

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

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

Therefore, we can write

$$y_i = \widehat{y}_{g,-ii'} y_{i'} a^2 d + a\left(1 + 2\mathbf{b}'\mathbf{x}_i d\right)\widehat{y}_{g,-ii'} + \mathbf{b}'\mathbf{x}_i + \mathbf{b}'\mathbf{x}_i\mathbf{x}_i'\mathbf{b}d + v_g + u_i + \widetilde{\varepsilon}_{gii'}, \tag{37}$$

where

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

Formally, we make the following assumptions.

Assumption A6: 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 A7: 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 A8: 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 Assumptions A6-A8 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

$$E(v_g + u_i + \widetilde{\varepsilon}_{gii'} \mid \overline{\mathbf{x}}_g, \overline{\mathbf{x}\mathbf{x}'_g}, \mathbf{x}_i) = \mu - da^2\sigma^2 = v_0. \tag{38}$$

By construction $v_g + u_i + \widetilde{\varepsilon}_{gii'}$ is also independent of $\mathbf{r}_g$. Given this, equation (22) then follows from equations (37) and (38).

A.5 Identification and Estimation 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 (9) 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 (9) 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 $\alpha = \mathbf{A}'\mathbf{p}$, $\beta = \mathbf{p}^{1/2\prime}\mathbf{R}\mathbf{p}^{1/2}$, $\widetilde{\gamma} = \widetilde{\mathbf{C}}'\mathbf{p}$, $\kappa = \mathbf{D}'\mathbf{p}$, $\delta = \mathbf{b}/\mathbf{p}$, $\mathbf{C}\mathbf{z}_i = \widetilde{\mathbf{C}}\widetilde{\mathbf{z}}_i + \mathbf{D}\widetilde{\mathbf{z}}_g$, $r_j = r_{jj} + 2\sum_{k>j} r_{jk} p_j^{-1/2} p_k^{1/2}$, 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 (9) reduces to the system of Engel curves

$$\begin{aligned}\mathbf{q}_i = &\left(x_i - \beta - \alpha'\overline{\mathbf{q}}_g - \widetilde{\gamma}'\widetilde{\mathbf{z}}_i - \kappa'\widetilde{\mathbf{z}}_g\right)^2 \mathbf{m} + \left(x_i - \beta - \alpha'\overline{\mathbf{q}}_g - \widetilde{\gamma}'\widetilde{\mathbf{z}}_i - \kappa'\widetilde{\mathbf{z}}_g\right)\delta \\ &+ \mathbf{r} + \mathbf{A}\overline{\mathbf{q}}_{\mathbf{g}} + \widetilde{\mathbf{C}}\widetilde{\mathbf{z}}_i + \mathbf{D}\widetilde{\mathbf{z}}_g + \mathbf{v}_g + \mathbf{u}_i,\end{aligned} \tag{39}$$

This has a very similar structure to the generic multiple equation system of equations (34), 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 (39) can be rewritten more simply as

$$\mathbf{q}_i = \left(x_i - \widetilde{\gamma}'\widetilde{\mathbf{z}}_i\right)^2 \mathbf{m} - 2\left(x_i - \widetilde{\gamma}'\widetilde{\mathbf{z}}_i\right)\left(\alpha'\overline{\mathbf{q}}_g + \beta + \kappa'\widetilde{\mathbf{z}}_g\right)\mathbf{m} + \left(x_i - \widetilde{\gamma}'\widetilde{\mathbf{z}}_i\right)\delta + \widetilde{\mathbf{C}}\widetilde{\mathbf{z}}_i + \widetilde{\mathbf{v}}_g + \mathbf{u}_i, \tag{40}$$

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 (40) 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

$$\begin{aligned} 0 = E\{[(q_{ji} - q_{ji'}) &- \left((x_i - \widetilde{\gamma}'\widetilde{\mathbf{z}}_i)^2 - (x_{i'} - \widetilde{\gamma}'\widetilde{\mathbf{z}}_{i'})^2\right) m_j - \widetilde{\mathbf{c}}_j'(\widetilde{\mathbf{z}}_i - \widetilde{\mathbf{z}}_{i'}) \\ &- (\delta_j - 2m_j(\alpha'\widehat{\mathbf{q}}_{g,-ii'} + \beta + \kappa'\widetilde{\mathbf{z}}_g))\left((x_i - \widetilde{\gamma}'\widetilde{\mathbf{z}}_i) - (x_{i'} - \widetilde{\gamma}'\widetilde{\mathbf{z}}_{i'})\right)]\mathbf{r}_{gii'}\}. \end{aligned} \tag{41}$$

