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.

April 2020

Abstract

We construct a peer effects model of consumption where mean expenditures of consumers in one's peer group affect one's utility through perceived consumption needs. We prove model identification using standard household-level consumer expenditure survey microdata, even with group fixed effects, and when most peer group members are not observed. Using Indian expenditure data, we find that each additional rupee spent by one's peers increases the cost of one's perceived needs, reducing money-metric utility by at least 0.25 rupees. These costs may be larger for richer households, meaning transfers from rich to poor could improve even inequality-neutral social welfare. We show welfare gains of billions of dollars per year might be possible by replacing government transfers of prvate goods to households with the provision of public goods or services.

*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 2005; 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 existence of peer effects in consumption, there has been much less work evaluating their welfare impacts. In this paper we study the effect of changes in peer mean expenditures on utility, asking how much one's own expenditure would have to increase to compensate for a unit increase in peer group expenditures.

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

Most empirical studies of consumption peer effects instead directly model individual consumption as a function of mean peer group consumption and other covariates (Chao and Schor, 1998; 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.

To study the welfare effects of peer consumption, we propose a "keeping-up-with-the-Joneses" 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 demand decisions to prices and budgets, but additionally allows the average expenditures of one's peers to affect demand decisions and welfare.

In our model, one's perceived required expenditures, or "needs," depend on, among other things, the mean 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, we find that consumers lose utility from feeling poorer when their peers get richer, and so consumers feel they must spend and consume more when their peers consume more (Luttmer 2005, Ravina 2008, Clark and Senik 2010). In contrast to those papers, which use direct data on reported well-being, we apply revealed preference methods to ordinary consumer demand data to identify the money-metric costs of peer consumption.

Much progress has been made in overcoming the econometric issue of 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 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, in our data we typically observe only 5 or 6 members of each peer group. This gives rise to some unusual econometric obstacles to identification that must be overcome.

One such obstacle is that we cannot consistently estimate within peer group mean expenditures, because so few members of each peer group are observed. Another is allowing for unobserved heterogeneity in group behavior, in the form of group level fixed effects or random effects. A third is coping with nonlinearities associated with maximization of empirically plausible utility functions. We propose some novel identifying moments to obtain model identification despite these obstacles, and provide an associated GMM estimator. These innovations in the econometric identification of peer effects could prove useful in many other applications of social interactions models.

We estimate our model using consumption survey data from India. Our groups are defined at a very local geographic level; roughly a small neighborhood. Within neighborhoods, we also group people by religion and caste. This results in very small groups, typically with a few hundred members, with only 5 to 10 members of the group sampled.

Empirically, we find that dealing with peer group mean measurement error issue is particularly important. Failing to account for these errors leads to attentuation biases so large that the estimated peer effects are reduced by up to 90% in some specifications.

Our empirical estimates of peer effects suggest that an increase in spending by one's peers of four rupees has the same effect on one's utility as a decrease in one's own expenditures of about one rupee. We also find some evidence that our estimated peer effects may be smaller for lower socio-economic status groups.

These results have at least 3 implications for tax and redistribution policy. First, they suggest that consumption or income taxes may be less costly in terms of social welfare and utility than is implied by standard demand models, which ignore peer effects. Since consumption has a negative externality through the peer effects channel, reducing consumption via taxation does not reduce welfare as much it would in the absence of that externality.

Second, if the utility associated with public goods or government services are not subject to these peer effects (or engender smaller peer effects), then governments can increase welfare by substituting the provision of public goods for the provision of private goods. This effect can be very large: we perform a rough calculation which shows that replacing India's Public Distribution System food subsidy program with more generous provision of public goods and services, such as public sanitation or cleaner air and water, could increase money metric welfare by up to 180 billion rupees (2.5 billion US dollars) per year at no additional cost.

Third, the finding that poorer households may have smaller peer effects suggests that transfers from higher to lower status groups can 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 thus give a new reason for tax progressivity, even if all consumers have the same marginal utility of money, and even if social welfare is inequality-neutral.

Our strategy to identify peer effects in consumption to estimate the parameters of a structural model of utility and demand, tailored to deal with the complexity of our data environment. But, the core of the approach is simple. Consider a model where each consumer, indexed by i, is a member of a peer group, indexed by g. Let overbars indicate true within-group means, and hats indicate sample averages. Let $\mathbf{q}_i$ be the vector of (continuous) quantities of goods that consumer i consumes. 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. This is $U_i = U(\mathbf{q}_i - \mathbf{f}_i)$, 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. Gorman (1976) analyzed the general form $U_i = U(\mathbf{q}_i - \mathbf{f}_i)$ 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 extend Gorman (1976) by letting 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 $\mathbf{f}_i = \mathbf{f}\left(\mathbf{z}_i,\overline{\mathbf{q}}_g\right)$ for some 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 given by

$$U_i = U\left(\mathbf{q}_i - \mathbf{f}\left(\mathbf{z}_i,\overline{\mathbf{q}}_g\right)\right) \tag{1}$$

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

Given these demand and needs 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 need to spend (i.e., 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 is the fundamental question our analysis seeks to answer.

Consistent estimation of how the needs function $\mathbf{f}\left(\mathbf{z}_i,\overline{\mathbf{q}}_g\right)$ responds to changes in $\overline{\mathbf{q}}_g$ is frustrated by several econometric obstacles, the most important of which is measurement error in average peer spending $\overline{\mathbf{q}}_g$. Our model introduces suitable observed and unobserved heterogeneity across consumers and across groups in the function U. Then, via revealed preference, we derive demand equations to estimate of the general form

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

Here $\mathbf{h}$ is a vector valued function (quadratic in x_i 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 complicates some aspects of identification. But, it helps overcome other identification issues: importantly, it allows us to include group-level fixed effects in the model.

Second, we cannot base identification on detailed network information (because we don't observe it). Most social interactions models make use of network information for identification. 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 data on friends of friends provides instruments for identification (e.g., Bramoullé, Djebbari and Fortin 2009; Jochmans and Weidner 2016; or de Giorgi, Frederiksen, and Pistaferri 2016). We use data consumer expenditure data, of the type that many countries collect for constructing consumer price indices. Since such surveys do not contain social network data, we can only define peer groups based on demographic characteristics and geography, and cannot exploit any network structure to help identification.

Third, as is in typical survey data, we only observe a small number of the members in each peer group (in our application, we observe fewer than 6 members of each group on average). As a result, we cannot consistently estimate group means $\overline{\mathbf{q}}_g$. For each group 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 both endogenous and correlated with other components of the model. This measurement error is further exacerbated by potential nonlinearity of $\mathbf{h}$, resulting in errors that contain interaction terms like $\left(\widehat{\mathbf{q}}_g - \overline{\mathbf{q}}_g\right) x_i$.

The remainder of this paper proceeds as follows. In Section II we expand on the structural model of utility, demand and peer effects introduced above. Section III illustrates our general procedure for dealing with the above econometric issues using a simple quadratic model. This procedure should be of independent interest to others wishing to estimate peer effects using survey data. We provide a sequence of formal theorems in a separate appendix, proving that our identification method and proposed estimators work both for the simple quadratic generic model (with some extensions), and for our general utility derived demand model. Results using this structural model are presented in Section IV, with policy implications provided 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, appears 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 this negative correlation still stands. Similar results hold for much smaller reference groups, e.g., Luttmer (2005) finds that an increase of the average income in one's neighbors reduces self-reported well being.

The possible mechanisms for this correlation 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 $\overline{\mathbf{q}}_g$.[1,2,3] If we could observe utility and consumption quantities of individuals and groups, 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 $\overline{x}_g$. Separate from our main empirical application, we estimate a similar regression, using data from India and groups that are roughly comparable to those in our main empirical analysis. The results agree with Luttmer (2005) and support our main model's underlying assumption that increases in peer expenditures decrease rather than increase utility. Our main model does not depend on crude utility measures, but instead identifies 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 use empirical approaches 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 one would need to increase x_i 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 Utility-Derived Demand Model

Ignoring unobserved preference heterogeneity for now, we begin with utility given by equation (1) where $U_i = U(\mathbf{q}_i - \mathbf{f}_i)$ where $\mathbf{f}_i = \mathbf{f}\left(\mathbf{z}_i, \overline{\mathbf{q}}_g\right)$. 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 a Cobb-Douglas utility function U with needs equalling a constant vector $\mathbf{f}$.

One can equivalently represent preferences using an indirect utility function, defined as

[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.

[2]One could imagine utility positively correlated with $\overline{\mathbf{q}}_g$, for example, through happiness for the success of one's peers. But the empirical evidence, including our own results, suggest that the correlation is negative.

[3]Our groups are defined (in the main) by geography. This implies a substantial risk of misspecifying how consumers are assigned to peer groups. We mitigate this risk in part by constructing very small groups, since defining groups that are too small creates inefficiency but not bias. We also show how misspecification of groups will generally leads to downward bias in peer effects estimates, so our estimated effects are likely to be conservative.

the maximum utility attainable with a given budget x_i when facing prices $\mathbf{p}$. Gorman (1976) shows[4] that for any regular utility function in this form, there exists a corresponding indirect utility function V such that

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

Indirect utility functions of this form can be shown to have many desirable properties for welfare calculations.[5] 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. We show later that we can also uniquely identify how $\mathbf{f}$ depends on $\overline{\mathbf{q}}_g$.

Luttmer (2005) regresses a self-reported measure of happiness on x_i, $\mathbf{z}_i$ and $\overline{x}_g$ (the within-group average income). We can interpret his regression as a simplified and linearized version of equation (3), where self-reported happiness is assumed to proxy for U_i. Table 1 (column 3) in Luttmer (2005) gives endogeneity-corrected estimates of the coefficients of $\overline{x}_g$ and x_i of -0.296 and 0.361, respectively. The negative ratio of these is 0.82, meaning that a 100 dollar increase in group-average income has the same effect on reported happiness as an 82 dollar reduction in own-income. We estimate an object that has a comparable interpretation to this relative coefficient. But instead of assuming that U_i equals an observed happiness measure that can be compared across individuals and regressed on covariates, we let U_i be unobserved. We instead derive demand equations from utility maximization, and recover the implied peer effects on utility using revealed preference methods.

The demand functions that result from maximizing our utility function can be obtained by applying Roy's (1947) identity to the indirect utility function of equation (3). These 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 $\mathbf{h}(\mathbf{p}, x) = -\nabla_{\mathbf{p}} V(\mathbf{p}, x)/\nabla_x V(\mathbf{p}, x)$. See, e.g., Pollak and Wales (1981) and Pendakur (2005). We take the function $\mathbf{f}$ to be linear, so

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

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, so we extend that linearity to the additional variables $\overline{\mathbf{q}}_g$.

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

[4] His version did not include $\overline{\mathbf{q}}_g$.

[5] Blackorby and Donaldson (1994) show that indirect utility functions $U_i = V\left(\mathbf{p}, x_i - \mathbf{p}'\mathbf{f}_i\right)$ satisfy Absolute Equivalence Scale Exactness (AESE). For preferences that satisfy AESE, one can 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.

effects and $\mathbf{u}_i$ is a $J-$vector of individual specific error terms that are assumed to have zero means conditional on x_i, $\mathbf{z}_i$, and $\mathbf{p}$.

The terms $\mathbf{v}_g + \mathbf{u}_i$ can be interpreted either as departures from utility maximization by individuals, or as unobserved preference heterogeneity. In the latter intepretation, 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 $\mathbf{Cz}_i$ with $(\mathbf{Cz}_i + \mathbf{v}_g + \mathbf{u}_i)$ in (4), and treat error terms as unobserved preference heterogeneity. However, we do not take a stand on whether $\mathbf{v}_g + \mathbf{u}_i$ represents preference heterogeneity or departures from utility maximization.

In the fixed effects model, $\mathbf{v}_g$ can be correlated in unknown ways with regressors including $\mathbf{p}$, x and $\overline{\mathbf{q}}_g$. The random effects model imposes the additional restriction that $\mathbf{v}_g$ be independent of regressors. As a result, the random effects model will be much more efficient, but at the cost of imposing these possibly questionable independence restrictions.

