# Supplement to "Consumption peer effects and utility needs in India"

(*Quantitative Economics*, Vol. 13, No. 3, July 2022, 1257–1295)

Arthur Lewbel
Department of Economics, Boston College

Samuel Norris
Department of Economics, University of British Columbia

Krishna Pendakur
Department of Economics, Simon Fraser University

Xi Qu
Antai College of Economics and Management, Shanghai Jiao Tong University

## 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. More precisely, what we are assuming is that there is an underlying distribution of the (infinite) population of potential group members. Everyone who is actually in the group in the real world population is a draw from this underlying potential population. Each individual in the group knows the true mean of this distribution that individuals are drawn from, and bases their behavior on that true mean. This model implies that the individual's own choice has zero effect on the group mean.

If group members observe each other's private information and make decisions simultaneously (corresponding to a complete information game), then we assume that each individual's actual behavior will only depend on others through the group mean. Estimation of complete games typically depends 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 behavior 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

Arthur Lewbel: lewbel@bc.edu
Samuel Norris: sam.norris@ubc.ca
Krishna Pendakur: pendakur@sfu.ca
Xi Qu: xiqu@sjtu.edu.cn

 https://doi.org/10.3982/QE1760

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 the calculating group means are negligible.[1]

### A.2 *Generic model identification and estimation with fixed effects*

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

$$y_i = \left(\overline{y}_g a + \mathbf{x}'_i\mathbf{b}\right)^2 d + \left(\overline{y}_g a + \mathbf{x}'_i\mathbf{b}\right) + v_g + u_i, \tag{A1}$$

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

We could have written the seemingly more general model

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

where $c$ and $k$ are additional constants to be estimated. However, 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 (A1), 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 $\widetilde{d} = d/(2cd + k)^2$. If $2cd + k = 0$, then by letting $\widetilde{v}_g = c^2 d + ck + v_g$, this equation becomes $y_i = (\overline{y}_g a + \mathbf{x}'_i\mathbf{b})^2 d + \widetilde{v}_g + u_i$. Since this pure quadratic form equation is strictly easier to identify and estimate, and is

[1] 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.

irrelevant for our empirical application, we will rule it out and, therefore, without loss of generality replace the more general model with equation (A1).

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 the 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 A.1 to A.5 below.

ASSUMPTION A.1. *Each individual $i$ in group $g$ satisfies equation (A1). $\mathbf{x}_i$ is a $K$-dimensional vector of covariates. For each $k \in \{1, \ldots, K\}$, 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 A.1 essentially defines the model. Note that Assumption A.1 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 A.2. *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 A.2 $d \neq 0$ 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 A.2 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 (A1) for $y_i$ into $\overline{y}_g = E(y_i \mid i \in g)$, we have

$$y_i = \overline{y}_g^2 da^2 + a\big(2d\mathbf{x}_i'\mathbf{b} + 1\big)\overline{y}_g + \mathbf{b}'\mathbf{x}_i\mathbf{x}_i'\mathbf{b}d + \mathbf{x}_i'\mathbf{b} + v_g + u_i \tag{A2}$$

Taking the within group expected value of this expression gives

$$\overline{y}_g = \overline{y}_g^2 da^2 + a\big(2d\mathbf{b}'\overline{\mathbf{x}}_g + 1\big)\overline{y}_g + d\mathbf{b}'\overline{\mathbf{x}\mathbf{x}'_g}\mathbf{b} + \mathbf{b}'\overline{\mathbf{x}}_g + v_g. \tag{A3}$$

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 (A3) is a quadratic with roots

$$\overline{y}_g = \frac{1 - a\big(2\mathbf{b}'\overline{\mathbf{x}}_g d + 1\big) \pm \sqrt{\big[1 - a\big(2\mathbf{b}'\overline{\mathbf{x}}_g d + 1\big)\big]^2 - 4a^2 d\big[d\mathbf{b}'\overline{\mathbf{x}\mathbf{x}'_g}\mathbf{b} + \mathbf{b}'\overline{\mathbf{x}}_g + v_g\big]}}{2a^2 d}. \tag{A4}$$

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 A.2 is satisfied, 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 are not any peer effects. We do not take a stand on which root of equation (A4) is chosen by consumers, and we just make the following assumption.

ASSUMPTION A.3. *Individuals within each group agree on an equilibrium selection rule.*

The equilibrium of $\overline{y}_g$ therefore exists under Assumption A.2 and is unique under Assumption A.3.

For identification, we need to remove the fixed effect from equation (A1), which we do by subtracting off another individual in the same group. For each $(i, i') \in g$, consider pairwise difference

$$\begin{aligned} y_i - y_{i'} &= 2ad\overline{y}_g\mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'}) + d\mathbf{b}'\big(\mathbf{x}_i\mathbf{x}'_i - \mathbf{x}_{i'}\mathbf{x}'_{i'}\big)\mathbf{b} + \mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'}) + u_i - u_{i'} \\ &= 2ad\widehat{y}_{g,-ii'}\mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'}) + d\mathbf{b}'\big(\mathbf{x}_i\mathbf{x}'_i - \mathbf{x}_{i'}\mathbf{x}'_{i'}\big)\mathbf{b} + \mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'}) \\ &\quad + u_i - u_{i'} - 2ad\varepsilon_{yg,-ii'}\mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'}), \end{aligned} \tag{A5}$$

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 (A5) 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 A.1, $u_i - u_{i'}$ is conditionally mean independent of $\mathbf{x}_i$ and $\mathbf{x}_{i'}$. It can also be shown that