Combining common terms, we have

$$\begin{aligned} 0 = E\{[(q_{ji} - q_{ji'}) &- (x_i^2 - x_{i'}^2)m_j + 2\left(x_i\widetilde{\mathbf{z}}_i - x_{i'}\widetilde{\mathbf{z}}_{i'}\right)'\widetilde{\gamma}m_j - \widetilde{\gamma}'(\widetilde{\mathbf{z}}_i\widetilde{\mathbf{z}}_i' - \widetilde{\mathbf{z}}_{i'}\widetilde{\mathbf{z}}_{i'}')\widetilde{\gamma}m_j \\ &- \left(\widetilde{\mathbf{c}}_j' - (\delta_j - 2m_j\beta)\widetilde{\gamma}'\right)(\widetilde{\mathbf{z}}_i - \widetilde{\mathbf{z}}_{i'}) - (\delta_j - 2m_j\beta)(x_i - x_{i'}) \\ &+ 2m_j\left(\alpha'\widehat{\mathbf{q}}_{g,-ii'} + \kappa'\widetilde{\mathbf{z}}_g\right)(x_i - x_{i'}) - 2\left(\widetilde{\mathbf{z}}_i - \widetilde{\mathbf{z}}_{i'}\right)'\widetilde{\gamma}m_j\left(\alpha'\widehat{\mathbf{q}}_{g,-ii'} + \kappa'\widetilde{\mathbf{z}}_g\right)]\mathbf{r}_{gii'}\}. \end{aligned} \tag{42}$$

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 $x_i(\widetilde{\mathbf{z}}_{i'} - \widetilde{\mathbf{z}}_i)$, $\delta_j - 2m_j\beta$ 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 $\widehat{\mathbf{q}}_{g,-ii'}(x_i - x_{i'})$ and $\widetilde{\mathbf{z}}_g(x_i - x_{i'})$. 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

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

from which we can construct

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

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

Hence,

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

Pairwise differencing gives

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

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 (39). 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 $\mathbf{b'1} = 1$, $\mathbf{d'1} = 0$, $\mathbf{d} \neq 0$. There exist values of $\overline{\mathbf{q}}_g$ that satisfy

$$\begin{aligned}\overline{\mathbf{q}}_g = \overline{x_g^2}\mathbf{m} + \left(\widetilde{\gamma}'\overline{\mathbf{zz}'}_g\widetilde{\gamma}\right)\mathbf{m} + \mathbf{m}\left(\alpha'\overline{\mathbf{q}}_g + \kappa'\widetilde{\mathbf{z}}_g + \beta\right)^2 - 2\mathbf{m}\left(\alpha'\overline{\mathbf{q}}_g + \kappa'\widetilde{\mathbf{z}}_g + \beta\right)\left(\overline{x}_g - \widetilde{\gamma}'\overline{\mathbf{z}}_g\right) \\ - 2\mathbf{m}\widetilde{\gamma}'\overline{x\mathbf{z}}_g + \left(\overline{x}_g - \beta - \alpha'\overline{\mathbf{q}}_g - \widetilde{\gamma}'\overline{\mathbf{z}}_g - \kappa'\widetilde{\mathbf{z}}_g\right)\delta + \mathbf{r} + \mathbf{A}\overline{\mathbf{q}}_\mathbf{g} + \widetilde{\mathbf{C}}\overline{\mathbf{z}}_g + \mathbf{D}\widetilde{\mathbf{z}}_g + \mathbf{v}_g. \qquad (43)\end{aligned}$$