The above derivations yield demand functions of the general form

$$\mathbf{q}_i = \mathbf{h}(\mathbf{p}, x_i - \mathbf{p}'\mathbf{A}\overline{\mathbf{q}}_g - \mathbf{p}'\mathbf{C}\mathbf{z}_i) + \mathbf{A}\overline{\mathbf{q}}_g + \mathbf{C}\mathbf{z}_i + \mathbf{v}_g + \mathbf{u}_i. \tag{5}$$

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

Based on a long empirical literature,[6] we assume

$$V(\mathbf{p}, x) = -(x - R(\mathbf{p}))^{-1} B(\mathbf{p}) - D(\mathbf{p}) \tag{6}$$

for some differentiable functions R, B and D. Applying Roy's identity to obtain the function $\mathbf{h}$ and equation (5) yields demand equations

$$\begin{aligned} \mathbf{q}_i = {} & \left(x_i - R(\mathbf{p}) - \mathbf{p}'(\mathbf{A}\overline{\mathbf{q}}_g + \mathbf{C}\mathbf{z}_i)\right)^2 \frac{\nabla D(\mathbf{p})}{B(\mathbf{p})} \\ & + \left(x_i - R(\mathbf{p}) - \mathbf{p}'(\mathbf{A}\overline{\mathbf{q}}_g + \mathbf{C}\mathbf{z}_i)\right) \frac{\nabla B(\mathbf{p})}{B(\mathbf{p})} + \nabla R(\mathbf{p}) + \mathbf{A}\overline{\mathbf{q}}_g + \mathbf{C}\mathbf{z}_i + \mathbf{v}_g + \mathbf{u}_i. \end{aligned} \tag{7}$$

Rationality (consistency with utility maximization) 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 $\mathbf{p}$. Standard

[6] Many studies of commodity demands have found that observed demand functions are close to polynomial. See, e.g. Lewbel (1991), Banks, Blundell, and Lewbel (1997), and references therein. Gorman (1981) shows that any polynomial demand system has 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 are those given by equation (6).

functions that satisfy these conditions 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, e.g., Lewbel (1997).[7]

For each good j, the resulting demand model is

$$q_{ji} = Q_j\left(\mathbf{p}, x_i, \overline{\mathbf{q}}_g, \mathbf{z}_i\right) + v_{jg} + u_{ji}, \tag{8}$$

where each Q_j function is given by

$$\begin{aligned} Q_j\left(\mathbf{p}, x_i, \overline{\mathbf{q}}_g, \mathbf{z}_i\right) &= \left(x_i - \mathbf{p}^{1/2\prime}\mathbf{R}\mathbf{p}^{1/2} - \mathbf{p}'\mathbf{A}\overline{\mathbf{q}}_g - \mathbf{p}'\mathbf{C}\mathbf{z}_i\right)^2 e^{-\mathbf{b}'\ln\mathbf{p}}\frac{d_j}{p_j} \\ + \left(x_i - \mathbf{p}^{1/2\prime}\mathbf{R}\mathbf{p}^{1/2} - \mathbf{p}'\mathbf{A}\overline{\mathbf{q}}_g - \mathbf{p}'\mathbf{C}\mathbf{z}_i\right)&\frac{b_j}{p_j} + R_{jj} + \sum_{k\neq j} R_{jk}\sqrt{p_k/p_j} + \mathbf{A}_j'\overline{\mathbf{q}}_g + \mathbf{C}_j'\mathbf{z}_i \end{aligned} \tag{9}$$

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.[8]

In our data, prices vary geographically by state, but are fixed within each group, so we can subscript prices by g[9]. More generally, our model would permit observing groups in multiple time periods, with prices varying by time instead of, or in addition to, varying geographically. In the Appendix we derive results at this added level of generality, including t subscripts for time and price regimes.

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 $q_{Ji} = \left(x_i - \sum_{j=1}^{J-1} p_j q_{ji}\right)/p_J$. While we report some results using $J = 3$ goods, most of our analyses will be based on $J = 2$, with the two goods being food and non-food. In this case $J - 1 = 1$ so we only need to estimate the demand equation for one good, which we choose to be food. Most of our analyses will also assume $\mathbf{A}$ is diagonal. With these simplifications, equation (9) reduces to the single equation

$$Q_1\left(\mathbf{p}_g, x_i, \overline{\mathbf{q}}_g, \mathbf{z}_i\right) = X_i^2 e^{-(b_1 \ln p_{1g} + (1-b_1)\ln p_{2g})} d_1/p_{1g} + X_i b_1/p_{1g} + R_{11} + A_{11}\overline{q}_{g1} + \mathbf{C}_1'\mathbf{z}_i,$$

[7]To avoid multicollinearity, in our application we restrict $\mathbf{R}$ to be diagonal. Since $J \leq 3$, our model remains Diewert-flexible in own and cross price effects.

[8]There is one straightforward extension to the demand model that we consider in some of our estimates, but do not include above to save on notation. We allow a few discrete group-level characteristics (such as religion dummies) to interact with $\overline{\mathbf{q}}_g$, thereby allowing $\mathbf{A}$ to vary with these characteristics. Identification 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 be applied separately for each value of these characteristics.

[9]More generally, our model would permit observing groups in multiple time periods, with prices varying by time instead of, or in addition to, varying geographically. In the Appendix we derive results at this added level of generality, including t subscripts for time and price regimes.

where

$$X_i = X(\mathbf{p}_g, x, \overline{\mathbf{q}}_g, \mathbf{z}_i) = x_i - R_{11}p_{1g} - R_{22}p_{2g} - \left(A_{11}\overline{q}_{g1} + \mathbf{C}_1'\mathbf{z}_i\right)p_{1g} - \left(A_{22}\overline{q}_{g2} + \mathbf{C}_2'\mathbf{z}_i\right)p_{2g}. \quad (10)$$

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 (8) and multiplying by p_{1g} yields our primary estimation model:

$$p_{1g}q_{1i} = X_i^2 e^{-(b_1 \ln p_{1g} + (1-b_1)\ln p_{2g})} d_1 + X_i b_1 + R_{11}p_{1g} + A_{11}p_{1g}\overline{q}_{g1} + \mathbf{C}_1' p_{1g}\mathbf{z}_i + p_{1g}v_{1g} + p_{1g}u_{1i} \quad (11)$$

The goal will be estimation of the set of parameters $\{\mathbf{A},\ \mathbf{C},\ \mathbf{R},\ \mathbf{d},\ \mathbf{b}\}$. In particular, $\mathbf{A}$ 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. Although we solve all 5 issues, most of our econometric novelty relates to how we deal with issue 5, measurement error in the group means.

To illustrate how we overcome these econometric issues, we first consider a very simple model that suffers from all these same problems. Below we show informally how we identify and estimate this simple generic model. In the Appendix we provide formal proofs of our identification method and associated estimator asymptotics, for both a multivariate extension of this generic model, and for our full consumer demand model given by equation (11).

Our model starts with cross section data, where each observed individual i is assumed to be in a peer group $g \in \{1, ...G\}$. The number of peer groups G is large, so we assume $G \to \infty$. In our data we will only observe a small number n_g of the individuals who are actually in each peer group g, so asymptotics assuming $n_g \to \infty$ (or assuming that n_g grows to the total number of people in each group) are inapplicable. We therefore assume n_g is fixed and does not grow with the sample size.

The 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, assume there's a single scalar covariate x_i that affects y_i (we extend the generic model to vectors of y_i and x_i in the appendix).

A typical peer specification with such data would be linear, e.g., $y_i = \overline{y}_g a + x_i b + u_i$, 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 the nonlinearity and heterogeneity issues associated with our demand model, consider the more general specification

$$y_i = \left(\overline{y}_g a + x_i b\right)^2 d + \left(\overline{y}_g a + x_i b\right) + v_g + u_i \tag{12}$$

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 (12) is in some way fundamental. Rather, it's just a simple nonlinear specification that nests the standard linear model as a special case, resembles our full demand model, and can be used to demonstrate all the issues (and solutions) associated with identification and estimation of our demand model.

Equation (12) differs from the linear model both by the squared index term and by including a group-level fixed or random effect v_g. As discussed earlier, in social interaction models, typical ways of obtaining identification with such effects is to exploit specialized data that includes observable network structures like the "intransitive triads" of Bramoullé, Djebbari, and Fortin (2009). We do not have such network information in our data. Alternatively, one might obtain identification using typical panel data methods, such as by differencing out fixed effects over time. However, we only have cross section data, not a panel.

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 (12) equal to

$$y_i = \left(\widehat{y}_g a + x_i b\right)^2 d + \left(\widehat{y}_g a + x_i b\right) + v_g + u_i + \varepsilon_{gi}, \tag{13}$$

where the difference between $\widehat{y}_g$ and $\overline{y}_g$ results in the additional error term ε_{gi}. By construction, ε_{gi} is given by

$$\varepsilon_{gi} = \left(\overline{y}_g^2 - \widehat{y}_g^2\right) a^2 d + 2\left(\overline{y}_g - \widehat{y}_g\right) x_i abd + \left(\overline{y}_g - \widehat{y}_g\right) a. \tag{14}$$

Inspection of equations (13) and (14) shows many of the obstacles to identifying and estimating the model parameters a, b, and d. First, with either fixed or random effects, v_g could be correlated with $\widehat{y}_g$. 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. 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 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 covariates and the unobservables v_g, u_i, 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.[10] 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 would cause 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 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 Level Fixed Effects

In the fixed effects model, we make no assumptions about how v_g may correlate with other covariates (incluing $\overline{y}_g$) 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 the nonlinearity of demand, and so we must assume 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 consumers i and i' observed in in the same group g (and, if we have time variation, in the same time period). To remove correlation issues, we define the leave-two-out group mean estimator

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

[10] 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 inclusion of group quantities (corresponding to $\overline{y}_g$ in the generic 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 our model to overcome this issue. See the Appendix for details.

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'. Replacing $\widehat{y}_g$ in equation (13) with $\widehat{y}_{g,-ii'}$, and differencing equation (13) between the individuals i and i' gives

$$y_i - y_{i'} = 2\widehat{y}_{g,-ii'}\left(x_i - x_{i'}\right)abd + \left(x_i^2 - x_{i'}^2\right)b^2d + \left(x_i - x_{i'}\right)b + u_i - u_{i'} + \varepsilon_{gi} - \varepsilon_{gi'}. \quad (15)$$

where

$$\varepsilon_{gi} - \varepsilon_{gi'} = 2\left(\overline{y}_g - \widehat{y}_{g,-ii'}\right)\left(x_i - x_{i'}\right)abd. \quad (16)$$

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

$$E\left(u_i - u_{i'} + \varepsilon_{gi} - \varepsilon_{gi'} \mid x_i, x_{i'}\right) = 0, \quad (17)$$

which we can then use to construct some moments for estimation of equation (15).

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 can be shown to be conditionally mean zero.

Equation (15) contains functions of $\widehat{y}_{g,-ii'}$, x_i, and $x_{i'}$ as regressors, and equation (17) shows that we can use functions of x_i and $x_{i'}$ as instruments. However, we still require an instrument for $\widehat{y}_{g,-ii'}$, because of it's correlation with $\varepsilon_{gi} - \varepsilon_{gi'}$. Since each y depends on x, an obvious candidate instrument for an average of y's in a group (that is, $\widehat{y}_{g,-ii'}$) would be an average of x's in the group, that is, some estimate $\widehat{x}_g$ of the mean group value $\overline{x}_g$. However, 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 asymptotically 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 information. One possible source is repeated cross section data, which are typically available in consumption surveys. Usually the same consumers are not sampled more than once (so no panel data is available), but we may have observations of other consumers in the same group from different time periods. It doesn't matter that these other consumers may or may not have the same fixed effects v_g or mean expenditures $\overline{y}_g$ as in our main sample. All we need is an 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 our instruments for the corresponding functions of $\widehat{y}_{g,-ii'}$ that we require.

Even if one's survey data is only available from a single cross section, other data sets could alternatively provide these group level instruments. For example, if x_i is a demographic variable, then instead of observing individuals from the same group in another time period, we could use census data to provide an estimate of $\widehat{x}_g$. 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. It is not even necessary that we observe the exact same groups in other time periods or surveys. All we need is some overlap between the group definition in our main data and in the data used to construct the instrument, and some correlation between the variable used to construct the instrument and x.

Let $\mathbf{r}_g$ denote a scalar or vector of the above described group level instruments. 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

$$E\left[\left(y_i - y_{i'} - 2\widehat{y}_{g,-ii'}\left(x_i - x_{i'}\right) abd - \left(x_i^2 - x_{i'}^2\right) b^2 d - \left(x_i - x_{i'}\right) b\right) \mathbf{r}_{gii'}\right] = 0. \tag{18}$$

Based on equation (18), 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, so our sample becomes all such pairs i and i'. The estimator is equivalent to linearly regressing each pair $y_i - y_{i'}$ 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 pairs of individuals within each group, so we must cluster standard errors at the group level to obtain proper inference.

Theorem 1 in Appendix A.2 describes these results formally, including extending this model to allow for vector $\mathbf{x}_i$, providing formal conditions for proving that an equilibrium exists, and showing 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 Appendix A.5 then gives a final extension of these results, showing identification, consistent estimation, and inference of our full utility-derived demand model, given by equations (8) and (9) for each good j.

III.B Generic Model With Group Level Random Effects

A drawback of the fixed effects estimator is that differencing across individuals, which was needed to remove the fixed effects, results in a substantial loss of information. In this section we add the additional assumptions that v_g is homoskedastic and independent of x_i, and develop a more efficient random effects estimator that does not entail differencing. This random effects estimator does not require nonlinearity for identification, and so is still consistent when $d = 0$.

To describe the random effects estimator it will be convenient to rewrite equation (12) as

$$y_i = \overline{y}_g^2 a^2 d + (a + 2x_i abd)\,\overline{y}_g + \left(x_i b + x_i^2 b^2 d\right) + v_g + u_i. \tag{19}$$

As before, we 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 $\overline{y}_g^2$.