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

where

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

Substituting this expression into equation (A5) 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 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 A.4. *Let $\mathbf{r}_g$ be a vector (possibly empty) of observed group level instruments that are independent of each $u_i$. Assume $E((\mathbf{x}_i - \overline{\mathbf{x}}_g) \mid i \in g, \overline{\mathbf{x}}_g, \overline{\mathbf{xx}'}_g, v_g, \mathbf{r}_g) = 0$, $E((\mathbf{x}_i\mathbf{x}_i' - \overline{\mathbf{xx}'}_g) \mid i \in g, \mathbf{r}_g) = 0$, 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 A.4 corresponds to (but is a little stronger than) standard instrument validity assumptions. A sufficient condition for the equalities in Assumption A.4 to hold is to let $\boldsymbol{\varepsilon}_{ix} = \mathbf{x}_i - \overline{\mathbf{x}}_g$ be independent across individuals, and assume that $E(\boldsymbol{\varepsilon}_{ix} \mid \overline{\mathbf{x}}_g, \overline{\mathbf{xx}'}_g, v_g, \mathbf{r}_g \text{ for } i \in g) = 0$ and $E(\boldsymbol{\varepsilon}_{ix}\boldsymbol{\varepsilon}_{ix}' \mid \overline{\mathbf{x}}_g, \mathbf{r}_g \text{ for } i \in g) = E(\boldsymbol{\varepsilon}_{ix}\boldsymbol{\varepsilon}_{ix}' \mid i \in g)$. To see this, we have

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

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 A.4 are used to show that $E(\varepsilon_{yg,-ii'} \mid \mathbf{r}_g) = 0$, and the independence of measurement errors across individuals is used to show $E(\varepsilon_{yg,-ii'}(\mathbf{x}_i - \mathbf{x}_{i'}) \mid \mathbf{r}_g, \mathbf{x}_i, \mathbf{x}_{i'}) = (\mathbf{x}_i - \mathbf{x}_{i'})E(\varepsilon_{yg,-ii'} \mid \mathbf{r}_g) = 0$, so that $\mathbf{x}_i$ and $\mathbf{x}_{i'}$ are valid instruments. Given Assumptions A.1 and A.4, one can directly verify that

$$
\begin{aligned}
E\big[y_i - y_{i'} - \big(2ad\widehat{y}_{g,-ii'}\mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'}) + d\mathbf{b}'\big(\mathbf{x}_i\mathbf{x}_i' - \mathbf{x}_{i'}\mathbf{x}_{i'}'\big)\mathbf{b} \\
+ \mathbf{b}'(\mathbf{x}_i - \mathbf{x}_{i'})\big) \mid \mathbf{r}_g, \mathbf{x}_i, \mathbf{x}_{i'}\big] = 0.
\end{aligned} \tag{A6}
$$

Under Assumptions A.1 to A.4, $(\mathbf{x}_i - \mathbf{x}_{i'})E(\widehat{y}_{g,-ii'}|\mathbf{r}_g, \mathbf{x}_i, \mathbf{x}_{i'})$ 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 $2ad\mathbf{b}$, $b_1 d\mathbf{b}, \ldots, b_K d\mathbf{b}$, 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 (A4) 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 $\mathbf{x}_i = \boldsymbol{\epsilon}_{ix} + \overline{\mathbf{x}}_g$.

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 (A6) that

$$E\left[\left(y_i - y_{i'} - (1 + 2ad\widehat{y}_{g,-ii'})\sum_{k=1}^{K} b_k(x_{ki} - x_{ki'}) \right.\right.$$
$$\left.\left. - d\sum_{k=1}^{K}\sum_{k'=1}^{K} b_k b_{k'}(x_{ki}x_{k'i} - x_{ki'}x_{k'i'})\right)\mathbf{r}_{gii'}\right] = 0. \tag{A7}$$

Let

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

Equation (A7) 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

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

So, $\Gamma_g$ is the set of all observed pairs of individuals $i$ and $i'$ in the group $g$. For $\ell \in \{1, 2k, 3k, 4kk' \mid k, k' = 1, \ldots, K\}$, 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 (A7) over all $(i, i') \in \Gamma_g$ gives the unconditional group level moment vector

$$E\left(\mathbf{Y}_{1g} - \sum_{k=1}^{K} b_k \mathbf{Y}_{2kg} - 2ad\sum_{k=1}^{K} b_k \mathbf{Y}_{3kg} - d\sum_{k=1}^{K}\sum_{k'=1}^{K} b_k b_{k'} \mathbf{Y}_{4kk'g}\right) = 0. \tag{A8}$$

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

ASSUMPTION A.5. *$E(\mathbf{Y}_g')E(\mathbf{Y}_g)$ is nonsingular.*

THEOREM 1. *Given Assumptions A.1–A.5, the coefficients $a$, $\mathbf{b}$, $d$ are identified from*

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

*As noted earlier, Assumptions A.1 to A.4 should generally suffice for identification. Assumption A.5 is used to obtain more convenient identification based on unconditional moments. Assumption A.5 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 (A8).*

Given the identification above, based on equation (A8) we can immediately construct a corresponding group level GMM estimator

$$\begin{aligned}(\widehat{a}, \widehat{b}_1, \ldots, \widehat{b}_K, \widehat{d}) &= \arg\min\left[\frac{1}{G}\sum_{g=1}^{G}\left(\mathbf{Y}_{1g} - \sum_{k=1}^{K} b_k\mathbf{Y}_{2kg} - 2ad\sum_{k=1}^{K} b_k\mathbf{Y}_{3kg} - d\sum_{k=1}^{K}\sum_{k'=1}^{K} b_k b_{k'}\mathbf{Y}_{4kk'g}\right)\right]' \\ &\quad \times \widehat{\Omega}\left[\frac{1}{G}\sum_{g=1}^{G}\left(\mathbf{Y}_{1g} - \sum_{k=1}^{K} b_k\mathbf{Y}_{2kg} - 2ad\sum_{k=1}^{K} b_k\mathbf{Y}_{3kg} - d\sum_{k=1}^{K}\sum_{k'=1}^{K} b_k b_{k'}\mathbf{Y}_{4kk'g}\right)\right] \qquad \text{(A9)}\end{aligned}$$

for some positive definite moment weighting matrix $\widehat{\Omega}$. In equation (A9), 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 (A7), 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,

$$(\widehat{a}, \widehat{b}_1, \ldots, \widehat{b}_K, \widehat{d}) = \arg\min\left(\frac{\sum_{g=1}^{G}\sum_{(i,i')\in\Gamma_g}\mathbf{m}_{gii'}}{\sum_{g=1}^{G}\sum_{(i,i')\in\Gamma_g}1}\right)'\widehat{\Omega}\left(\frac{\sum_{g=1}^{G}\sum_{(i,i')\in\Gamma_g}\mathbf{m}_{gii'}}{\sum_{g=1}^{G}\sum_{(i,i')\in\Gamma_g}1}\right), \qquad \text{(A10)}$$

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 (A7), $\mathbf{r}_{gii'}$ 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 (A4) 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, for example, $x_{ki}(x_{ki} - x_{ki'})$.