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

$$\begin{aligned}\overline{q}_g = \frac{1}{2m\,(Ap)^2}\{(2mAp(\overline{x}_g - \widetilde{\gamma}'\overline{\mathbf{z}}_g - \kappa'\widetilde{\mathbf{z}}_g - \beta) + 1 - A + pA\delta) \pm [(2mAp(\overline{x}_g - \widetilde{\gamma}'\overline{\mathbf{z}}_g - \kappa'\widetilde{\mathbf{z}}_g - \beta) \\ + 1 - A + pA\delta)^2 - 4m\,(Ap)^2\left(m\overline{x_g^2} + m\gamma'\overline{\mathbf{zz}'}_g\gamma + m(\kappa'\widetilde{\mathbf{z}}_g + \beta)^2 - 2m(\kappa'\widetilde{\mathbf{z}}_g + \beta)(\overline{x}_g - \widetilde{\gamma}'\overline{\mathbf{z}}_g) \right. \\ \left. -2m\widetilde{\gamma}'\overline{x\mathbf{z}}_g + (\overline{x}_g - \beta - \widetilde{\gamma}'\overline{\mathbf{z}}_g - \kappa'\widetilde{\mathbf{z}}_g)\,\delta + r + \widetilde{C}\overline{\mathbf{z}}_g + D\widetilde{\mathbf{z}}_g + v_g\right)]^{1/2}\}, \qquad (44)\end{aligned}$$

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 (44), we can see $\overline{\mathbf{q}}_g$ is an implicit function of $\overline{x_g^2}, \overline{x}_g, \overline{\mathbf{z}}_g, \widetilde{\mathbf{z}}_g, \overline{\mathbf{zz}'}_g, \overline{x\mathbf{z}}_g$, and $\mathbf{v}_g$. In the case of multiple equilibria, we do not take a stand on which root of equation (43) 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 $E\left((\varepsilon_{ix}, \varepsilon_{iz})|\overline{\mathbf{z}}_g, \widetilde{\mathbf{z}}_g, \overline{x\mathbf{z}}_g, \overline{\mathbf{zz}'}_g, \overline{x}_g, \overline{x_g^2}, \mathbf{v}_g, \mathbf{r}_g\right) = 0$ 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

$$E\{[\mathbf{q}_i - \mathbf{q}_{i'} + 2\mathbf{m}\,(\alpha'\widehat{\mathbf{q}}_{g,-ii'} + \kappa'\widetilde{\mathbf{z}}_g)\,[(x_i - x_{i'}) - \widetilde{\gamma}'(\widetilde{\mathbf{z}}_i - \widetilde{\mathbf{z}}_{i'})] - (x_i^2 - x_{i'}^2)\mathbf{m} - \widetilde{\gamma}'\,(\widetilde{\mathbf{z}}_i\widetilde{\mathbf{z}}_i' - \widetilde{\mathbf{z}}_{i'}\widetilde{\mathbf{z}}_{i'}')\,\widetilde{\gamma}\mathbf{m}$$
$$+2\mathbf{m}\widetilde{\gamma}'(\widetilde{\mathbf{z}}_i x_i - \widetilde{\mathbf{z}}_{i'} x_{i'}) - \eta(x_i - x_{i'}) + (\eta\widetilde{\gamma}' - \widetilde{\mathbf{C}})(\widetilde{\mathbf{z}}_i - \widetilde{\mathbf{z}}_{i'})]|x_i, x_{i'}, \mathbf{z}_i, \mathbf{z}_{i'}, \mathbf{r}_g\} = 0 \quad (45)$$