To obtain valid moments for identification now, we employ a variant of the method we used before. Again let i' denote an individual other than i in group g, construct $\widehat{y}_{g,-ii'}$ as before, and again replace $\overline{y}_g$ with $\widehat{y}_{g,-ii'}$. 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 $\overline{y}_g^2$ 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 combine the resulting moments over all individuals i' in g, thereby once again exploiting all of the information in the group. With this replacement, equation (19) 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 $E(\widetilde{\varepsilon}_{gii'} | x_i, \mathbf{r}_g) = -da^2 Var\,(v_g)$ 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

$$E\left[y_i - \widehat{y}_{g,-ii'} y_{i'} a^2 d - (a + 2x_i abd)\, \widehat{y}_{g,-ii'} - \left(x_i b + x_i^2 b^2 d\right) - v_0 \mid x_i, \mathbf{r}_g\right] = 0 \tag{20}$$

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

$$E\left[\left(y_i - \widehat{y}_{g,-ii'} y_{i'} a^2 d - (a + 2x_i abd)\, \widehat{y}_{g,-ii'} - \left(x_i b + x_i^2 b^2 d\right) - v_0\right) \mathbf{r}_{gi}\right] = 0 \tag{21}$$

which we can estimate using GMM exactly as before, treating every pair of individuals in each group as observations and clustering standard errors at the group level.

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 Data

For our main empirical analysis, we use household consumption data from the 61$^{\text{st}}$ round of the National Sample Survey (NSS) of India, which was conducted from July 2004 to June 2005. This survey contains information on household demographics and spending for a representative sample of the country.

To define appropriate peer groups, we exploit a property of multi-stage sampling, which is a standard feature of the NSS and other consumption surveys. To cut down on surveying costs, consumers are sampled from small geographic areas like villages and neighborhoods. These areas are particularly small and relevant in urban areas, where they're constructed to be compact and bounded by well-defined, clear-cut natural boundaries whenever possible, and so generally correspond to recognizable neighborhoods (NSS, 2019). Households in the same neighborhood are likely to be similar to each other in observable and unobservable ways because of assortative geographic selection, and are likely to be in at least indirect contact.

This makes them appropriate candidates for defining our groups, and crucially are available as a byproduct of the sampling design in many consumption surveys.

We restrict our attention to urban households, where the geographic sampling areas are particularly small. Each sub-block, the smallest geographic unit available in the data, has a population of roughly 150 to 400 households. In each sub-block in our data, up to 10 households are sampled. We call this level of geography the *neighbourhood*. To reflect the fact that much social activity is within religion and caste groups, we interact the neighborhoods with indicators of religion (Hindu or not) and caste (NSS scheduled caste/tribe or not). We refer to these groups defined by neighborhood, religion, and caste as *neighborhood-subcastes*, and use them as the peer groups in our analysis.[11]

Our sample includes all urban households in groups where we observe at least three households, the minimum required for our method of identification and estimation. To avoid expenditure outliers, we include only households that are between the 1$^{\text{st}}$ and 99$^{\text{th}}$ percentiles of household expenditure in each state. We also restrict our sample to households with 12 or fewer members, whose head is aged 20 or more. Together, these restrictions drop roughly 4% of the sample.

Table 1 shows summary statistics for our sample. The number of observed households in each group averages around 5 (with a range from 3 to 10), which is a small share of the several hundred households that comprise each group in the population. These small within group samples illustrate the importance of showing identification and consistent estimation without assuming that many of the members of each group are observed.

For our main sample, we have a total of 4,599 distinct groups, and 24,757 distinct households. Our estimators use all unique household-pairs within each group, and we have a total of 128,640 such pairs.

The NSS collects item-level household spending and quantities for a large number of items. We consider only nondurable consumption items, and compute total expenditure x_i as the sum of spending on these goods. Our main results use a two-good demand system of food and non-food. On average 47% of nondurable expenditure is on food. An alternative specification we consider uses a three-good demand system of food, fuel, and other.

For instruments, we use three other survey rounds (the 59$^{\text{th}}$, 60$^{\text{th}}$, and 62$^{\text{nd}}$) to construct neighbourhood average expenditure in other years, $\widehat{x}_{g,-t}$, where $-t$ denotes years (survey rounds) other than the one that our model is estimated using. We use functions of $\widehat{x}_{g,-t}$ as instruments for neighbourhood average food and non-food consumption $\widehat{y}_{g,-ii'}$. However,

[11] The NSS contains information on whether the household is in a scheduled caste or tribe, but not the exact subcaste. However, since subcastes are typically geographically concentrated, we expect that the neighborhood-religion-scheduled caste groups will mostly capture subcastes as well.

since the neighbourhood identifiers in the NSS are not consistent over time (and are not linked to external information like neighborhood name), we cannot identify the exact same neighborhoods in other years. For each group g we therefore construct $\widehat{x}_{g,-t}$ using all observations from other years in the same district as g. As discussed earlier, these remain valid as instruments as long as they include some other members from the group g in other years. Our groups are spread across 535 districts, which are subunits of 20 states.

We construct prices of our demand aggregates at the state level, following Deaton (1998). We fist compute state-item average unit-value prices for the subset of items for which we have quantity data. Then, in a second stage, we aggregate these state-item-level unit value prices into state-level food and non-food prices using a Stone price index, with weights given by the overall sample average spending on each item. [12]

We condition on 7 demographic variables $\mathbf{z}$. These are household size minus 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 (equalling zero for uneducated or illiterate household heads).

Table 1 shows summary statistics at the level of the household, and at the level of the household-pairs used for estimation. Total expenditures and the spending components are expressed in units of average household expenditure. Only 26% of households have at least a high school education, and almost all households have married household heads. Roughly 14% of households have ration cards entitling them to subsidized basic foods.

IV.B Generic model

Our demand model assumes 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 $\mathbf{q}$.

Before proceeding with our main structural results, we implement the simpler generic model of equation (15) to examine these key assumptions. 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.

We use the same data and group definitions as in our main analysis, and similarly let y_i equal expenditures on food and x_i equal total household expenditure. We report the main

[12] In a typical state, these prices are computed as averages of roughly 2000 observations. Given this relatively large number of observations, we do not attempt to instrument for possible remaining measurement errors in these constructed price indices.

results of this analysis in Table 3. We confirm that that peer-average food expenditures significantly affects demand for food, and that both linear and quadratic terms in the budget x_i are statistically significant. The estimated peer effects in the generic fixed effects model are relatively imprecise, in part because the generic model does not exploit all the restrictions inherent in the structural demand model. We discuss these preliminary results in full in Appendix Section B.1.

IV.C Baseline Model

Our baseline structural model is a 2-good demand system (food vs other nondurable expenditure), as given by equation (11), and estimated by GMM using the associated moment conditions (18) and (21) for fixed- and random-effects, respectively. Both models use pairwise data based on all unique pairs of observations within each group, with standard errors clustered at the district level to obtain valid inference.[13]

Our fixed-effects approach involves substituting the leave-two-out within-group sample average quantity $\widehat{q}_{gj,-ii'}$ for the within-group mean $\overline{q}_{gj}$, and differencing across people within groups. Thus, we substitute $\widehat{q}_{gj,-ii'}$ for $\overline{q}_{gj}$ in the definition of X_i (eq. (10)) to create $\widehat{X}_i$:

$$\widehat{X}_i = x_i - R_{11}p_{1g} - R_{22}p_{2g} - \left(A_{11}\widehat{q}_{g1,-ii'} + \mathbf{C}_1'\mathbf{z}_i\right)p_{1g} - \left(A_{22}\widehat{q}_{g2,-ii'} + \mathbf{C}_2'\mathbf{z}_i\right)p_{2g},$$

and substitute $\widehat{q}_{gj,-ii'}$ for $\overline{q}_{gj}$ and $\widehat{X}_i$ for X_i in the demand equation (11). Then, we difference the demand equation across individuals within groups to generate a moment condition analogous to (18):

$$E[(p_{1g}q_{1i} - p_{1g}q_{1i'} - (\widehat{X}_i^2 - \widehat{X}_{i'}^2)e^{-(b_1 \ln p_{1g} + (1-b_1)\ln p_{2g})}d_1 - (\widehat{X}_i - \widehat{X}_{i'})b_1 + \mathbf{C}_1' p_{1g}\left(\mathbf{z}_i - \mathbf{z}_{i'}\right))\mathbf{r}_{gii'}] = 0. \tag{22}$$

Notice that, as in the generic model, many group-varying terms, including $A_{11}p_{1g}\overline{q}_{g1}$, drop out as a result of this differencing. Further, since $\left(\widehat{X}_i - \widehat{X}_{i'}\right) = x_i - x_{i'} - \mathbf{C}_1'\left(\mathbf{z}_i - \mathbf{z}_{i'}\right)p_{1g} - \mathbf{C}_2'\left(\mathbf{z}_i - \mathbf{z}_{i'}\right)p_{2g}$, such variables are present only in the quadratic term $\left(\widehat{X}_i^2 - \widehat{X}_{i'}^2\right)$ via interactions between group-average quantities $\overline{q}_{g1}$ 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 (11)

[13] The fact that we use pairwise estimation within-groups implies that we should cluster no smaller than the group level. However, because the instruments are computed at the district level, we cluster at the larger level of the district. Typical districts contrain about 10 groups.

and compute a new variable $\widetilde{X}_{ii'}$ for the squared term as follows:

$$\widetilde{X}_{ii'} = \widehat{X}_i[x_i - R_{11}p_{1g} - R_{22}p_{2g} - \left(A_{11}q_{g1i'} + \mathbf{C}_1'\mathbf{z}_i\right)p_{1g} - \left(A_{22}q_{g2i'} + \mathbf{C}_2'\mathbf{z}_i\right)p_{2g}].$$

Finally, we substitute $\widehat{q}_{g1,-ii'}$ for $\overline{q}_{gj}$, $\widehat{X}_i$ for X_i and $\widetilde{X}_{ii'}$ for X_i^2 in the demand equation (11) to generate a moment condition analogous to (21):

$$E[(p_{1g}q_{1i} - \widetilde{X}_{ii'}e^{-(b_1 \ln p_{1g} + (1-b_1)\ln p_{2g})}d_1 - \widehat{X}_i b_1 - R_{11}p_{1g} - A_{11}p_{1g}\widehat{q}_{g1,-ii'} - \mathbf{C}_1'p_{1g}\mathbf{z}_i - p_{1g}v_0)\mathbf{r}_{gi}] = 0. \tag{23}$$

These moments use pair-specific instruments that differ between our fixed- and random-effects models. As discussed earlier, to instrument for $\widehat{q}_{gj}$, we construct group-averages at the district level from other time periods. Recall that the subscript $-t$ indicate averages from all other time periods. For both the fixed- and random-effects models, we create a group-level instrument $\check{q}_{gj}$ equal to the OLS predicted value of $\widehat{q}_{gj}$ conditional on $\widehat{x}_{g,-t}, \widehat{x^2}_{g,-t}, \sqrt{\widehat{x}_{g,-t}}, \widehat{x}^2_{g,-t}, \widehat{\mathbf{z}}_{g,-t}$.[14]

Let $\widetilde{\mathbf{z}}_i$ and $\widetilde{\mathbf{z}}_g$ be, respectively, the individually-varying and group-level subvectors of $\mathbf{z}_i$. In our baseline model, $\widetilde{\mathbf{z}}_i$ includes all covariates; however, when we consider additional heterogeneity in peer effects, we will additionally include group-level covariates in $\widetilde{\mathbf{z}}_g$. Letting $\cdot$ denote element-wise multiplication, our complete instrument list for the fixed-effects model is:

$$\mathbf{r}_{gii'} = \left(x_i^2 - x_{i'}^2\right), \left(x_i - x_{i'}\right) \cdot \left(1, \mathbf{p}_g \cdot \check{\mathbf{q}}_g, \mathbf{p}_g \cdot \widetilde{\mathbf{z}}_g\right), \mathbf{p}_g \cdot \left(\widetilde{\mathbf{z}}_i - \widetilde{\mathbf{z}}_{i'}\right) \cdot \left(1, \mathbf{p}_g \cdot \check{\mathbf{q}}_g\right), x_i\mathbf{p}_g \cdot \left(\widetilde{\mathbf{z}}_i - \widetilde{\mathbf{z}}_{i'}\right).$$

Our instrument list for the random-effects model is:

$$\mathbf{r}_{gi} = \left(1, \mathbf{p}_g, \mathbf{p}_g \cdot \check{\mathbf{q}}_g, \mathbf{p}_g \cdot \mathbf{z}_i\right), x_i \cdot \left(1, \mathbf{p}_g, x_{it}, \mathbf{p}_g \cdot \check{\mathbf{q}}_g, \mathbf{p}_t \cdot \mathbf{z}_{ig}\right), \mathbf{p}_g \cdot \mathbf{p}_g.$$

The last term provides instruments for v_0 in equation (20).

Our primary focus is on the peer effects given by elements of the matrix $\mathbf{A}$. We start with the simplest and most interpretable version of this structural model, where $\mathbf{A} = a\mathbf{I}_J$ is a diagonal matrix with the scalar a replicated in each element of the main diagonal. In this specification, an increase in the group-average food quantity of δ increases needs for food by $a\delta$, and an increase in the group-average non-food quantity of δ increases needs for non-food nondurables by the same $a\delta$. Also, having $\mathbf{A}$ be diagonal means that group-average food quantities have no effect on needs for non-food nondurables (and vice versa). We relax these

[14] We do this instead of using all the separate variables as instruments for $\mathbf{q_g}$ to reduce the dimensionality of our instrument vector. This dimension reduction is needed for feasibility of our GMM estimator, because $\check{\mathbf{q}}_g$ is multiplied by the demographic controls to generate the final instrument vector.

restrictions later.