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 A.4. 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 A.4. The relevance of these instruments (the nonsingularity condition in Assumption A.5) 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 A.4 and A.5 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 (A4) 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 $\mathbf{y}_i = (y_{1i}, \ldots, y_{Ji})$ 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 multiequation that for each good $j$,

$$y_{ji} = \left(\overline{\mathbf{y}}_g'\mathbf{a}_j + \mathbf{x}_i'\mathbf{b}_j\right)^2 d_j + \left(\overline{\mathbf{y}}_g'\mathbf{a}_j + \mathbf{x}_i'\mathbf{b}_j\right) + v_{jg} + u_{ji}, \tag{A11}$$

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

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

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

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

we have the equilibrium

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

and the leave-two-out group average

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

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,

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

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

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

Denote

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

For $\ell \in \{1j, 2k, 3jk, 4kk' \mid j = 1, \ldots, J; k, k' = 1, \ldots, K\}$, 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 $k$th element of $\mathbf{b}_j$ and $a_{jj'}$ is the $j'$th element of $\mathbf{a}_j$.

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

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

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

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

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

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

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

where

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

## A.4 *Multiple equation generic model with random effects*

Here, we provide the derivation of equation (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 (A1).

$$y_i = \overline{y}_g^2 a^2 d + a\big(1 + 2\mathbf{b}'\mathbf{x}_i d\big)\overline{y}_g + \mathbf{b}'\mathbf{x}_i + \mathbf{b}'\mathbf{x}_i\mathbf{x}_i'\mathbf{b}d + v_g + u_i, \tag{A12}$$

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

$$\overline{y}_g = \overline{y}_g^2 da^2 + a\big(2d\mathbf{b}'\overline{\mathbf{x}}_g + 1\big)\overline{y}_g + d\mathbf{b}'\overline{\mathbf{x}\mathbf{x}_g'}\mathbf{b} + \mathbf{b}'\overline{\mathbf{x}}_g + \mu, \tag{A13}$$

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

$$y_i = \widehat{y}_{g,-ii'}y_{i'}a^2 d + a\big(1 + 2\mathbf{b}'\mathbf{x}_i d\big)\widehat{y}_{g,-ii'} + \mathbf{b}'\mathbf{x}_i + \mathbf{b}'\mathbf{x}_i\mathbf{x}_i'\mathbf{b}d + v_g + u_i + \widetilde{\varepsilon}_{gii'}, \tag{A14}$$

where

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

Formally, we make the following assumptions.

ASSUMPTION A.6. *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 A.7. *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 A.8. *For each group $g$, $v_g$ is independent across groups with $E(v_g|\mathbf{x}, \overline{\mathbf{x}}_g, \overline{\mathbf{x}\mathbf{x}'_g}) = \mu$ and we have the conditional homoskedasticity that* $\mathrm{Var}(v_g|\mathbf{x}, \overline{\mathbf{x}}_g, \overline{\mathbf{x}\mathbf{x}'_g}) = \sigma^2$.

Let $v_0 = \mu - da^2\sigma^2$. It follows from Assumptions A.6–A.8 that, for any $l \neq i$, $E(\overline{y}_g \varepsilon_{yl}|\mathbf{x}_i, \overline{\mathbf{x}}_g, \overline{\mathbf{x}\mathbf{x}'_g}) = 0$ and $E(\varepsilon_{yl}\mathbf{x}_i|\mathbf{x}_i, \overline{\mathbf{x}}_g, \overline{\mathbf{x}\mathbf{x}'_g}) = 0$. Hence, $E(\widetilde{\varepsilon}_{gii'}|x_i, \overline{\mathbf{x}}_g, \overline{\mathbf{x}\mathbf{x}'_g}) = -da^2 E(\varepsilon_{yg,-ii'}\varepsilon_{y,i'}|\mathbf{x}_i, \overline{\mathbf{x}}_g, \overline{\mathbf{x}\mathbf{x}'_g}) = -da^2\,\mathrm{Var}(v_g)$ and

$$E\big(v_g + u_i + \widetilde{\varepsilon}_{gii'} \mid \overline{\mathbf{x}}_g, \overline{\mathbf{x}\mathbf{x}'_g}, \mathbf{x}_i\big) = \mu - da^2\sigma^2 = v_0. \tag{A15}$$

By construction $v_g + u_i + \widetilde{\varepsilon}_{gii'}$ is also independent of $\mathbf{r}_g$. Given this, equation (20) then follows from equations (A14) and (A15).

### 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 $\mathbf{C} = (\widetilde{\mathbf{C}} : \mathbf{D})$ 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 $\mathbf{m} = (e^{-\mathbf{b}'\ln\mathbf{p}})\mathbf{d}/\mathbf{p}$ with constraints of $\mathbf{b}'\mathbf{1} = 1$ and $\mathbf{d}'\mathbf{1} = 0$. Then equation (9) reduces to the system of Engel curves

$$\begin{aligned}\mathbf{q}_i = {} & \big(x_i - \beta - \boldsymbol{\alpha}'\overline{\mathbf{q}}_g - \widetilde{\boldsymbol{\gamma}}'\widetilde{\mathbf{z}}_i - \boldsymbol{\kappa}'\widetilde{\mathbf{z}}_g\big)^2\mathbf{m} + \big(x_i - \beta - \boldsymbol{\alpha}'\overline{\mathbf{q}}_g - \widetilde{\boldsymbol{\gamma}}'\widetilde{\mathbf{z}}_i - \boldsymbol{\kappa}'\widetilde{\mathbf{z}}_g\big)\boldsymbol{\delta} \\ & + \mathbf{r} + \mathbf{A}\overline{\mathbf{q}}_g + \widetilde{\mathbf{C}}\widetilde{\mathbf{z}}_i + \mathbf{D}\widetilde{\mathbf{z}}_g + \mathbf{v}_g + \mathbf{u}_i, \end{aligned} \tag{A16}$$

This has a very similar structure to the generic multiple equation system of equations (A11), and we proceed similarly.