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

$$E\Bigg[\Bigg(\mathbf{q}_i - \mathbf{q}_{i'} + 2e^{-\mathbf{b}'\ln\mathbf{p}_t}\frac{\mathbf{d}}{\mathbf{p}_t}\,(\mathbf{p}_t'\mathbf{A}\widehat{\mathbf{q}}_{gt,-ii'} + \mathbf{p}_t'\mathbf{D}\widetilde{\mathbf{z}}_g)\,[(x_i - x_{i'}) - \mathbf{p}_t'\widetilde{\mathbf{C}}(\widetilde{\mathbf{z}}_i - \widetilde{\mathbf{z}}_{i'})] - e^{-\mathbf{b}'\ln\mathbf{p}_t}\frac{\mathbf{d}}{\mathbf{p}_t}$$
$$[(x_i^2 - x_{i'}^2) + \mathbf{p}_t'\widetilde{\mathbf{C}}\,(\widetilde{\mathbf{z}}_i\widetilde{\mathbf{z}}_i' - \widetilde{\mathbf{z}}_{i'}\widetilde{\mathbf{z}}_{i'}')\,\widetilde{\mathbf{C}}'\mathbf{p}_t - 2\mathbf{p}_t'\widetilde{\mathbf{C}}(\mathbf{z}_i x_i - \mathbf{z}_{i'} x_{i'})] - \left(\frac{\mathbf{b}}{\mathbf{p}_t} - 2e^{-\mathbf{b}'\ln\mathbf{p}_t}\frac{\mathbf{d}}{\mathbf{p}_t}\mathbf{p}_t^{1/2\prime}\mathbf{R}\mathbf{p}_t^{1/2}\right)$$
$$\cdot(x_i - x_{i'}) + [(\frac{\mathbf{b}}{\mathbf{p}_t} - 2e^{-\mathbf{b}'\ln\mathbf{p}_t}\frac{\mathbf{d}}{\mathbf{p}_t}\mathbf{p}_t^{1/2\prime}\mathbf{R}\mathbf{p}_t^{1/2})\widetilde{\mathbf{C}}'\mathbf{p}_t - \widetilde{\mathbf{C}}](\widetilde{\mathbf{z}}_i - \widetilde{\mathbf{z}}_{i'})|x_i, x_{i'}, \mathbf{z}_i, \mathbf{z}_{i'}, \mathbf{r}_g\Bigg] = 0. \quad (46)$$

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

$$\begin{aligned} L_{1jgii'} &= (q_{ji} - q_{ji'}),\;\; L_{2jgii'} = \widehat{q}_{jg,-ii'}(x_i - x_{i'}),\;\; L_{3jkgii'} = \widehat{q}_{jgt,-ii'}(\widetilde{z}_{ki} - \widetilde{z}_{ki'}), \\ L_{4k_2gii'} &= \widetilde{z}_{k_2g}(x_i - x_{i'}),\;\; L_{5kk_2gii'} = \widetilde{z}_{k_2g}(\widetilde{z}_{ki} - \widetilde{z}_{ki'}),\;\; L_{6gii'} = x_i^2 - x_{i'}^2, \\ L_{7kk'gii'} &= \widetilde{z}_{ki}\widetilde{z}_{k'i} - \widetilde{z}_{ki'}\widetilde{z}_{k'i'},\;\; L_{8kgii'} = \widetilde{z}_{ki}x_i - \widetilde{z}_{ki'}x_{i'},\;\; L_{9gii'} = x_i - x_{i'},\;\; L_{10kgii'} = \widetilde{z}_{ki} - \widetilde{z}_{ki'}, \end{aligned} \quad (47)$$

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

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

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

$$\mathbf{P} = (\mathbf{p}_1, ..., \mathbf{p}_T)' \text{ and } \mathbf{\Lambda} = (\Lambda_1', ..., \Lambda_T')'$$

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

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

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

$$\mathbf{A} = (\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}'(\alpha_1, ..., \alpha_T)' \text{ and } \mathbf{D} = (\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}'(\kappa_1, ..., \kappa_T)';$$

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

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

$\mathbf{d}$ 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 (39) 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 (9) are identified.

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

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

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.

A.6 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,

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

By rewriting q_{ji} as

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

where

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

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

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

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

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

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 $i \in g$,

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

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 B7: $E\left(\mathbf{H}_g'\right) E\left(\mathbf{H}_g\right)$ 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 B7, 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