In this restricted version of the model, the welfare implications of peer effects simplify. Needs are given by $\mathbf{f}_i = \mathbf{A}\overline{\mathbf{q}}_g + \mathbf{C}\mathbf{z}_i$ and group-average expenditure is given by $\overline{x}_g = \boldsymbol{p}'\overline{\mathbf{q}}_g$, so when $\mathbf{A} = a\mathbf{I}_J$, the cost of needs, $\boldsymbol{p}'\mathbf{f}_i$, simplifies to $\boldsymbol{p}'\mathbf{f}_i = a\overline{x}_g + \boldsymbol{p}'\mathbf{C}\mathbf{z}_i$. Consequently, the scalar a equals the increase in the rupee cost of needs, $\boldsymbol{p}'\mathbf{f}_i$, of a one rupee increase in group-average expenditure $\overline{x}_g$.

IV.D Baseline Estimates and Alternative Group Sizes

Table 2 gives estimates of the scalar a. In our baseline model, groups are defined by neighborhood-subcastes, that is, a group is people who live in the same neighborhood, are of the same religion (either Hindu or not), and are of the same caste status (either scheduled caste or not). For comparison, we also consider two larger group sizes: people who live in the same neighborhood regardless of religion and caste, and people who live in the same district regardless of religion and caste.

Note that neighborhoods have populations of roughly 150 to 400 households, of which at most 10 are observed in our sample. Districts are much larger than neighborhoods, with populations of roughly 500,000 to 3,000,000 households. In our data, we observe 5.4 households from the average neighborhood-subcaste, while with the larger group definitions we average 6.9 and 53.1 observed households per group, respectively.

We report results for two samples. The upper half of Table 2 (Panel A) uses all the data available for each of the three group definitions, and so ends up with somewhat different samples for each. Panel B holds the sample constant across the group definitions, using only the observations from our baseline model (the smallest group definition).

Table 2 reports both random effects (RE) and fixed effects (FE) estimates of the scalar a, for all three group sizes. Columns (1) to (3) give RE, Columns (4) to (6) give FE, and columns (7) to (9) give the difference RE minus FE.

A key implementation question is how to define our groups. If we define them at too large a level, we should expect the estimated peer effects to be biased towards zero, because our estimate of group consumption $\widehat{q}_{gj,-ii'}$ will include consumption from non-peers. We should similarly expect the significance level of the estimates to fall if the defined groups are too large. In contrast, if we define our groups at too small a level, the estimator will be consistent but inefficient, because although we are grouping only households that do indeed have peer effects on each other, in each group we will leaving out some informative peers who were placed in another group.

For both RE and FE, we find that the larger group sizes have estimates that are closer to zero and have lower t statistics than our baseline, suggesting that our baseline groups, while quite small, are the most appropriate size (the largest group size FE estimate actually flips sign to negative, but is not statistically significant). We therefore focus our remaining analyses on the baseline neighborhood-subcaste group definition, reported in columns (3) and (6), and the difference between them in column (9).

As expected, the RE estimates have far lower standard errors than the FE estimates, because they are based on much stronger assumptions, and do not lose information from differencing. The RE point estimate of 0.606 in column (3) also turns out to be much larger than the FE estimate of 0.266 in column (6), and we reject equality of the coefficients (column (9)).

Random effects imposes strong restrictions on unobserved heterogeneity that may not be valid, and that fixed effects do not impose, potentially biasing the RE estimates. In particular, our estimated positive difference between RE and FE estimates is consistent with group-level preferences for food consumption v_{gj} being correlated with group expenditure levels, causing upward bias in the RE peer effects estimates. This is easiest to see in a simplified version of equation (13). Suppose that the true model was linear (so $d = 0$), and we instrumented for $\widehat{y}_g$ only with other-period group consumption $\widehat{x}_{g,-t}$. Then, positive correlation between group expenditure and group tastes (conditional on x_i) would result in upwards bias in the estimated peer effects for normal goods like food.

Since RE has much lower variance (indicated by standard errors) and is likely to be biased, to reduce mean squared errors it is common to propose shrinkage estimators that equal weighted averages of RE and FE estimates, trading off the bias of RE with the higher variance of FE (a recent example is Armstrong, Kolesár, and Plagborg-Møller 2020). We report both the RE and FE estimates in our remaining empirical analyses, so one may implement such shrinkage if desired. However, for simplicity in our later policy discussions, we will focus on the smaller FE coefficients as the most conservative estimate of the magnitude of peer effects.

To interpret our estimate of a, imagine first that just one household in a group had, ceteris paribus, an additional s rupees to spend. Compare this to the case where everyone in the group each had an additional 1000 rupees to spend. What s would give the household in the first case the same utility as in the second case? The answer is less than a 1000 rupees, because in the second case, peer effects reduce the utility of the increased spending. By our model, the answer is $s = 1000(1 - a)$, which is 394 rupees in the RE model and 734 rupees in the FE model.

Economic theory requires that a lie between zero and, roughly, one. It is greater than zero because our model is one of peer effects increasing perceived needs that take the form of costs, and it is less than about one to ensure that an equilibrium exists.[15] An encouraging feature of our estimates is that they lie well within this required range, without any such constraint being enforced in estimation.

IV.E Measurement Error in Group Means

The neighborhood-subcaste groups in our baseline analysis each have between 3 and 10 observed households, out of an average of around 200 households in the population. This suggests that the group mean measurement errors $\widehat{q}_{gj} - \overline{q}_{gj}$ are likely to be substantial. Much of the complexity in our GMM estimator entails constructing moments that remain valid in the presence of these measurement errors. Table 3 considers the impact of our measurement error corrections on the estimated values for a in both the RE and FE models. We should expect that, the smaller are the group definitions, the larger are the measurement errors in the estimates of each $\overline{q}_{gj}$, and hence the larger should be the effect of correcting for these measurement errors.

Regarding the direction of bias, one might expect measurement error in $\overline{q}$ to induce the usual attenuation (i.e., bias towards zero) that is standard in linear models with measurement error. However, the nonlinearity of our models and our estimators could cause bias in either direction. A priori, we expect standard attenuation bias to play a larger role in the RE model, because in that model the parameter a is primarily identified as the coefficient of the estimate of $\overline{q}$ itself, while in the FE model, due to differencing, a is identified only off of differences of interactions between $\overline{q}$ and other covariates.

To assess the impact of our corrections for measurement error, we replace the instruments in our models with stronger instruments that would be valid in the absence of measurement error. In particular, instead of instrumenting $\widehat{q}_{gj}$ with district level-averages from other time periods, we instrument $\widehat{q}_{gj}$ with group-level averages from the current time period. So everywhere that $\widehat{x}_{g,-t}$ and $\widehat{\mathbf{z}}_{g,-t}$ appear in our estimators, we replace them with $\widehat{x}_g$ and $\widehat{\mathbf{z}}_g$. As a result, the total number and types of moments remains exactly the same as in our baseline estimates.

Table 3 is analogous to columns (1) to (6) of Table 2, but is estimated with the instruments that do not correct for measurement error. This should be compared to the corresponding entries in Table 2. Columns (3) and (6) are still our preferred group size specifications.

Both the RE and FE estimates show considerable differences between estimates with and without the measurement error correction. As expected, the smaller the group sizes, the larger the differences between the corrected and uncorrected estimates.

[15]The exact value that is necessary to ensure that an equilibrium exists has a complicated expression which we derive in the Appendix, but this value is near one.

In the RE models, we see standard attenuation bias dominating, and the magnitude of the bias appears very large: uncorrected estimates are about half the size of the corrected estimates for the largest group size, attenuating all the way to about one tenth the size of the corrected estimate for our baseline, which is the smallest group size. The measurement error corrected RE estimates also have larger standard errors than the uncorrected estimates, due to the fact that the instruments are less informative in the former case. Since both estimators would be consistent in the absence of measurement error, we can form a Hausman test to compare the estimators, and the uncorrected estimators are rejected.

The direction and size of bias is different for the FE estimator. Here, at all three group sizes, the uncorrected estimates are about twice as large as the corrected, suggesting a significant impact of nonlinearity and differencing on the size and direction of bias in the FE models. As with the RE models, the uncorrected FE estimates have smaller standard errors than the corrected estimates, and Hausman tests reject the uncorrected estimates. We conclude that our corrections for measurement errors due to small within group sample sizes are empirically justified and important.

IV.F Alternative Specifications and Robustness Checks

IV.F.1 Peer effects by demographic groups

In Tables 2 and 3, the peer effect parameter a is restricted to be the same for all types of households. In Table 4, we allow a to vary with observed household characteristics. In columns (1) and (5), we replicate columns (3) and (6) from Table 2, where the group is defined at the neighborhood-subcaste level, for the RE and FE models, and a is a fixed value. In columns (2) and (6), we allow a to depend on whether the household is Hindu or not, and whether they come from a scheduled (disadvantaged) caste. In columns (3) and (7), we define groups at the neighborhood-subcaste-landownership level, and allow a to depend on the landownership indicator variable. In columns (4) and (8), we define groups by neighborhood-subcaste-high-school attainment, and allow a to depend on the high-school attainment indicator.

Columns (2) and (6) show estimated differences in peer effects across Hindu vs non-Hindu and scheduled vs non-scheduled tribe/caste. The left-out category (picked up by the constant) is non-scheduled Hindu. The RE estimates show some significant differences in peer effects, but the FE estimates do not, and most of estimated differences have the opposite sign in the FE vs RE models.

Columns (3) and (7) allow a to depend on the household level land-ownership indicator. Both the FE and RE models show landowners having larger peer effects than landless households, but the magnitudes differ dramatically, with RE estimates implying a small difference, while FE showing the landless having almost no peer effects. As before, the standard errors on the FE models are all much larger than the RE standard errors.

Columns (4) and (8) allow a to depend on a household level high-school attainment indicator, defined to equal to 1 if the household head has at least high school education and zero otherwise. Here the FE and RE models disagree, with the FE model showing the more educated households having larger peer effects, while the RE model shows the opposite.

Particularly when focusing on the FE estimates, our estimated peer effects are larger for higher socio–economic status groups. A possible explanation is that the poorest households in India are close enough to subsistence that it is more costly to engage in status competitions. This is similar to Akay and Martinsson's (2011) finding for very poor Ethiopians.

IV.F.2 Peer effects by A matrix specification

Next, Table 5 considers what happens when we relax the restriction that $\mathbf{A} = a\mathbf{I}_J$ for a scalar a. Since needs are given by $\mathbf{f}_i = \mathbf{A}\overline{\mathbf{q}}_g + \mathbf{C}\mathbf{z}_i$, the money cost of the part of needs driven by peer effects is given by $\boldsymbol{p}'\mathbf{A}\overline{\mathbf{q}}_g$. In the previous subsections, with $\mathbf{A} = a\mathbf{I}_J$, this cost of needs due to peer effects is $\boldsymbol{p}'\mathbf{A}\overline{\mathbf{q}}_g = a\left(p_1\overline{q}_{1g} + p_2\overline{q}_{2g}\right) = a\overline{x}_g$, and so is proportional to group mean total expenditures $\overline{x}_g$. When we allow $\mathbf{A}$ to be an unconstrained diagonal matrix, this cost of needs becomes $\boldsymbol{p}'\mathbf{A}\overline{\mathbf{q}}_g = a_{11}p_1\overline{q}_{1g} + a_{22}p_2\overline{q}_{2g}$. This allows for the possibility that group-average food expenditure, $p_1\overline{q}_{1g}$, and group-average non-food nondurable expenditure, $p_2\overline{q}_{2g}$, have different effects on needs. Finally, when $\mathbf{A}$ is completely unrestricted, we get $\boldsymbol{p}'\mathbf{A}\overline{\mathbf{q}}_g = a_{11}p_1\overline{q}_{1g} + a_{21}p_2\overline{q}_{1g} + a_{12}p_1\overline{q}_{2g} + a_{22}p_2\overline{q}_{2g}$.

In columns (3) and (6) of Table 2, we reproduce columns (3) and (6) of Table 5, reporting the estimate of the scalar a where $\mathbf{A} = a\mathbf{I}_J$. In Columns (2) and (5) of Table 5, we let $\mathbf{A}$ be an unconstrained diagonal matrix, and report its two estimated diagonal elements, a_{11} and a_{22}. And in columns (1) and (4) of Table 5, we give estimates of all four elements of $\mathbf{A}$ where $\mathbf{A}$ is completely unrestricted. For these estimates, we again define groups as neighborhood-subcaste.

The main difficulty in estimating these more general models is multicollinearity. As people's income rises, they tend to spend more on both food and non-food items. As a result $p_1\overline{q}_{1g}$ and $p_2\overline{q}_{2g}$ (group average food and nonfood expenditures, respectively) in the diagonal $\mathbf{A}$ model tend to be highly correlated across groups. This problem is worse still in the unrestricted $\mathbf{A}$ model, where $p_1\overline{q}_{1g}$, $p_2\overline{q}_{1g}$, $p_1\overline{q}_{2g}$, and $p_2\overline{q}_{2g}$ are all highly multicollinear, both because group-average quantities of food and nonfood are positively correlated with each other, and because prices are positively correlated with each other across states.