Define $\widetilde{\mathbf{v}}_g = (\boldsymbol{\alpha}'\overline{\mathbf{q}}_g + \beta + \boldsymbol{\kappa}'\widetilde{\mathbf{z}}_g)^2\mathbf{m} - (\boldsymbol{\alpha}'\overline{\mathbf{q}}_g + \beta + \boldsymbol{\kappa}'\widetilde{\mathbf{z}}_g)\boldsymbol{\delta} + \mathbf{r} + \mathbf{A}\overline{\mathbf{q}}_g + \mathbf{D}\widetilde{\mathbf{z}}_g + \mathbf{v}_g$. Then equation (A16) can be rewritten more simply as

$$\begin{aligned}\mathbf{q}_i = {} & \big(x_i - \widetilde{\boldsymbol{\gamma}}'\widetilde{\mathbf{z}}_i\big)^2\mathbf{m} - 2\big(x_i - \widetilde{\boldsymbol{\gamma}}'\widetilde{\mathbf{z}}_i\big)\big(\boldsymbol{\alpha}'\overline{\mathbf{q}}_g + \beta + \boldsymbol{\kappa}'\widetilde{\mathbf{z}}_g\big)\mathbf{m} + \big(x_i - \widetilde{\boldsymbol{\gamma}}'\widetilde{\mathbf{z}}_i\big)\boldsymbol{\delta} \\ & + \widetilde{\mathbf{C}}\widetilde{\mathbf{z}}_i + \widetilde{\mathbf{v}}_g + \mathbf{u}_i, \end{aligned} \tag{A17}$$

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 $j \in \{1, \ldots, J\}$ and each pair of individuals $i$ and $i'$ in group g. This pairwise differencing of equation (A17) gives, for each good $j$,

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

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

$$\begin{aligned} 0 = E\big\{\big[(q_{ji} - q_{ji'}) - \big(\big(x_i - \widetilde{\boldsymbol{\gamma}}'\widetilde{\mathbf{z}}_i\big)^2 - \big(x_{i'} - \widetilde{\boldsymbol{\gamma}}'\widetilde{\mathbf{z}}_{i'}\big)^2\big)m_j - \widetilde{\mathbf{c}}_j'(\widetilde{\mathbf{z}}_i - \widetilde{\mathbf{z}}_{i'}) \\ - \big(\delta_j - 2m_j\big(\boldsymbol{\alpha}'\widehat{\mathbf{q}}_{g,-ii'} + \beta + \boldsymbol{\kappa}'\widetilde{\mathbf{z}}_g\big)\big)\big(\big(x_i - \widetilde{\boldsymbol{\gamma}}'\widetilde{\mathbf{z}}_i\big) - \big(x_{i'} - \widetilde{\boldsymbol{\gamma}}'\widetilde{\mathbf{z}}_{i'}\big)\big)\big]\mathbf{r}_{gii'}\big\}. \end{aligned} \tag{A18}$$

Combining common terms, we have

$$\begin{aligned} 0 = E\big\{\big[(q_{ji} - q_{ji'}) - \big(x_i^2 - x_{i'}^2\big)m_j + 2(x_i\widetilde{\mathbf{z}}_i - x_{i'}\widetilde{\mathbf{z}}_{i'})'\widetilde{\boldsymbol{\gamma}}m_j - \widetilde{\boldsymbol{\gamma}}'\big(\widetilde{\mathbf{z}}_i\widetilde{\mathbf{z}}_i' - \widetilde{\mathbf{z}}_{i'}\widetilde{\mathbf{z}}_{i'}'\big)\widetilde{\boldsymbol{\gamma}}m_j \\ - \big(\widetilde{\mathbf{c}}_j' - (\delta_j - 2m_j\beta)\widetilde{\boldsymbol{\gamma}}'\big)(\widetilde{\mathbf{z}}_i - \widetilde{\mathbf{z}}_{i'}) - (\delta_j - 2m_j\beta)(x_i - x_{i'}) \\ + 2m_j\big(\boldsymbol{\alpha}'\widehat{\mathbf{q}}_{g,-ii'} + \boldsymbol{\kappa}'\widetilde{\mathbf{z}}_g\big)(x_i - x_{i'}) - 2(\widetilde{\mathbf{z}}_i - \widetilde{\mathbf{z}}_{i'})'\widetilde{\boldsymbol{\gamma}}m_j\big(\boldsymbol{\alpha}'\widehat{\mathbf{q}}_{g,-ii'} + \boldsymbol{\kappa}'\widetilde{\mathbf{z}}_g\big)\big]\mathbf{r}_{gii'}\big\}. \end{aligned} \tag{A19}$$

From the above equation, for each $j = 1, \ldots, J-1$, $m_j$ 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 $x_i(\widetilde{\mathbf{z}}_{i'} - \widetilde{\mathbf{z}}_i)$, $\delta_j - 2m_j\beta$ 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 $\widehat{\mathbf{q}}_{g,-ii'}(x_i - x_{i'})$ and $\widetilde{\mathbf{z}}_g(x_i - x_{i'})$. 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 $j = 1, \ldots, J-1$, given sufficient variation in the covariates and instruments. Let $\boldsymbol{\eta} = \boldsymbol{\delta} - 2\mathbf{m}\beta$. As $\sum_{j=1}^J m_j p_j = (e^{-\mathbf{b}' \ln \mathbf{p}}) \sum_{j=1}^J d_j = 0$ 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 $\widetilde{\mathbf{c}}_J = (\widetilde{\boldsymbol{\gamma}} - \sum_{j=1}^{J-1} \widetilde{\mathbf{c}}_j p_j)/p_J$, 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 $t = 1, \ldots, T$. 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 $(r_{11}, \ldots, r_{JJ}, r_{12}, \ldots, r_{J-1,J})$ and $\mathbf{b}$ as $(b_1, \ldots, b_{J-1})$. 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: $\delta_{jt} = b_j/p_{jt}$, $m_{jt} = (e^{-\mathbf{b}' \ln \mathbf{p}_t}) d_j/p_{jt}$ 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 $j = 1, \ldots, J-1$ and $d_J = -\sum_{j=1}^{J-1} d_j$. In our data, prices vary by time and region, yielding $T$ much higher than necessary.