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

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

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

with

$$
\begin{aligned}
m_{jt}&=e^{-\mathbf{b}'\ln\mathbf{p}_t}\frac{d_j}{p_{jt}},\ \alpha_t=\mathbf{A}'\mathbf{p}_t,\ \widetilde{\gamma}_t=\widetilde{\mathbf{C}}'\mathbf{p}_t,\ \kappa_t=\mathbf{D}'\mathbf{p}_t,\ \beta_t=\mathbf{p}_t^{1/2\prime}\mathbf{R}\mathbf{p}_t^{1/2},\\
\eta_{jt}&=\frac{b_j}{p_{jt}}-2m_{jt}\mathbf{p}_t^{1/2\prime}\mathbf{R}\mathbf{p}_t^{1/2},\ \delta_{jt}=\frac{b_j}{p_{jt}},\ r_{jt}=R_{jj}+2\sum_{k>j}R_{jk}\sqrt{p_{kt}/p_{jt}},\\
v_{jt0}&=\mu_{jt}-e^{-\mathbf{b}'\ln\mathbf{p}_t}\frac{d_j}{p_{jt}}\sum_{j_1=1}^{J}\sum_{j_2=1}^{J}\sum_{j=1}^{J}\sum_{j'=1}^{J}A_{j_1j}p_{j_1t}A_{j_2j'}p_{j_2t}\mathbf{\Sigma}_{vt,jj'}.
\end{aligned}
$$

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

$$v_{jt0} = \mu_j - e^{-\mathbf{b}' \ln \mathbf{p}_t} \frac{d_j}{p_{jt}} \sum_{j_1=1}^{J} \sum_{j_2=1}^{J} \sum_{j=1}^{J} \sum_{j'=1}^{J} A_{j_1 j} p_{j_1 t} A_{j_2 j'} p_{j_2 t} \mathbf{\Sigma}_{v,jj'}$$

Moreover, since $\mathbf{v}_{gt}$ represents deviations from the utility 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.

Appendix B: Preliminary Data Analyses

B.1 Generic Model Estimates

Our first empirical is to estimate the generic peer effects model presentedin Section III. We do so using the 61$^{\text{st}}$ round of the NSS.

As in the presentation in (14), 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 (23) and fixed-effects unconditional moments (20). Define $\overline{x}_{g,-t}$ to be the group-average expenditure in other time periods. Fixed-effects instruments $\mathbf{r}_{gii'}$ 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 $\mathbf{r}_{gi}$ are: $\overline{x}_{g,-t}, x_i, x_i\overline{x}_{g,-t}, x_i^2, z_i, 1$. These instruments are constructed to mirror the sources of identification in the FE and RE cases, respectively. Resulting GMM estimates of the parameters are given in Table B2.

In the fixed, but not random effects specifications, peer luxury expenditure has a significant and substantial effect on own luxury expenditure. Higher levels of peer luxury expenditure work in the opposite direction of higher levels of own expenditure, effectively making the household behave (in a demand sense) as if it was poorer when peer expenditures rise. However, the magnitude of the peer effects varies dramatically across RE and FE specifications (although they do not vary much with different controls). Equality of peer effects is decisively rejected by Hausman tests. This is a natural consequence of the group-level unobservable taste for an expenditure category v_g being correlated with expenditure in that category. In our preferred FE estimate of column 8, a 100 rupee increase in peer luxury expenditures makes households behave as if they are over 50 rupees poorer (in terms of luxury demand), controlling for group level characteristics.

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.

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 on luxuries, but it cannot tell us if these spillovers are bad in the sense of lowering one's utility when one's peers spend more (though 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. This can most easily be seen by noting that b is typically smaller (albeit insignificantly) than a in the FE specifications, meaning that group expenditure has a larger effect on behavior than x_i. We showed in Appendix Section A.2 that this is inconsistent with a peer-spending equilibrium, and is a natural consequence of excluding group-average non-luxury spending from the right hand side. Fourth, it is not possible to derive welfare or utility implications of the resulting estimates.

In order to address the first of these issues, we now turn to a brief analysis of well-being data from a different survey. Dealing with the remaining issues requires our full structural model, which we present in the main paper.

B.2 Subjective well-being and peer consumption

Our generic model estimates above are consistent with a theory 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. However, the generic model only reveals the effect of peer consumption on one's own consumption, not on one's utility. For example, it is possible that the success of my peers makes me happy rather than envious. Or peer consumption could increase the utility I obtain from my own consumption, e.g., my own telephone becomes more useful when my friends also have telephones. In short, our needs model implies that peer expenditures induce negative rather than positive consumption externalities.

To directly check the sign of these peer 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.