Considering first the RE estimates with an unrestricted diagonal $\mathbf{A}$ matrix (column (2) of

Table 5), we see estimated values of 0.639 and 0.572 for a_{11} and a_{22}, respectively. These are similar in magnitude to each other, and similar to the estimated value of 0.606 for a in the baseline RE model. Although the two values are similar in magnitude, they are estimated precisely enough to reject the hypothesis that they are identical. Turning to the RE estimates with an unrestricted $\mathbf{A}$ matrix, column (1), we again find the estimated magnitudes of a_{11} and a_{22} are similar to each other (though lower than before), and the difference between them is now statistically insignificant. The off diagonal elements of this unrestricted $\mathbf{A}$ matrix are both statistically insignificant.

Taken together we interpret these results as evidence that imposing the restrictions $a_{11} = a_{22}$ and $a_{12} = a_{21} = 0$, as in our baseline model, is at least a reasonable approximation.

In contrast to the RE model, we see evidence that the above discussed multicollinearity overwhelms the FE model. Column (5) shows infeasibly large estimates of a_{11} and a_{22} with opposite signs and greatly increased standard errors, and even more extreme estimates in column (4) where all four elements of $\mathbf{A}$ have impossibly large magnitudes and varying signs. These are all common hallmarks of substantial positive multicollinearity.

We should expect that the multicollinearity issues among the $p_j\overline{q}_{kg}$ terms would be much more severe in the FE model, and not just because it is based on a weaker set of assumptions. The identification of $\mathbf{A}$ in the FE estimator comes only from interaction terms between each $p_j\overline{q}_{kg}$ and the budget x_i. This is due to the fact that the level terms for each $p_j\overline{q}_{kg}$ get differenced away. In contrast, the identifying variation for $\mathbf{A}$ in the RE estimator comes from both the level terms $p_j\overline{q}_{kg}$ and their interactions with x_i.

We take from these results that the multicollinearity of group-average expenditures is too severe in our data to get trustworthy estimates of variation in the elements of $\mathbf{A}$ in our preferred fixed effect specification, but that baseline restriction $\mathbf{A} = a\mathbf{I}_J$ is reasonable.

IV.F.3 A Three Goods Model

All the models presented so far have been demand systems with $J = 2$ goods (food and non-food). When $J = 2$, we only need to estimate a single demand equation (since the other is determined by the restriction that consumers exhaust their budget). However, our theorems show identification of peer effect parameters in demand systems where J is any number of goods. In Table 6, we present estimates of a $J = 3$ equation demand model, having two equations we need to estimate. The 3 goods are taken to be food, fuel and other nondurable goods. The former non-food category is now divided into fuel and other, so total expenditures x_i for each household remains the same as before.

We report estimates for the RE and FE models, with an unrestricted diagonal $\mathbf{A}$ matrix in columns (1) and (3) of Table 6, and with the restriction that $\mathbf{A} = a\mathbf{I}_J$ in columns (2) and

(4). As before, groups are defined at the neighborhood-subcaste level.

In the RE models, a in column (2) and the varying diagonal elements of $\mathbf{A}$ in column (1) are all significant and larger than before, ranging from 0.740 to 0.938. Since adding more goods should not increase the magnitude of the overall peer effects, we take this as additional evidence that the restrictions imposed by the RE model may not hold, and are likely inducing an upward bias. We also perform a Hausman test of the RE model against the FE model, and again reject the additional restrictions imposed by the RE model.

In the FE model, we again see evidence of multicollinearity in column (3), with two elements of the estimated $\mathbf{A}$ diagonal being extremely large and positive, and one being extremely large and negative. However, in column (4), we see a statistically significant estimate of a of 0.296, which agrees very well with the FE estimate of 0.266 we had in the two goods baseline model. We take this as additional evidence in favor of the FE model with $\mathbf{A} = a\mathbf{I}_J$.

IV.G Are Peer Expenditures Really Negative Externalities?

Our findings suggest that higher peer expenditures makes consumers behave, at the margin, as if they were poorer. We take this to mean that, in a welfare sense, they feel poorer. An alternative explanation could be the presence of network effects from peer consumption, where peer consumption directly increases the utility of the goods one consumes. An example could be something like a phone, which becomes more valuable when other consumers also have phones.

We address this concern by directly estimating the effect of peer expenditure on subjective well-being data, and confirm that, conditional on household income, higher levels of peer group expenditures are associated with lower satisfaction. We take this as confirmation that our demand estimates do indeed reflect lower welfare resulting from increasing peer expenditures.

For this exercise we use data from the Indian modules of the World Values Surveys (WVS).[16] 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. The WVS also includes information on household income (not precise income levels, but a few income brackets). We divide the sample into groups based on state and religion (since neighborhood and caste identifiers are not available in the WVS), and merge in estimates of average expenditure for each group from the NSS data. Summary statistics for this data are reported in Table 1 in the Appendix.

[16]We use the 4$^{\text{th}}$ (2001), 5$^{\text{th}}$ (2006), and 6$^{\text{th}}$ (2014) waves.

Interpreting ordinal self-reported well-being as a crude measure of utility, we regress this self-reported well-being (both linearly and by ordered logit) on one's own income bin and on the average expenditure in one's group. The results are reported in Table 2 in the Appendix. We find that the resulting coefficient estimates have signs that are consistent with our theory: higher income increases self reported well being, but higher group expenditure decreases it. A 1000 rupee increase in peer group expenditure (relative to a mean of 5554, with standard deviation of 2580) decreases self reported well being by 15% of a standard deviation, which is in line with the welfare effects we found using our structural model.[17]

We also use the WVS data to test another implication of our model, which is that peer expenditures only affect utility through a linear index (this index being the function that equals the cost of needs). Columns (3) and (6) of Table 2 depart from a linear index by interacting group expenditure with dummies for being in the top two income groups. We find no differential effect of group expenditure for the top income groups (results are similar if we interact group expenditure with one or three top income groups). We take this as evidence in favor of our linear index structure for peer effects.

V Implications for Tax and Transfers Policy

Our finding that perceived needs rise with peer group average consumption 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 negative externalities on one's 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. Here we apply the same logic to our estimated consumption peer effects, and in particular show how potentially large free lunch gains are possible.

[17]In principle, one could use self reported well being data to estimate a, the effect of peer expenditure in money-metric terms. There are two issues with this approach. First, no existing data record both consumption and self reported well being. Second, and more importantly, this approach (as well as that of other papers in the literature, such as Luttmer (2005), that apply this approach) relies on a random-effects assumption that expenditure and group expenditure are uncorrelated with other determinants of self reported well being. A key advantage of our utility-derived demand model is that the FE approach allows identification even when group preferences that are correlated with group expenditure. Given these two issues, we take the self reported well being results here only evidence in favor of negative consumption externalities, and do not attempt to use them to back out other measures of the welfare cost of peer effects.

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, all else equal, 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 simplicity, round our baseline estimate of $a = 0.266$ to $1/4$. Suppose you experience a 4 rupee tax increase, and for simplicity let your marginal propensity to consume be 100%. If your peers also have their taxes increase by the same amount, then your loss in utility will only be equivalent to that of a 3 rupee tax increase. The reason is that although your net income, and therefore expenditure, will have dropped by 4 rupees, so will have that of your peers. Consequently, your needs will have dropped by $1/4 * 4 = 1$ rupee, so that your net loss in money-metric utility is 3 rupees.

However, to fully evaluate the effect of the tax increase, we must also consider 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 $a = 1/4$, then the welfare gains from reduced expenditures on needs by the taxed consumers will be 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 or government services.

We found evidence that the size of the peer effects may be smaller for poorer and less 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 the marginal utility of money is the same for all consumers.

An alternative way to exploit the potential free lunch associated with reduced perceived needs from taxing peers is to spend the resulting tax revenues on public goods or government services. To the extent that jealousy or envy are the underlying cause of the peer externalities we identify, public goods and services may not invoke those effects (or at least induce smaller peer effects), because by definition public goods are consumed by all members of the group. This suggests that public goods and services may provide at least 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), which 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. 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 the provision of public goods or services that provide utility to the poor but are equally available to all households. Such public goods and services might include clean water, public sanitation, better air quality, or improved fire or police protection.

A 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 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). Then, the public cost of providing 5kg of rice at the subsidized price of 3 rupees per kg is 60 rupees per month per person. Suppose there is no waste, so that the private consumption of people increases by 60 Rupees per month per person. Using $a = 1/4$, this implies that needs rise by 15 Rupees per month per person. Thus the government's expenditures of 60 rupees only increases money metric utility by 45 rupees per person per month. This is in contrast to a benefit of up to 60 rupees per person per month that might be obtained by provision of public goods. 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 up to roughly 180 billion rupees (over 2 billion US dollars) per year.

This toy calculation comes with many caveats,[18] and is not intended to be a rigorous analysis of alternatives to the PDS. It is only intended to highlight the potentially enormous impacts that accounting for peer effects can have on the evaluation of tax and transfer policies.

VI Conclusions

We show identification and GMM estimation of peer effects in a model where most members of each peer group are not 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

[18] Some caveats are that the benefits of this alternative might be reduced to the extent that some households derive less utility from the public good than others, or if the public good also generates negative peer effects, or if households in total derive less utility from the public good than it costs to provide. The effects might also be reduced to the extent that households given ration cards, who are therefore relatively poor, may have smaller than average peer effects, and some are likely to be at subsistence level where few such effects are possible. On the other hand benefits could 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.

small and 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, each member is only observed once, and detailed network structure is not available. We obtain these results both for a generic quadratic model, and for a utility-derived demand model. The methods we use to identify and estimate these effects could potentially have broad application to other social network models.

We propose a utility-derived consumer demand model where a consumer's perceived needs for each good depends in part on the average consumption of goods among the other members of the consumer's peer group. We show how this model can be used for welfare analyses, and in particular to identify what fraction of total expenditure increases are spent on "keeping up with the Joneses" type peer effects.

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 about 26 rupees or more in most peer groups. This means that the increase in utility you experience if you and everyone else in your peer group spends 100 more rupees (say, because of a tax cut) is the same as the increase in utility you would get from spending only $100 - 26 = 74$ more rupees if no one else in your peer group increased their spending.

These results can at least partly explain the Easterlin (1974) paradox, in that income growth over time, which increases people's consumption budgets, results in lower utility growth than is implied by standard demand models that ignore peer effects.

These results also suggest that income or consumption taxes can have far lower negative effects on consumer welfare than are implied by standard models. This is because a tax that reduces my expenditures by 100 rupees will, if applied to everyone in my peer group, have the same effect on my utility as a tax of only 74 rupees that ignores the peer effects. This implies that about a fourth of the money people might get back from an across the board tax cut doesn't increase utility, but instead is spent on increased perceived needs due to peer effects. 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 or government services (that are less likely to induce peer effects themselves) 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 welfare gains of hundreds of billions of rupees could 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 effects than richer households.

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.

Armstrong, T. B., M. Kolesár, and M. Plagborg-Møller (2020), " Robust Empirical Bayes Confidence Intervals," Unpublished manuscript, Yale and Princeton Universities.

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.

NSS (2019), "Urban Frame Survey" http://mospi.nic.in/urban-frame-surveyufs

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,_2013.

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 consumption data

	Observations (N=24,757)				Pairs (N=128,640)			
	Mean	SD	Min	Max	Mean	SD	Min	Max
x_i	.99	.59	.072	4.7	1	.59	.072	4.7
q_i food	.44	.21	0	2	.44	.21	0	2
q_i non-food	.44	.33	.0069	2.7	.44	.32	.0069	2.7
$\widehat{q}_{g,ii'}$ food					.44	.15	.027	1.7
$\widehat{q}_{g,ii'}$ non-food					.44	.24	.02	2.4
p food	1.1	.08	.94	1.3	1.1	.083	.94	1.3
p non-food	1.2	.11	.94	1.5	1.2	.12	.94	1.5
(Household size -1)/10	.38	.21	0	1.1	.38	.21	0	1.1
Age (household head, in 10 years)	.39	.11	.17	.82	.4	.11	.17	.82
Household head married	.84	.36	0	1	.84	.36	0	1
Log land owned	.15	.35	0	2.3	.16	.35	0	2.3
Ration card	.14	.35	0	1	.13	.34	0	1
Literate but no HS	.46	.5	0	1	.47	.5	0	1
High school or greater	.26	.44	0	1	.26	.44	0	1

Table reports summary statistics for estimation sample.