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

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

from which we can construct

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

Hence,

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

Pairwise differencing gives

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

where the composite error is

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

Make the following assumptions.

ASSUMPTION B1. *Each individual $i$ in group $g$ satisfies equation* (A16). *Unobserved errors $\mathbf{u}_i$'s are independent across groups and have zero mean conditional on all $(x_l, \mathbf{z}_l)$ for $l \in g$, and $\mathbf{v}_g$ are unobserved group level fixed effects. The number of observed groups $G \to \infty$. For each observed group $g$, a sample of $n_g$ observations of $\mathbf{q}_i$, $x_i$, $\mathbf{z}_i$ is observed. Each sample size $n_g$ is fixed and does not go to infinity. The true number of individuals comprising each group is unknown.*

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

$$\begin{aligned}\overline{\mathbf{q}}_g = \overline{x_g^2}\mathbf{m} + \big(\widetilde{\boldsymbol{\gamma}}'\overline{\mathbf{z}\mathbf{z}'}_g\widetilde{\boldsymbol{\gamma}}\big)\mathbf{m} + \mathbf{m}\big(\boldsymbol{\alpha}'\overline{\mathbf{q}}_g + \boldsymbol{\kappa}'\widetilde{\mathbf{z}}_g + \beta\big)^2 - 2\mathbf{m}\big(\boldsymbol{\alpha}'\overline{\mathbf{q}}_g + \boldsymbol{\kappa}'\widetilde{\mathbf{z}}_g + \beta\big)\big(\overline{x}_g - \widetilde{\boldsymbol{\gamma}}'\overline{\mathbf{z}}_g\big) \\ - 2\mathbf{m}\widetilde{\boldsymbol{\gamma}}'\overline{x\mathbf{z}}_g + \big(\overline{x}_g - \beta - \boldsymbol{\alpha}'\overline{\mathbf{q}}_g - \widetilde{\boldsymbol{\gamma}}'\overline{\mathbf{z}}_g - \boldsymbol{\kappa}'\widetilde{\mathbf{z}}_g\big)\boldsymbol{\delta} + \mathbf{r} + \mathbf{A}\overline{\mathbf{q}}_g + \widetilde{\mathbf{C}}\overline{\mathbf{z}}_g + \mathbf{D}\widetilde{\mathbf{z}}_g + \mathbf{v}_g. \quad \text{(A20)}\end{aligned}$$

Assumption B1 just defines the model. Assumption B2 ensures that an equilibrium exists for each group, thereby avoiding Tamer's (2003) potential incoherence problem. To see this, observe that if $A \neq 0$ and $J = 2$, i.e., $q$ is a scalar, then $\overline{\mathbf{q}}_g$ has the solution

$$\begin{aligned}\overline{q}_g = \frac{1}{2m(Ap)^2}\big\{&\big(2mAp\big(\overline{x}_g - \widetilde{\boldsymbol{\gamma}}'\overline{\mathbf{z}}_g - \boldsymbol{\kappa}'\widetilde{\mathbf{z}}_g - \beta\big) + 1 - A + pA\delta\big) \\ &\pm \big[\big(2mAp\big(\overline{x}_g - \widetilde{\boldsymbol{\gamma}}'\overline{\mathbf{z}}_g - \boldsymbol{\kappa}'\widetilde{\mathbf{z}}_g - \beta\big) \\ &+ 1 - A + pA\delta\big)^2 - 4m(Ap)^2\big(m\overline{x_g^2} + m\boldsymbol{\gamma}'\overline{\mathbf{z}\mathbf{z}'}_g\boldsymbol{\gamma} + m\big(\boldsymbol{\kappa}'\widetilde{\mathbf{z}}_g + \beta\big)^2 \\ &- 2m\big(\boldsymbol{\kappa}'\widetilde{\mathbf{z}}_g + \beta\big)\big(\overline{x}_g - \widetilde{\boldsymbol{\gamma}}'\overline{\mathbf{z}}_g\big) \\ &- 2m\widetilde{\boldsymbol{\gamma}}'\overline{x\mathbf{z}}_g + \big(\overline{x}_g - \beta - \widetilde{\boldsymbol{\gamma}}'\overline{\mathbf{z}}_g - \boldsymbol{\kappa}'\widetilde{\mathbf{z}}_g\big)\delta + r + \widetilde{C}\overline{\mathbf{z}}_g + D\widetilde{\mathbf{z}}_g + v_g)\big)\big]^{1/2}\big\}, \quad \text{(A21)}\end{aligned}$$

while if $A$ does equal zero, then the model will be trivially identified because in that case there aren't any peer effects. From equation (A21), 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 (A20) 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 $(\overline{x}_g, \overline{\mathbf{z}}_g) = E((x_i, \widetilde{\mathbf{z}}_i) \mid i \in g)$ 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 $\sigma^2_{xg} = \mathrm{Var}(x_i \mid i \in g)$, $\sigma_{x\mathbf{z}g} = \mathrm{Cov}(x_i, \widetilde{\mathbf{z}}_i \mid i \in g)$ and $\Sigma_{\mathbf{z}g} = \mathrm{Var}(\widetilde{\mathbf{z}}_i \mid i \in g)$. Denote $\varepsilon_{ix} = x_i - \overline{x}_g$ and $\boldsymbol{\varepsilon}_{iz} = \widetilde{\mathbf{z}}_i - \overline{\mathbf{z}}_g$. Assume $E((\varepsilon_{ix}, \boldsymbol{\varepsilon}_{iz}) | \overline{\mathbf{z}}_g, \widetilde{\mathbf{z}}_g, \overline{x\mathbf{z}}_g, \overline{\mathbf{z}\mathbf{z}'}_g, \overline{x}_g, \overline{x_g^2}, \mathbf{v}_g, \mathbf{r}_g) = 0$ 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