Appendix Table B1 summarizes 3,236 observations from the 5^{th} (2006) and 6^{th} (2014) waves of the World Values Survey, two recent waves with 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. Average expenditure is 2,200 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 3 presents estimates of well-being as a function of both own total expenditures and group total expenditures, specified as

$$U_i = \beta_1 \widehat{x}_{igt} + \beta_2 \overline{\overline{x}}_{gt} + Z_{igt}\alpha + \gamma_g + \phi_t + \varepsilon_{igt}, \tag{48}$$

where U_i is the z-normalized well-being indicator, $\widehat{x}_{igt}$ is imputed individual expenditures, $\overline{\overline{x}}_{gt}$ is imputed group expenditures, Z_{igt} is vector of individual level controls, γ_g is a group level fixed effect (groups are defined within states, so this effectively includes a state fixed effect as well), and ϕ_t is a year fixed effect. Identification of β_2 comes from group-level changes in expenditure between rounds, and corresponds to the change in self-reported utility as group income is rising versus falling, holding own income constant. We also repeat this analysis using an ordered logit specification.

Results in the second column of Appendix Table B3 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 $\overline{\overline{x}}_{gt}$ *decreases* satisfaction by 0.19 standard deviations. Other specifications in Table B3 give similar results. The signs of these effects are consistent with our model of peer expenditures as negative consumption externalities. They are also consistent with Luttmer's (2005) finding of "neighbours as negatives" with US data, where increases in group income holding individual income constant reduces individual's reported well-being.

The ratio of the peer-expenditure and own-expenditure effects, $-\beta_2/\beta_1 = 19/13 = 1.45$, says that one must increase one's own expenditures by 145 rupees to compensate for the loss of utility that results from a 100 rupees increase in group expenditure levels. This point estimate is unreasonably large, as we show in Appendix Section A.2 that equilibrium requires that this ratio be less than 1. However, the standard error of this estimate is 0.85, meaning that we cannot reject any value in the reasonable range of zero to one. The corresponding ratio estimate in Luttmer (2005) is 0.76.

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 columns 4 and 5 of Table B3). The results are qualitatively the same, suggesting that our results are not being determined by the normalizations implicit in z-scoring the satisfaction responses.

Finally, we include an interaction term (the product of the budget and peer expenditures) in the regression in columns 3 and 6, and find its coefficient to be insignificantly different from zero, which is consistent with our linear index modeling assumption.

This analysis support our structural model assumptions that utility is increasing in household expenditure and decreasing in group average expenditure. Moreover, we cannot reject the assumption that the marginal rate of substitution between the two lies between zero and one, consistent with our main structural model finding in the neighborhood of 1/2.

Table B1: 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 B2: Luxury spending as a function of group spending, generic model estimates

	RE				Peer group FE			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
a (peer mean expenditure)	0.002	-0.068	-0.107	-0.112	-0.053	-0.324***	-0.657***	-0.586***
	(0.038)	(0.117)	(0.114)	(0.107)	(0.047)	(0.124)	(0.157)	(0.132)
b (own expenditure)	0.187***	0.439***	0.436***	0.428***	0.205***	0.445***	0.446***	0.379***
	(0.013)	(0.011)	(0.011)	(0.011)	(0.013)	(0.011)	(0.011)	(0.025)
d (curvature)	2.263***	0.289***	0.295***	0.308***	1.847***	0.289***	0.302***	0.413***
	(0.420)	(0.032)	(0.034)	(0.036)	(0.314)	(0.029)	(0.030)	(0.071)
-a/b	-0.010	0.156	0.245	0.261	0.258	0.727	1.474	1.546
	(0.203)	(0.266)	(0.259)	(0.247)	(0.225)	(0.267)	(0.328)	(0.342)
P(a = -b)	0.000	0.001	0.003	0.002	0.001	0.299	0.157	0.110
Hausman for a					4.400	3.644	12.470	13.885
P-value					0.036	0.056	0.000	0.000
Individual controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Group controls	No	No	Yes	Yes	No	No	Yes	Yes
Price controls	No	No	No	Yes	No	No	No	Yes
Number of groups	2,354	2,354	2,354	2,354	2,354	2,354	2,354	2,354
Number of pairs	2,055,776	2,055,776	2,055,776	2,055,776	2,055,776	2,055,776	2,055,776	2,055,776