Table 2: Estimated peer effects by group definition

	RE			FE			Difference		
	District	Neighborhood	Neighbor-						
hood- caste	District	Neighborhood	Neighbor-						
hood- caste	District	Neighborhood	Neighbor-						
hood- caste									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: All data									
A (group consumption)	0.334***	0.558***	0.606***	-0.228*	0.088	0.266**	0.562***	0.470***	0.341***
	(0.044)	(0.036)	(0.036)	(0.138)	(0.121)	(0.119)	(0.131)	(0.115)	(0.114)
J overid stat	14138.36	1264.97	653.76	22426.79	2130.89	1305.88			
p-value	0.000	0.000	0.000	0.000	0.000	0.000			
N pairs	3,761,688	195,282	128,640	3,761,688	195,282	128,640	3,761,688	195,282	128,640
N households	30,184	29,462	24,757	30,184	29,462	24,757	30,184	29,462	24,757
Average group size	53.14	6.88	5.38	53.14	6.88	5.38	53.14	6.88	5.38
Panel B: Consistent sample									
A (group consumption)	0.367***	0.568***	0.606***	-0.284*	0.156	0.266**	0.651***	0.412***	0.341***
	(0.045)	(0.034)	(0.036)	(0.153)	(0.133)	(0.119)	(0.147)	(0.128)	(0.114)
J overid stat	9877.93	985.24	653.76	16108.87	1583.52	1305.88			
p-value	0.000	0.000	0.000	0.000	0.000	0.000			
N pairs	2,564,578	150,184	128,640	2,564,578	150,184	128,640	2,564,578	150,184	128,640
N households	24,757	24,757	24,757	24,757	24,757	24,757	24,757	24,757	24,757
Average group sample size	43.90	6.28	5.38	43.90	6.28	5.38	43.90	6.28	5.38

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Selected estimates for structural demand model, Controls include household size, age, marital status, land owned, ration card indicator, education, religion, and group size. Standard errors clustered at the district level. ** $p < 0.05$, *** $p < 0.01$

Table 3: Peer effects without measurement error correction

	RE			FE		
	District	Neighborhood	Neighbor-			
hood- caste	District	Neighborhood	Neighbor-			
hood- caste						
	(1)	(2)	(3)	(4)	(5)	(6)
A (group consumption)	0.143***	0.038**	0.054***	0.470**	0.559***	0.529***
	(0.031)	(0.017)	(0.016)	(0.215)	(0.089)	(0.090)
J overid stat	11354.53	1340.51	1013.93	17386.44	1651.41	1300.67
p-value	0.000	0.000	0.000	0.000	0.000	0.000
N pairs	2,564,578	150,184	128,640	2,564,578	150,184	128,640
N households	24,757	24,757	24,757	24,757	24,757	24,757
Average group sample size	43.90	6.28	5.38	43.90	6.28	5.38

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Selected estimates for structural demand model, Controls include household size, age, marital status, land owned, ration card indicator, education, religion, and group size. Standard errors clustered at the district level. ** $p < 0.05$, *** $p < 0.01$

Table 4: Peer effects by demographic group

	RE				FE			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	0.606***	0.606***	0.600***	0.694***	0.266**	0.255**	0.078	0.145
	(0.036)	(0.035)	(0.057)	(0.041)	(0.119)	(0.121)	(0.125)	(0.143)
Scheduled non-Hindu		0.168**				0.130		
		(0.066)				(0.237)		
Scheduled Hindu		0.247***				-0.285		
		(0.087)				(0.330)		
Non-scheduled non-Hindu		0.179***				-0.075		
		(0.055)				(0.153)		
Owns land			0.031				0.446***	
			(0.054)				(0.136)	
High school or greater				-0.186***				0.454***
				(0.057)				(0.163)
p-value heterogeneity		0.00	0.56	0.00		0.68	0.00	0.01
N pairs	128,640	128,640	100,756	84,052	128,640	128,640	100,756	84,052
N households	24,757	24,757	21,696	20,233	24,757	24,757	21,696	20,233

Selected estimates for structural demand model, Controls include household size, age, marital status, land owned, ration card indicator, education, religion, and group size. A (group consumption) represents the structural effect of group consumption on own consumption, constrained to be the same for all goods. Standard errors clustered at the district level. ** $p < 0.05$, *** $p < 0.01$

Table 5: Peer effects by A matrix specification

	RE			FE		
	(1)	(2)	(3)	(4)	(5)	(6)
A (group food on food consumption)	0.411**	0.639***	0.606***	9.741***	2.228***	0.266**
	(0.171)	(0.036)	(0.036)	(2.066)	(0.382)	(0.119)
A (group non-food on non-food consumption)	0.452***	0.572***	0.606***	5.400***	-0.911***	0.266**
	(0.171)	(0.034)	(0.036)	(1.577)	(0.276)	(0.119)
A (group food on own non-food consumption)	-0.397			-7.695***		
	(0.275)			(1.828)		
A (group non-food on own food consumption)	-0.095			-6.383***		
	(0.102)			(1.860)		
p-value equality	0.896	0.001		0.000	0.000	
p-value diagonal	0.002			0.000		
N pairs	128,640	128,640	128,640	128,640	128,640	128,640
N households	24,757	24,757	24,757	24,757	24,757	24,757

Selected estimates for structural demand model, Controls include household size, age, marital status, land owned, ration card indicator, education, religion, and group size. Standard errors clustered at the district level. ** $p < 0.05$, *** $p < 0.01$

Table 6: Estimated peer effects in a three-good demand system

	RE		FE	
	(1)	(2)	(3)	(4)
A (group food on food consumption)	0.848***	0.932***	2.393***	0.296***
	(0.023)	(0.014)	(0.426)	(0.100)
A (group fuel on own fuel consumption)	0.938***	0.932***	2.820***	0.296***
	(0.018)	(0.014)	(0.913)	(0.100)
A (group other on own other consumption)	0.740***	0.932***	-1.387***	0.296***
	(0.023)	(0.014)	(0.334)	(0.100)
Hausman H			42.151	41.701
p-value			0.00	0.00
p-value equality	0.000		0.000	
N pairs	128,640	128,640	128,640	128,640
N households	24,757	24,757	24,757	24,757

Selected estimates for structural demand model, Controls include household size, age, marital status, land owned, ration card indicator, education, religion, and group size. Standard errors clustered at the district level. ** $p < 0.05$, *** $p < 0.01$

Appendices for online publication

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.[19]

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 $\overline{y}_g = E\left(y_i \mid i \in g\right)$, $\widehat{y}_{g,-ii'} = \sum_{l \in g, l \neq i,i'} y_l/(n_g - 2)$, and $\varepsilon_{yg,-ii'} = \widehat{y}_{g,-ii'} - \overline{y}_g$, so $\overline{y}_g$ is the true

[19] 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.

group mean outcome and $\widehat{y}_{g,-ii'}$ is the observed leave-two-out group average outcome in our data, and $\varepsilon_{yg,-ii'}$ is the estimation error in the leave-two-out sample group average. Define $\overline{\mathbf{x}}_g = E(\mathbf{x}_i \mid i \in g)$, $\overline{\mathbf{x}\mathbf{x}'_g} = E(\mathbf{x}_i\mathbf{x}'_i \mid i \in g)$, and similarly define $\widehat{\mathbf{x}}_{g,-ii'}$, $\widehat{\mathbf{x}\mathbf{x}'}_{g,-ii'}$, $\varepsilon_{\mathbf{x}g,-ii'}$ and $\varepsilon_{\mathbf{xx}g,-ii'}$ analogously to $\widehat{y}_{g,-ii'}$, and $\varepsilon_{yg,-ii'}$.

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

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

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 $i' \in g$, $\Pr(\mathbf{x}_{ik} \neq \mathbf{x}_{i'k}) > 0$. Unobserved v_g are group level fixed effects. Unobserved errors u_i are independent across groups g and have $E(u_i \mid \text{all } \mathbf{x}_{i'} \text{ having } i' \in g \text{ where } i \in g) = 0$. The number of observed groups $G \to \infty$. For each observed group g, we observe a sample of $n_g \geq 3$ observations of $y_i, \mathbf{x}_i$.

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

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

In Assumption A2 $d \neq 0$ 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 $\overline{y}_g = E(y_i \mid i \in g)$, 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

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

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

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

where

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

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

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

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

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

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

The above derivations outline how we obtain identification, while the formal proof is given in Theorem 1 below. 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

$$\left(\widehat{a}, \widehat{b}_1, \ldots \widehat{b}_K, \widehat{d}\right) =$$

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

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,

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

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

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

Using the same analysis as in Appendix A.2,

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

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

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

Denote

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

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

$$\left(\widehat{\mathbf{a}}_1', ..., \widehat{\mathbf{a}}_J', \widehat{\mathbf{b}}_1', ...\widehat{\mathbf{b}}_J', \widehat{d}_1, ..., \widehat{d}_J\right)'$$

$$= \arg\min \left(\frac{\sum_{g=1}^{G}\sum_{(i,i')\in\Gamma_g}\mathbf{m}_{gii'}}{\sum_{g=1}^{G}\sum_{(i,i')\in\Gamma_g}1}\right)' \widehat{\Omega} \left(\frac{\sum_{g=1}^{G}\sum_{(i,i')\in\Gamma_g}\mathbf{m}_{gii'}}{\sum_{g=1}^{G}\sum_{(i,i')\in\Gamma_g}1}\right), \text{where}$$

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

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

A.4 Multiple Equation Generic Model With Random Effects

Here we provide the derivation of equation (20), 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).

$$y_i = \overline{y}_g^2 a^2 d + a\left(1 + 2\mathbf{b}'\mathbf{x}_i d\right)\overline{y}_g + \mathbf{b}'\mathbf{x}_i + \mathbf{b}'\mathbf{x}_i\mathbf{x}_i'\mathbf{b}d + v_g + u_i, \tag{34}$$

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

$$\overline{y}_g = \overline{y}_g^2 da^2 + a(2d\mathbf{b}'\overline{\mathbf{x}}_g + 1)\overline{y}_g + d\mathbf{b}'\overline{\mathbf{x}\mathbf{x}'_g}\mathbf{b} + \mathbf{b}'\overline{\mathbf{x}}_g + \mu, \tag{35}$$

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 $\varepsilon_{\mathbf{x}l} = \mathbf{x}_l - \overline{\mathbf{x}}_g$, $\varepsilon_{\mathbf{x}\mathbf{x}l} = \mathbf{x}_l\mathbf{x}_l' - \overline{\mathbf{x}\mathbf{x}'_g}$, 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

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

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

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

Therefore, we can write

$$y_i = \widehat{y}_{g,-ii'}y_{i'}a^2 d + a\left(1 + 2\mathbf{b}'\mathbf{x}_i d\right)\widehat{y}_{g,-ii'} + \mathbf{b}'\mathbf{x}_i + \mathbf{b}'\mathbf{x}_i\mathbf{x}_i'\mathbf{b}d + v_g + u_i + \widetilde{\varepsilon}_{gii'}, \tag{36}$$

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 $\varepsilon_{\mathbf{x}\mathbf{x}l}$.

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 $\varepsilon_{\mathbf{x}\mathbf{x}l}$ are independent across individuals and have zero means.

Assumption A8: For each group g, v_g is independent across groups with $E(v_g|\mathbf{x}, \overline{\mathbf{x}}_g, \overline{\mathbf{xx}'_g}) = \mu$ and we have the conditional homoskedasticity that $Var(v_g|\mathbf{x}, \overline{\mathbf{x}}_g, \overline{\mathbf{xx}'_g}) = \sigma^2$.

Let $v_0 = \mu - da^2\sigma^2$. It follows from Assumptions A6-A8 that, for any $l \neq i$, $E(\overline{y}_g\varepsilon_{yl}|\mathbf{x}_i, \overline{\mathbf{x}}_g, \overline{\mathbf{xx}'_g}) = 0$ and $E(\varepsilon_{yl}\mathbf{x}_i|\mathbf{x}_i, \overline{\mathbf{x}}_g, \overline{\mathbf{xx}'_g}) = 0$. Hence, $E(\widetilde{\varepsilon}_{gii'}|x_i, \overline{\mathbf{x}}_g, \overline{\mathbf{xx}'_g}) = -da^2E(\varepsilon_{yg,-ii'}\varepsilon_{y,i'}|\mathbf{x}_i, \overline{\mathbf{x}}_g, \overline{\mathbf{xx}'_g}) = -da^2Var\left(v_g\right)$ and

$$E(v_g + u_i + \widetilde{\varepsilon}_{gii'} \mid \overline{\mathbf{x}}_g, \overline{\mathbf{xx}'_g}, \mathbf{x}_i) = \mu - da^2\sigma^2 = v_0. \tag{37}$$

By construction $v_g + u_i + \widetilde{\varepsilon}_{gii'}$ is also independent of $\mathbf{r}_g$. Given this, equation (20) then follows from equations (36) and (37).

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 (8) 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 (8) 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.

We distinguish between elements of z that vary at the individual versus group level, writing $\mathbf{C}$ as $\mathbf{C} = \left(\widetilde{\mathbf{C}} : \mathbf{D}\right)$ for submatrices $\widetilde{\mathbf{C}}$ and $\mathbf{D}$, and replacing $\mathbf{C}\mathbf{z}_i$ in Equation 9 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.

Let $\boldsymbol{\alpha} = \mathbf{A}'\mathbf{p}$, $\beta = \mathbf{p}^{1/2\prime}\mathbf{R}\mathbf{p}^{1/2}$, $\widetilde{\boldsymbol{\gamma}} = \widetilde{\mathbf{C}}'\mathbf{p}$, $\boldsymbol{\kappa} = \mathbf{D}'\mathbf{p}$, $\boldsymbol{\delta} = \mathbf{b}/\mathbf{p}$, $\mathbf{C}\mathbf{z}_i = \widetilde{\mathbf{C}}\widetilde{\mathbf{z}}_i + \mathbf{D}\widetilde{\mathbf{z}}_g$, $r_j = r_{jj} + 2\sum_{k>j} r_{jk} p_j^{-1/2} p_k^{1/2}$, 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 (8) reduces to the system of Engel curves

$$\begin{aligned}\mathbf{q}_i = {} & \left(x_i - \beta - \boldsymbol{\alpha}'\overline{\mathbf{q}}_g - \widetilde{\boldsymbol{\gamma}}'\widetilde{\mathbf{z}}_i - \boldsymbol{\kappa}'\widetilde{\mathbf{z}}_g\right)^2 \mathbf{m} + \left(x_i - \beta - \boldsymbol{\alpha}'\overline{\mathbf{q}}_g - \widetilde{\boldsymbol{\gamma}}'\widetilde{\mathbf{z}}_i - \boldsymbol{\kappa}'\widetilde{\mathbf{z}}_g\right)\boldsymbol{\delta} \\ & + \mathbf{r} + \mathbf{A}\overline{\mathbf{q}}_{\mathbf{g}} + \widetilde{\mathbf{C}}\widetilde{\mathbf{z}}_i + \mathbf{D}\widetilde{\mathbf{z}}_g + \mathbf{v}_g + \mathbf{u}_i,\end{aligned} \tag{38}$$

This has a very similar structure to the generic multiple equation system of equations (34), and we proceed similarly.