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

Then we have the moment condition

$$\begin{aligned} 0 = E\big\{\big[\mathbf{q}_i - \mathbf{q}_{i'} + 2\mathbf{m}\big(\boldsymbol{\alpha}'\widehat{\mathbf{q}}_{g,-ii'} + \boldsymbol{\kappa}'\widetilde{\mathbf{z}}_g\big)\big[(x_i - x_{i'}) - \widetilde{\boldsymbol{\gamma}}'(\widetilde{\mathbf{z}}_i - \widetilde{\mathbf{z}}_{i'})\big] - \big(x_i^2 - x_{i'}^2\big)\mathbf{m} \\ - \widetilde{\boldsymbol{\gamma}}'\big(\widetilde{\mathbf{z}}_i\widetilde{\mathbf{z}}_i' - \widetilde{\mathbf{z}}_{i'}\widetilde{\mathbf{z}}_{i'}'\big)\widetilde{\boldsymbol{\gamma}}\mathbf{m} + 2\mathbf{m}\widetilde{\boldsymbol{\gamma}}'(\widetilde{\mathbf{z}}_i x_i - \widetilde{\mathbf{z}}_{i'} x_{i'}) - \boldsymbol{\eta}(x_i - x_{i'}) + \big(\boldsymbol{\eta}\widetilde{\boldsymbol{\gamma}}' - \widetilde{\mathbf{C}}\big)(\widetilde{\mathbf{z}}_i - \widetilde{\mathbf{z}}_{i'})\big] | \\ x_i, x_{i'}, \mathbf{z}_i, \mathbf{z}_{i'}, \mathbf{r}_g\big\} \end{aligned} \tag{A22}$$

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

$$\begin{aligned} E\Bigg[ & (\mathbf{q}_i - \mathbf{q}_{i'} + 2e^{-\mathbf{b}'\ln\mathbf{p}_t}\frac{\mathbf{d}}{\mathbf{p}_t}\big(\mathbf{p}_t'\mathbf{A}\widehat{\mathbf{q}}_{gt,-ii'} + \mathbf{p}_t'\mathbf{D}\widetilde{\mathbf{z}}_g\big)\big[(x_i - x_{i'}) - \mathbf{p}_t'\widetilde{\mathbf{C}}(\widetilde{\mathbf{z}}_i - \widetilde{\mathbf{z}}_{i'})\big] - e^{-\mathbf{b}'\ln\mathbf{p}_t}\frac{\mathbf{d}}{\mathbf{p}_t} \\ & \times \big[\big(x_i^2 - x_{i'}^2\big) + \mathbf{p}_t'\widetilde{\mathbf{C}}\big(\widetilde{\mathbf{z}}_i\widetilde{\mathbf{z}}_i' - \widetilde{\mathbf{z}}_{i'}\widetilde{\mathbf{z}}_{i'}'\big)\widetilde{\mathbf{C}}'\mathbf{p}_t - 2\mathbf{p}_t'\widetilde{\mathbf{C}}(\mathbf{z}_i x_i - \mathbf{z}_{i'} x_{i'})\big] \\ & - \left(\frac{\mathbf{b}}{\mathbf{p}_t} - 2e^{-\mathbf{b}'\ln\mathbf{p}_t}\frac{\mathbf{d}}{\mathbf{p}_t}\mathbf{p}_t^{1/2\prime}\mathbf{R}\mathbf{p}_t^{1/2}\right) \\ & \times (x_i - x_{i'}) + \left[\left(\frac{\mathbf{b}}{\mathbf{p}_t} - 2e^{-\mathbf{b}'\ln\mathbf{p}_t}\frac{\mathbf{d}}{\mathbf{p}_t}\mathbf{p}_t^{1/2\prime}\mathbf{R}\mathbf{p}_t^{1/2}\right)\widetilde{\mathbf{C}}'\mathbf{p}_t - \widetilde{\mathbf{C}}\right](\widetilde{\mathbf{z}}_i - \widetilde{\mathbf{z}}_{i'})| \\ & x_i, x_{i'}, \mathbf{z}_i, \mathbf{z}_{i'}, \mathbf{r}_g\Bigg] = 0 \end{aligned} \tag{A23}$$

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

$$\begin{aligned}
L_{1jgii'} &= (q_{ji} - q_{ji'}), \qquad L_{2jgii'} = \widehat{q}_{jg,-ii'}(x_i - x_{i'}),\\
L_{3jkgii'} &= \widehat{q}_{jgt,-ii'}(\widetilde{z}_{ki} - \widetilde{z}_{ki'}), \qquad L_{4k_2gii'} = \widetilde{z}_{k_2g}(x_i - x_{i'}),\\
L_{5kk_2gii'} &= \widetilde{z}_{k_2g}(\widetilde{z}_{ki} - \widetilde{z}_{ki'}), \qquad L_{6gii'} = x_i^2 - x_{i'}^2,\\
L_{7kk'gii'} &= \widetilde{z}_{ki}\widetilde{z}_{k'i} - \widetilde{z}_{ki'}\widetilde{z}_{k'i'}, \qquad L_{8kgii'} = \widetilde{z}_{ki}x_i - \widetilde{z}_{ki'}x_{i'},\\
L_{9gii'} &= x_i - x_{i'}, \qquad L_{10kgii'} = \widetilde{z}_{ki} - \widetilde{z}_{ki'},
\end{aligned} \tag{A24}$$

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

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

where $\widetilde{\gamma}_k$ is the $k$th element of $\widetilde{\boldsymbol{\gamma}} = \widetilde{\mathbf{C}}'\mathbf{p}$, $\kappa_{k_2}$ is the $k_2$th element of $\boldsymbol{\kappa} = \mathbf{D}'\mathbf{p}$, and $\widetilde{c}_{jk}$ is the $(j,k)$th element of $\widetilde{\mathbf{C}}$.