Model estimated is $y_i = d(\hat{y}_g a + x_i b + X\beta)^2 + (\hat{y}_g a + x_i c + X\beta)$. 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 B3: Satisfaction on household and peer income

	OLS (SDs)			Ordered logit		
	(1)	(2)	(3)	(4)	(5)	(6)
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***	0.141*		0.335***	0.359*
		(0.025)	(0.079)		(0.058)	(0.198)
Group expenditure, deflated		-0.190*	-0.182		-0.424*	-0.407
		(0.107)	(0.114)		(0.256)	(0.285)
Own X group expenditure			-0.003			-0.006
			(0.018)			(0.044)
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Ratio	1.47	1.45	1.29	1.13	1.27	1.13
	(0.764)	(0.850)	(1.249)	(0.684)	(0.803)	(1.202)
P(Own + group = 0)	0.528	0.588	0.799	0.848	0.734	0.908
Dependent mean	0.00	0.00	0.00	3.07	3.07	3.07
Dependent SD	1.00	1.00	1.00	1.22	1.22	1.22
Observations	3236	3236	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 B4: Structural demand model, full estimates for fixed effects model

		Same A		Diagonal A	
		est	*std err*	est	*std err*
A	luxuries	0.502	*0.110*	-2.628	*0.395*
	necessities	0.502	*0.110*	2.992	*0.276*
R	luxuries	8.228	*4.228*	6.936	*2.387*
	necessities	-1.899	*2.462*	-17.609	*3.418*
C luxuries	(hhsize-1)/10	0.607	*0.049*	0.317	*0.030*
	headage/120	0.013	*0.085*	0.054	*0.044*
	married	0.070	*0.030*	0.010	*0.016*
	ln(land+1)	0.021	*0.016*	-0.010	*0.012*
	ration card	0.047	*0.027*	-0.020	*0.013*
	Educ med	-0.604	*0.792*	0.655	*0.857*
	Educ high	-1.754	*1.062*	0.165	*1.592*
C necessities	(hhsize-1)/10	1.476	*0.053*	1.138	*0.037*
	headage/120	0.102	*0.095*	0.129	*0.051*
	married	0.093	*0.031*	0.030	*0.018*
	ln(land+1)	0.088	*0.017*	0.051	*0.013*
	ration card	0.030	*0.031*	-0.050	*0.015*
	Educ med	0.323	*0.773*	-0.858	*0.868*
	Educ high	1.211	*1.041*	-0.350	*1.607*
b	luxuries	1.466	*0.233*	-0.870	*0.154*
d	luxuries	0.073	*0.004*	0.070	*0.004*

Table B5: Structural demand model, full estimates for random effects model

		Same A		Diagonal A	
		est	*std err*	est	*std err*
A	luxuries	0.547	*0.015*	0.461	*0.019*
	necessities	0.547	*0.015*	0.572	*0.016*
R	luxuries	-0.101	*0.180*	-0.766	*0.409*
	necessities	-3.674	*0.348*	-0.197	*1.517*
C luxuries	(hhsize-1)/10	0.596	*0.058*	0.598	*0.059*
	headage/120	-0.058	*0.080*	-0.074	*0.080*
	married	0.005	*0.030*	0.008	*0.028*
	ln(land+1)	0.055	*0.016*	0.056	*0.016*
	ration card	-0.054	*0.021*	-0.053	*0.020*
	Educ med	-0.112	*0.027*	-0.100	*0.026*
	Educ high	-0.205	*0.042*	-0.208	*0.046*
C necessities	(hhsize-1)/10	1.505	*0.070*	1.480	*0.068*
	headage/120	0.034	*0.095*	0.024	*0.091*
	married	0.026	*0.035*	0.031	*0.031*
	ln(land+1)	0.114	*0.019*	0.113	*0.019*
	ration card	-0.095	*0.025*	-0.092	*0.023*
	Educ med	-0.127	*0.033*	-0.119	*0.031*
	Educ high	-0.210	*0.043*	-0.231	*0.044*
b	luxuries	-0.176	*0.036*	0.352	*0.325*
d	luxuries	0.091	*0.004*	0.085	*0.005*
v	luxuries	1.022	*0.554*	-2.898	*0.845*
	necessities	4.119	*1.406*	-2.165	*3.656*