Define $\widetilde{\mathbf{v}}_g = \left(\boldsymbol{\alpha}'\overline{\mathbf{q}}_g + \beta + \boldsymbol{\kappa}'\widetilde{\mathbf{z}}_g\right)^2 \mathbf{m} - \left(\boldsymbol{\alpha}'\overline{\mathbf{q}}_g + \beta + \boldsymbol{\kappa}'\widetilde{\mathbf{z}}_g\right)\boldsymbol{\delta} + \mathbf{r} + \mathbf{A}\overline{\mathbf{q}}_{\mathbf{g}} + \mathbf{D}\widetilde{\mathbf{z}}_{\mathbf{g}} + \mathbf{v}_g$. Then equation (38) can be rewritten more simply as

$$\mathbf{q}_i = \left(x_i - \widetilde{\boldsymbol{\gamma}}'\widetilde{\mathbf{z}}_i\right)^2 \mathbf{m} - 2\left(x_i - \widetilde{\boldsymbol{\gamma}}'\widetilde{\mathbf{z}}_i\right)\left(\boldsymbol{\alpha}'\overline{\mathbf{q}}_g + \beta + \boldsymbol{\kappa}'\widetilde{\mathbf{z}}_g\right)\mathbf{m} + \left(x_i - \widetilde{\boldsymbol{\gamma}}'\widetilde{\mathbf{z}}_i\right)\boldsymbol{\delta} + \widetilde{\mathbf{C}}\widetilde{\mathbf{z}}_i + \widetilde{\mathbf{v}}_g + \mathbf{u}_i, \tag{39}$$

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 (39) gives, for each good j,

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

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

$$\begin{aligned}0 = E\{ & [(q_{ji} - q_{ji'}) - \left(\left(x_i - \widetilde{\boldsymbol{\gamma}}'\widetilde{\mathbf{z}}_i\right)^2 - \left(x_{i'} - \widetilde{\boldsymbol{\gamma}}'\widetilde{\mathbf{z}}_{i'}\right)^2\right) m_j - \widetilde{\mathbf{c}}_j'(\widetilde{\mathbf{z}}_i - \widetilde{\mathbf{z}}_{i'}) \\ & - (\delta_j - 2m_j(\boldsymbol{\alpha}'\widehat{\mathbf{q}}_{g,-ii'} + \beta + \boldsymbol{\kappa}'\widetilde{\mathbf{z}}_g))\left(\left(x_i - \widetilde{\boldsymbol{\gamma}}'\widetilde{\mathbf{z}}_i\right) - \left(x_{i'} - \widetilde{\boldsymbol{\gamma}}'\widetilde{\mathbf{z}}_{i'}\right)\right)]\mathbf{r}_{gii'}\}.\end{aligned} \tag{40}$$

Combining common terms, we have

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

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

Assume we observe data from T different price regimes. In the main text, each group is observed only once, in a single price regime, so prices could just be subscripted by g. Here we allow for the greater generality of repeated cross section data, where groups could be observed more than once in different price regimes. To allow for this greater generality, we add a t subscript to prices. 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

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

from which we can construct

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

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

Hence,

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

Pairwise differencing gives

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

where the composite error is

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

Make the following assumptions.

Assumption B1: Each individual i in group g satisfies equation (38). Unobserved errors $\mathbf{u}_i$'s are independent across groups and have zero mean conditional on all $(x_l, \mathbf{z}_l)$ for $l \in g$, and $\mathbf{v}_g$ are unobserved group level fixed effects. The number of observed groups $G \to \infty$. For each observed group g, a sample of n_g observations of $\mathbf{q}_i, x_i, \mathbf{z}_i$ is observed. Each sample size n_g is fixed and does not go to infinity. The true number of individuals comprising each group is unknown.

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

$$\overline{\mathbf{q}}_g = \overline{x_g^2}\mathbf{m} + \left(\widetilde{\boldsymbol{\gamma}}'\overline{\mathbf{z}\mathbf{z}'}_g\widetilde{\boldsymbol{\gamma}}\right)\mathbf{m} + \mathbf{m}\left(\boldsymbol{\alpha}'\overline{\mathbf{q}}_g + \boldsymbol{\kappa}'\widetilde{\mathbf{z}}_g + \beta\right)^2 - 2\mathbf{m}\left(\boldsymbol{\alpha}'\overline{\mathbf{q}}_g + \boldsymbol{\kappa}'\widetilde{\mathbf{z}}_g + \beta\right)\left(\overline{x}_g - \widetilde{\boldsymbol{\gamma}}'\overline{\mathbf{z}}_g\right)$$
$$- 2\mathbf{m}\widetilde{\boldsymbol{\gamma}}'\overline{x\mathbf{z}}_g + \left(\overline{x}_g - \beta - \boldsymbol{\alpha}'\overline{\mathbf{q}}_g - \widetilde{\boldsymbol{\gamma}}'\overline{\mathbf{z}}_g - \boldsymbol{\kappa}'\widetilde{\mathbf{z}}_g\right)\boldsymbol{\delta} + \mathbf{r} + \mathbf{A}\overline{\mathbf{q}}_{\mathbf{g}} + \widetilde{\mathbf{C}}\overline{\mathbf{z}}_g + \mathbf{D}\widetilde{\mathbf{z}}_g + \mathbf{v}_g. \tag{42}$$

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

$$\overline{q}_g = \frac{1}{2m\,(Ap)^2}\{\left(2mAp(\overline{x}_g - \widetilde{\boldsymbol{\gamma}}'\overline{\mathbf{z}}_g - \boldsymbol{\kappa}'\widetilde{\mathbf{z}}_g - \beta) + 1 - A + pA\delta\right) \pm [(2mAp(\overline{x}_g - \widetilde{\boldsymbol{\gamma}}'\overline{\mathbf{z}}_g - \boldsymbol{\kappa}'\widetilde{\mathbf{z}}_g - \beta)$$
$$+ 1 - A + pA\delta)^2 - 4m\,(Ap)^2\left(m\overline{x_g^2} + m\boldsymbol{\gamma}'\overline{\mathbf{z}\mathbf{z}'}_g\boldsymbol{\gamma} + m(\boldsymbol{\kappa}'\widetilde{\mathbf{z}}_g + \beta)^2 - 2m(\boldsymbol{\kappa}'\widetilde{\mathbf{z}}_g + \beta)(\overline{x}_g - \widetilde{\boldsymbol{\gamma}}'\overline{\mathbf{z}}_g)\right.$$
$$\left. -2m\widetilde{\boldsymbol{\gamma}}'\overline{x\mathbf{z}}_g + \left(\overline{x}_g - \beta - \widetilde{\boldsymbol{\gamma}}'\overline{\mathbf{z}}_g - \boldsymbol{\kappa}'\widetilde{\mathbf{z}}_g\right)\delta + r + \widetilde{C}\overline{\mathbf{z}}_g + D\widetilde{\mathbf{z}}_g + v_g)\right)]^{1/2}\}, \tag{43}$$

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 (43), we can see $\overline{\mathbf{q}}_g$ is an implicit function of $\overline{x_g^2}, \overline{x}_g, \overline{\mathbf{z}}_g, \widetilde{\mathbf{z}}_g, \overline{\mathbf{z}\mathbf{z}'}_g, \overline{x\mathbf{z}}_g$, and $\mathbf{v}_g$. In the case of multiple equilibria, we do not take a stand on which root of equation (42) 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 $\Sigma_{x\mathbf{z}g} = \begin{pmatrix} \sigma^2_{xg} & \sigma_{x\mathbf{z}g} \\ \sigma'_{x\mathbf{z}g} & \Sigma_{\mathbf{z}g} \end{pmatrix}$ where $\sigma^2_{xg} = Var(x_i \mid i \in g)$, $\sigma_{x\mathbf{z}g} = Cov(x_i, \widetilde{\mathbf{z}}_i \mid i \in g)$ and $\Sigma_{\mathbf{z}g} = Var(\widetilde{\mathbf{z}}_i \mid i \in g)$. Denote $\varepsilon_{ix} = x_i - \overline{x}_g$ and $\boldsymbol{\varepsilon}_{iz} = \widetilde{\mathbf{z}}_i - \overline{\mathbf{z}}_g$. Assume $E\left((\varepsilon_{ix}, \boldsymbol{\varepsilon}_{iz})|\overline{\mathbf{z}}_g, \widetilde{\mathbf{z}}_g, \overline{x\mathbf{z}}_g, \overline{\mathbf{z}\mathbf{z}'}_g, \overline{x}_g, \overline{x_g^2}, \mathbf{v}_g, \mathbf{r}_g\right) = 0$ and is independent across individual i's.

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

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

Then we have the moment condition

$$0 = E\{[\mathbf{q}_i - \mathbf{q}_{i'} + 2\mathbf{m}\left(\boldsymbol{\alpha}'\widehat{\mathbf{q}}_{g,-ii'} + \boldsymbol{\kappa}'\widetilde{\mathbf{z}}_g\right)[(x_i - x_{i'}) - \widetilde{\boldsymbol{\gamma}}'(\widetilde{\mathbf{z}}_i - \widetilde{\mathbf{z}}_{i'})] - (x_i^2 - x_{i'}^2)\mathbf{m} \tag{44}$$
$$-\widetilde{\boldsymbol{\gamma}}'\left(\widetilde{\mathbf{z}}_i\widetilde{\mathbf{z}}_i' - \widetilde{\mathbf{z}}_{i'}\widetilde{\mathbf{z}}_{i'}'\right)\widetilde{\boldsymbol{\gamma}}\mathbf{m} + 2\mathbf{m}\widetilde{\boldsymbol{\gamma}}'(\widetilde{\mathbf{z}}_i x_i - \widetilde{\mathbf{z}}_{i'}x_{i'}) - \boldsymbol{\eta}(x_i - x_{i'}) + (\boldsymbol{\eta}\widetilde{\boldsymbol{\gamma}}' - \widetilde{\mathbf{C}})(\widetilde{\mathbf{z}}_i - \widetilde{\mathbf{z}}_{i'})]|x_i, x_{i'}, \mathbf{z}_i, \mathbf{z}_{i'}, \mathbf{r}_g\}$$

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

$$E\left[\left(\mathbf{q}_i - \mathbf{q}_{i'} + 2e^{-\mathbf{b}'\ln \mathbf{p}_t}\frac{\mathbf{d}}{\mathbf{p}_t}\left(\mathbf{p}_t'\mathbf{A}\widehat{\mathbf{q}}_{gt,-ii'} + \mathbf{p}_t'\mathbf{D}\widetilde{\mathbf{z}}_g\right)\left[(x_i - x_{i'}) - \mathbf{p}_t'\widetilde{\mathbf{C}}(\widetilde{\mathbf{z}}_i - \widetilde{\mathbf{z}}_{i'})\right] - e^{-\mathbf{b}'\ln \mathbf{p}_t}\frac{\mathbf{d}}{\mathbf{p}_t}\right.\right.$$
$$\left[(x_i^2 - x_{i'}^2) + \mathbf{p}_t'\widetilde{\mathbf{C}}\left(\widetilde{\mathbf{z}}_i\widetilde{\mathbf{z}}_i' - \widetilde{\mathbf{z}}_{i'}\widetilde{\mathbf{z}}_{i'}'\right)\widetilde{\mathbf{C}}'\mathbf{p}_t - 2\mathbf{p}_t'\widetilde{\mathbf{C}}(\mathbf{z}_i x_i - \mathbf{z}_{i'}x_{i'})\right] - \left(\frac{\mathbf{b}}{\mathbf{p}_t} - 2e^{-\mathbf{b}'\ln \mathbf{p}_t}\frac{\mathbf{d}}{\mathbf{p}_t}\mathbf{p}_t^{1/2\prime}\mathbf{R}\mathbf{p}_t^{1/2}\right)$$
$$\left.\left.\cdot(x_i - x_{i'}) + \left[\left(\frac{\mathbf{b}}{\mathbf{p}_t} - 2e^{-\mathbf{b}'\ln \mathbf{p}_t}\frac{\mathbf{d}}{\mathbf{p}_t}\mathbf{p}_t^{1/2\prime}\mathbf{R}\mathbf{p}_t^{1/2}\right)\widetilde{\mathbf{C}}'\mathbf{p}_t - \widetilde{\mathbf{C}}\right](\widetilde{\mathbf{z}}_i - \widetilde{\mathbf{z}}_{i'})\right| x_i, x_{i'}, \mathbf{z}_i, \mathbf{z}_{i'}, \mathbf{r}_g\right] = 0. \quad (45)$$