ASSUMPTION B5. *$E(\mathbf{Q}_g')E(\mathbf{Q}_g)$ is nonsingular, where*

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

Under Assumption B5, we can identify

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

for each $j = 1, \ldots, J-1$. 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

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

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

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

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

where $\boldsymbol{\eta}_t = (\eta_{1t}, \ldots, \eta_{J-1,t})'$. Hence, we need the $T \times J$ matrix, $\mathbf{P}$ has full column rank to further identify parameters in $\mathbf{A}$ and $\mathbf{D}$; we 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 identify 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 $\mathbf{p}_1, \ldots, \mathbf{p}_T$ such that the $T \times J$ matrix $\mathbf{P} = (\mathbf{p}_1, \ldots, \mathbf{p}_T)'$ and the $(J-1)T \times [J-1+J(J+1)/2]$ matrix, $\Lambda$ both have full column rank.*

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

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

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

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

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

To illustrate, in the two-goods system, that is, $J = 2$, 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* (A16) *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

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

So, $\Gamma_{ngt}$ 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

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

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

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

with

$$m_{jt} = e^{-\mathbf{b}'\ln \mathbf{p}_t}\frac{d_j}{p_{jt}}, \qquad \boldsymbol{\alpha}_t = \mathbf{A}'\mathbf{p}_t, \qquad \widetilde{\boldsymbol{\gamma}}_t = \widetilde{\mathbf{C}}'\mathbf{p}_t, \qquad \boldsymbol{\kappa}_t = \mathbf{D}'\mathbf{p}_t,$$

$$\beta_t = \mathbf{p}_t^{1/2\prime}\mathbf{R}\mathbf{p}_t^{1/2}, \qquad \delta_{jt} = \frac{b_j}{p_{jt}}.$$

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 $\varepsilon$'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 $\varepsilon$'s, we assume that true group level variables like $\overline{x}_{gt}$ and $\overline{\mathbf{z}}_{gt}$ are correlated over time, for example, 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

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

because

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

and

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

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

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

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

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

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

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

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

Therefore,

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

By rewriting $q_{ji}$ as

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

where

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

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

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

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

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

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:

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

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 $i \in g$,

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

Denote

$$L_{1jgi} = q_{ji}, \qquad L_{2jj'gi} = \frac{1}{n_g-1}\sum_{i'\in g, i'\neq i}\widehat{q}_{jg,-ii'}q_{j'i'}, \qquad L_{3gi} = x_i^2, \qquad L_{4kk'gi} = \widetilde{z}_{ki}\widetilde{z}_{k'i},$$

$$L_{5k_2k_2'gi} = \widetilde{z}_{k_2g}\widetilde{z}_{k_2'g}, \qquad L_{6kgi} = \widetilde{z}_{ki}x_i, \qquad L_{7k_2gi} = \widetilde{z}_{k_2g}x_i, \qquad L_{8jgi} = \widehat{q}_{jg,-i}x_i,$$

$$L_{9jkgi} = \widehat{q}_{jg,-i}\widetilde{z}_{ki}, \qquad L_{10jk_2gi} = \widehat{q}_{jg,-i}\widetilde{z}_{k_2g}, \qquad L_{11kk_2gi} = \widetilde{z}_{ki}\widetilde{z}_{k_2g}, \qquad L_{12jgi} = \widehat{q}_{jg,-i},$$

$$L_{13gi} = x_i, \qquad L_{14kgi} = \widetilde{z}_{ki}, \qquad L_{15k_2gi} = \widetilde{z}_{k_2g}, \qquad L_{16gi} = 1.$$

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

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

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

Assumption B7. *$E(\mathbf{H}_g')E(\mathbf{H}_g)$ is nonsingular, where*

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

Under Assumptions B1–B4 and Assumption B7, we can identify

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

for each $j = 1, \ldots, J-1$. 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 $j = 1, \ldots, J-1$ are all identified. Then $\mathbf{A}_J = (\boldsymbol{\alpha} - \sum_{j=1}^{J-1}\mathbf{A}_j p_j)/p_J$, $\widetilde{\mathbf{c}}_J = (\widetilde{\boldsymbol{\gamma}} - \sum_{j=1}^{J-1}\widetilde{\mathbf{c}}_j p_j)/p_J$, 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 $\mathbf{A}\overline{\mathbf{q}}_g$, which is differenced out in the fixed-effects

model, and only $\widetilde{\mathbf{C}}$ can be identified from the cross-product of $\overline{\mathbf{q}}_{\boldsymbol{g}}$ 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 $\boldsymbol{\alpha}$, $\beta$, $\widetilde{\boldsymbol{\gamma}}$, $\boldsymbol{\kappa}$, $\boldsymbol{\delta}$, $r_j$, and $\mathbf{m}$ with their corresponding full demand system expressions, that is, $\boldsymbol{\alpha} = \mathbf{A}'\mathbf{p}$, $\beta = \mathbf{p}^{1/2\prime}\mathbf{R}\mathbf{p}^{1/2}$, 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

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

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

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

with

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

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

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

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 $\boldsymbol{\Sigma}_v = \text{Var}(\mathbf{v}_{gt})$ do not vary by $t$. Then we can replace $v_{jt0}$ with

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

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 $\boldsymbol{\Sigma}_v$ instead of all the separate $v_{jt0}$ constants.

### A.7 *Generic model identification with observed network structure*

Here, we consider extending the model to an application where network structure is observed. Instead of assigning people to groups, the outcome of each individual now depends on the mean of that specific person's set of friends. As before, we may only observe a small subset of each person's friends, so the mean outcome of each person's friends is observed with error. We show that our method of handling nonlinearity, endogeneity, and measurement error extends to this framework. Note, however, that since this structure no longer has groups, in this extension we drop the presence of group level fixed or random effects.

Suppose we have full or partial information of a network $W_N$, with $W_{ij} = 1$ if $i$ and $j$ are friends:

$$y_i = (\overline{y}_{-iN} a + x_i b)^2 d + (\overline{y}_{-iN} a + x_i b) + u_i,$$

where $\overline{y}_{-iN}$ is the mean outcome of $i$'s friends, that is,

$$\overline{y}_{-iN} = \frac{1}{N_i} \sum_{j \in N} W_{ij} \overline{y}_{jN}$$

with $N_i = \sum_{j \in N} W_{ij}$ and

$$\overline{y}_{jN} = (\overline{y}_{-jN} a + x_j b)^2 d + (\overline{y}_{-jN} a + x_j b).$$