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

$$\begin{aligned} &L_{1jgii'} = (q_{ji} - q_{ji'}), \;\; L_{2jgii'} = \widehat{q}_{jg,-ii'}(x_i - x_{i'}), \;\; L_{3jkgii'} = \widehat{q}_{jgt,-ii'}(\widetilde{z}_{ki} - \widetilde{z}_{ki'}), \\ &L_{4k_2gii'} = \widetilde{z}_{k_2g}(x_i - x_{i'}), \;\; L_{5kk_2gii'} = \widetilde{z}_{k_2g}(\widetilde{z}_{ki} - \widetilde{z}_{ki'}), \;\; L_{6gii'} = x_i^2 - x_{i'}^2, \\ &L_{7kk'gii'} = \widetilde{z}_{ki}\widetilde{z}_{k'i} - \widetilde{z}_{ki'}\widetilde{z}_{k'i'}, \;\; L_{8kgii'} = \widetilde{z}_{ki}x_i - \widetilde{z}_{ki'}x_{i'}, \;\; L_{9gii'} = x_i - x_{i'}, \;\; L_{10kgii'} = \widetilde{z}_{ki} - \widetilde{z}_{ki'}, \end{aligned} \quad (46)$$

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

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

Then for each good j, the identification is based on

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

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

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

Under Assumption B5, we can identify

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

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

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

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

$$\mathbf{PA} = (\boldsymbol{\alpha}_1, ..., \boldsymbol{\alpha}_T)', \ \mathbf{PD} = (\boldsymbol{\kappa}_1, ..., \boldsymbol{\kappa}_T)', \text{ and}$$

$$\boldsymbol{\Lambda}\left(b_1, ...b_{J-1}, r_{11}, ..., r_{JJ}, r_{12}, ..., r_{J-1,J}\right)' = \begin{pmatrix} \boldsymbol{\eta}_1 \\ \vdots \\ \boldsymbol{\eta}_T \end{pmatrix},$$

where $\boldsymbol{\eta}_t = (\eta_{1t}, ..., \eta_{J-1,t})'$. 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 $\boldsymbol{\Lambda}$ 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

$$\mathbf{A} = (\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}'(\boldsymbol{\alpha}_1, ..., \boldsymbol{\alpha}_T)' \text{ and } \mathbf{D} = (\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}'(\boldsymbol{\kappa}_1, ..., \boldsymbol{\kappa}_T)';$$

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

$$(b_1, ...b_{J-1}, r_{11}, ..., r_{JJ}, r_{12}, ..., r_{J-1,J})' = (\mathbf{\Lambda}'\mathbf{\Lambda})^{-1}\mathbf{\Lambda}'(\boldsymbol{\eta}_1', ..., \boldsymbol{\eta}_T')';$$

$\mathbf{d}$ 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 $\widetilde{\mathbf{C}}$, $\boldsymbol{\alpha}$, $\widetilde{\boldsymbol{\gamma}}$, $\boldsymbol{\kappa}$, $\mathbf{m}$, and $\boldsymbol{\eta} = \boldsymbol{\delta} - 2\mathbf{m}\beta$ in the Engel curve system (38) 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 (8) 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

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

with

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

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

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

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

because

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

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

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

Therefore,

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

By rewriting q_{ji} as

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

where

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

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

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

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

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

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

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

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

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

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

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

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

for each $j = 1, ..., J-1$. From this, $\widetilde{\boldsymbol{\gamma}}$, $\boldsymbol{\kappa}$, $\boldsymbol{\alpha}$, $\mathbf{m}$, $\boldsymbol{\eta} = \boldsymbol{\delta} - 2\mathbf{m}\beta$, $\mathbf{A}_j$, $\widetilde{\mathbf{c}}_j$, $\mathbf{D}_j$, and $m_j\beta^2 - \delta_j\beta + r_j + v_{j0}$ for $j = 1, ..., J-1$ are all identified. Then, $\mathbf{A}_J = \left(\boldsymbol{\alpha} - \sum_{j=1}^{J-1} \mathbf{A}_j p_j\right)/p_J$, $\widetilde{\mathbf{c}}_J = (\widetilde{\boldsymbol{\gamma}} - \sum_{j=1}^{J-1} \widetilde{\mathbf{c}}_j p_j)/p_J$, 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 $\boldsymbol{\alpha}$, β, $\widetilde{\boldsymbol{\gamma}}$, $\boldsymbol{\kappa}$, $\boldsymbol{\delta}$, r_j, and $\mathbf{m}$ with their corresponding full demand system expressions, i.e., $\boldsymbol{\alpha} = \mathbf{A}'\mathbf{p}$, $\beta = \mathbf{p}^{1/2\prime}\mathbf{R}\mathbf{p}^{1/2}$, 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

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

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

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

with

$$
\begin{aligned}
m_{jt} &= e^{-\mathbf{b}'\ln\mathbf{p}_t}\frac{d_j}{p_{jt}},\ \boldsymbol{\alpha}_t = \mathbf{A}'\mathbf{p}_t,\ \widetilde{\boldsymbol{\gamma}}_t = \widetilde{\mathbf{C}}'\mathbf{p}_t,\ \boldsymbol{\kappa}_t = \mathbf{D}'\mathbf{p}_t,\ \beta_t = \mathbf{p}_t^{1/2\prime}\mathbf{R}\mathbf{p}_t^{1/2}, \\
\eta_{jt} &= \frac{b_j}{p_{jt}} - 2m_{jt}\mathbf{p}_t^{1/2\prime}\mathbf{R}\mathbf{p}_t^{1/2},\ \delta_{jt} = \frac{b_j}{p_{jt}},\ r_{jt} = R_{jj} + 2\sum_{k>j} R_{jk}\sqrt{p_{kt}/p_{jt}}, \\
v_{jt0} &= \mu_{jt} - e^{-\mathbf{b}'\ln\mathbf{p}_t}\frac{d_j}{p_{jt}}\sum_{j_1=1}^{J}\sum_{j_2=1}^{J}\sum_{j=1}^{J}\sum_{j'=1}^{J} A_{j_1 j}p_{j_1 t}A_{j_2 j'}p_{j_2 t}\boldsymbol{\Sigma}_{vt,jj'}.
\end{aligned}
$$

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

$\boldsymbol{\mu} = E(\mathbf{v}_{gt})$ 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-derived demand functions, it may be reasonable to assume that $\boldsymbol{\mu} = 0$. With these substitutions we only need to estimate the parameters $\mathbf{\Sigma}_v$ instead of all the separate v_{jt0} constants.

Appendices for online publication

Appendix B: Preliminary Data Analyses

B.1 Generic Model Estimates

In other, non-demand settings, the generic peer effects model of Section III may be more appropriate than the structural demand model. We implemented this model in Section IV.B, but in this section describe the results in more detail.

As in the presentation in (12), y_i is expenditures on food, $\overline{y}_g$ is the true group-mean expenditure on food, $\widehat{y}_g$ is the observed sample average, and x_i is total expenditures.

We provide estimates using random-effects unconditional moments (21) and fixed-effects unconditional moments (18). 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 3.

In the RE model, higher levels of peer food expenditure work in the same direction as own expenditure; in effect making the household behave (in a demand sense) as if it was richer when peer expenditures rise. Since this is not sustainable in equilibrium, it is reassuring that in the FE specification, higher peer expenditure makes households reduce their demand for food.

This difference between the models is a natural consequence of the group-level unobservable taste for an expenditure category v_g being correlated with expenditure in that category. Unsurprisingly, the Hausman tests decisively reject the RE specification.

However, the peer effects in the FE specification are very large. Variation in peer expenditure has over twice the effect of own expenditure on demand behavior (see the estimates of $-a/b$), but we cannot reject equivalence of the two effects given the imprecision of the peer effect estimates. This is a potential consequence of excluding group-average non-food spending from the right hand side. We take this as a reason to focus on the structural estimates, which restrict behavior (including price responses) in a way consistent with economic theory.

In both models, the estimated values of b is positive, and d is negative. As a result, food budget shares are declining with expenditure, consistent with Engel's Law..

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 food, 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.

Table 1 summarizes the 4$^{\text{th}}$ (2001), 5$^{\text{th}}$ (2006), and 6$^{\text{th}}$ (2014) waves of the World Values Survey. 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$^{\text{th}}$ wave, so we collapse all waves 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. The exact cutpoints are undocumented, so we collapse the scale to five points for interpretability and use dummies for the income groupings directly in our analysis.

For this analysis we define groups by religion (Hindu vs non-Hindu) 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. This is for two reasons: first, we do not observe caste or geographic indicators smaller than states; and second, larger groups are needed here because the WVS sample size is much smaller than the NSS and we have no asymptotic theory to deal with small group sizes in this part of the analysis.

Table 2 presents estimates of well-being as a function of both own total expenditures and group total expenditures, specified as

$$U_i = \sum_{s=2}^{5} \beta_g 1[I_i = s] + \pi \overline{\widehat{x}}_{gt} + X_{igt}\alpha + \gamma_g + \phi_t + \varepsilon_{igt}, \tag{47}$$

where U_i is the z-normalized well-being indicator, $1[g_i = s]$ is an indicator for individual i belonging to income group s, $\overline{\widehat{x}}_{gt}$ is imputed group expenditures, X_{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 π 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 Table 2 imply that satisfaction is increasing over the entire range of individual expenditures, but that a 100 rupee increase in peer expenditure $\overline{\overline{x}}_{gt}$ *decreases* satisfaction by 0.16 standard deviations. Other specifications in Table 2 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.

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 2). The results are qualitatively the same, suggesting that our results are not being determined by the normalizations implicit in z-scoring the satisfaction responses. We conclude that welfare is indeed increasing in household expenditure and decreasing in peer expenditure.

Finally, we include an interaction term (the product of peer expenditures and the individual being in the top two income groups) 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.

Table 1: Subjective well-being summary statistics

	Mean	SD	Min	Max
Income group 2 (=1)	.4	.49	0	1
Income group 3 (=1)	.21	.4	0	1
Income group 4 (=1)	.087	.28	0	1
Income group 5 (=1)	.041	.2	0	1
Group expenditure (1000 rupees)	5.6	2.6	2.8	18
Age	.34	.12	.15	.77
Sex	1.4	.49	1	2
Household size	.32	.19	0	.9
Married (=1)	.84	.36	0	1
Primary education (=1)	.095	.29	0	1
Secondary education (=1)	.13	.34	0	1
Observations	5,084			

Life satisfaction variable from World Values Survey. Participants were 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 2001 and 2006, and 1-10 in 2014. We collapsed to a 1-5 scale in 2014. Group income measured in thousands of Rs/month.

Table 2: Satisfaction on household and peer income

	OLS (SDs)			Ordered logit		
	(1)	(2)	(3)	(4)	(5)	(6)
Income group 2 (=1)	0.14**	0.12**	0.12**	0.33***	0.30**	0.30**
	(0.06)	(0.06)	(0.06)	(0.11)	(0.12)	(0.12)
Income group 3 (=1)	0.36***	0.33***	0.33***	0.80***	0.74***	0.75***
	(0.07)	(0.08)	(0.08)	(0.15)	(0.15)	(0.15)
Income group 4 (=1)	0.40***	0.39***	0.21	0.95***	0.93***	0.47
	(0.10)	(0.10)	(0.19)	(0.23)	(0.23)	(0.42)
Income group 5 (=1)	0.52***	0.51***	0.33*	1.19***	1.17***	0.71
	(0.17)	(0.17)	(0.19)	(0.42)	(0.40)	(0.45)
Group expenditure (1000 rupees)	-0.15**	-0.15**	-0.16**	-0.35**	-0.34*	-0.37**
	(0.07)	(0.07)	(0.07)	(0.17)	(0.18)	(0.18)
Group expend X top 2 quintiles			0.03			0.07
			(0.03)			(0.06)
Controls	No	Yes	Yes	No	Yes	Yes
Observations	5,084	5,084	5,084	5,084	5,084	5,084

Dependent variable as noted in column header, in SD. Subjective well being data from World Values Survey, imputated group income from NSS. Peer groups defined as intersection of state and religion (Hindu and non-Hindu). Controls include household size, age, sex, marital status and education. All columns include year fixed effects. Standard errors in parentheses and clustered at the group level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

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

	RE		FE	
	(1)	(2)	(3)	(4)
a (peer mean expenditure)	0.179***	0.131***	-1.076***	-1.077**
	(0.042)	(0.046)	(0.392)	(0.442)
b (own expenditure)	0.410***	0.415***	0.466***	0.456***
	(0.009)	(0.011)	(0.015)	(0.018)
d (curvature)	-0.177***	-0.182***	-0.092***	-0.067***
	(0.008)	(0.010)	(0.013)	(0.012)
$-a/b$	-0.437	-0.315	2.309	2.361
	(0.108)	(0.115)	(0.847)	(0.975)
p for $-a/b = 1$	0.000	0.000	0.122	0.163
Hausman for a			10.374	7.536
P-value			0.001	0.006
Individual controls	No	Yes	No	Yes
Number of groups	568	564	568	564
Number of pairs	221,642	128,640	221,642	128,640

Dependent variable is household food spending. Individual controls include household size, age, marital status and amount of land owned. All models include price controls. Standard errors in parentheses and clustered at the group level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.