But we can only sample outcomes for a small subset $n \subset N$, and hence replace $\overline{y}_{-iN}$ with an estimate

$$\widehat{y}_{-in} = \frac{1}{n_i} \sum_{j \in n} W_{ij} y_j.$$

Notice that $\widehat{y}_{-in}$ differs from $\overline{y}_{-iN}$ in two ways: (1) the expected outcome of $j$'s friend $\overline{y}_{jN}$ is the replaced by the observed outcome $y_j$; (2) we can only take the weighted average of the surveyed sample $n$, not the true sample $N$. We then get the decomposition:

$$\overline{y}_{-iN} - \widehat{y}_{-in} = \frac{1}{N_i} \sum_{j \in N} W_{ij} \overline{y}_{jN} - \frac{1}{n_i} \sum_{j \in n} W_{ij} y_j = \frac{1}{N_i n_i} \left( n_i \sum_{j \in N} W_{ij} \overline{y}_{jN} - N_i \sum_{j \in n} W_{ij} y_j \right)$$

$$= \frac{1}{N_i n_i}\left( n_i \sum_{j\in n} W_{ij}\overline{y}_{jN} + n_i \sum_{j\notin n} W_{ij}\overline{y}_{jN} - n_i \sum_{j\in n} W_{ij} y_j - (N_i - n_i)\sum_{j\in n} W_{ij} y_j \right)$$

$$= \frac{1}{N_i n_i}\left( N_i \sum_{j\in n} W_{ij}\overline{y}_{jN} - (N_i - n_i)\sum_{j\in n} W_{ij}\overline{y}_{jN} - N_i \sum_{j\in n} W_{ij} y_j + n_i \sum_{j\notin n} W_{ij}\overline{y}_{jN} \right)$$

$$= \frac{1}{N_i n_i}\left( N_i \sum_{j\in n} W_{ij}(\overline{y}_{jN} - y_j) + n_i \sum_{j\notin n} W_{ij}\overline{y}_{jN} - (N_i - n_i)\sum_{j\in n} W_{ij}\overline{y}_{jN} \right)$$

$$= \frac{1}{n_i}\sum_{j\in n} W_{ij} u_j + \frac{N_i - n_i}{N_i}\left( \frac{1}{N_i - n_i}\sum_{j\notin n} W_{ij}\overline{y}_{jN} - \frac{1}{n_i}\sum_{j\in n} W_{ij}\overline{y}_{jN} \right).$$

Denote

$$\eta_{iN} = \frac{1}{n_i}\sum_{j\in n} W_{ij}\overline{y}_{jN} - \frac{1}{N_i - n_i}\sum_{j\notin n} W_{ij}\overline{y}_{jN}.$$

Then $\eta_{iN}$ is difference between the average expectation of $i$'s observed friends and the average expectation of $i$'s unobserved friends. If the sampled data is a random draw from the full network, we will have $E(\eta_{iN}) = 0$ and $\eta_{iN}$ independent of $u$ for all $i$ as the sampling is purely random. From this construction, we can see that

$$\widehat{y}_{-in} = \overline{y}_{-iN} + \left(1 - \frac{n_i}{N_i}\right)\eta_{iN} - \frac{1}{n_i}\sum_{j\in n} W_{ij} u_j$$

is again an unbiased estimator of $\overline{y}_{-iN}$. The measurement error contains two parts, where the first part is purely random but the second part is constructed by the model error $u$ and will not vanish even if both $N_i$ and $n_i$ increases.

If we now replace $\overline{y}_{-iN}$ with $\widehat{y}_{-in}$, we have

$$y_i = (\widehat{y}_{-in} a + x_i b)^2 d + (\widehat{y}_{-in} a + x_i b) + u_i + \varepsilon_i, \tag{A25}$$

where $\varepsilon_i$ is given by

$$\begin{aligned}\varepsilon_i &= \left(\overline{y}^2_{-iN} - \widehat{y}^2_{-in}\right) a^2 d + 2(\overline{y}_{-iN} - \widehat{y}_{-in}) x_i a b d + (\overline{y}_{-iN} - \widehat{y}_{-in}) a \\ &= \left( \frac{1}{n_i}\sum_{j\in n} W_{ij} u_j - \left(1 - \frac{n_i}{N_i}\right)\eta_{iN} \right) \\ &\quad \times \left[ \left( 2\overline{y}_{-iN} + \left(1 - \frac{n_i}{N_i}\right)\eta_{iN} - \frac{1}{n_i}\sum_{j\in n} W_{ij} u_j \right) a^2 d + 2 x_i a b d + a \right]. \end{aligned} \tag{A26}$$

As before, this $\varepsilon_i$ does not have zero conditional mean due to the quadratic term.

Since we no longer have groups, we cannot look at all pairs of observations within a group. Instead, we can randomly split $i$'s observed friends into two subsets $n_i = n_i^{(1)} + n_i^{(2)}$ and construct the sample mean from each subset

$$\widehat{y}^{(1)}_{-in} = \frac{1}{n_i^{(1)}}\sum_{j\in n_i^{(1)}} W_{ij} y_j \quad \text{and} \quad \widehat{y}^{(2)}_{-in} = \frac{1}{n_i^{(2)}}\sum_{j\in n_i^{(2)}} W_{ij} y_j.$$

Then

$$\widehat{y}_{-in}^{(1)} = \overline{y}_{-iN} + \left(1 - \frac{n_i^{(1)}}{N_i}\right)\eta_{iN}^{(1)} - \frac{1}{n_i^{(1)}} \sum_{j \in n_i^{(1)}} W_{ij}u_j,$$

$$\widehat{y}_{-in}^{(2)} = \overline{y}_{-iN} + \left(1 - \frac{n_i^{(2)}}{N_i}\right)\eta_{iN}^{(2)} - \frac{1}{n_i^{(2)}} \sum_{j \in n_i^{(2)}} W_{ij}u_j,$$

and there are no common $u_j$'s in the two subsample averages. Then our model becomes

$$y_i = \big(\widehat{y}_{-in}^{(1)}a + x_i b\big)\big(\widehat{y}_{-in}^{(2)}a + x_i b\big)d + (\widehat{y}_{-in}a + x_i b) + u_i + \widetilde{\varepsilon}_i, \tag{A27}$$

where

$$\begin{aligned}
\widetilde{\varepsilon}_i &= \big(\overline{y}_{-iN}^2 - \widehat{y}_{-in}^{(1)}\widehat{y}_{-in}^{(2)}\big)a^2 d + \big(2\overline{y}_{-iN} - \widehat{y}_{-in}^{(1)} - \widehat{y}_{-in}^{(2)}\big)x_i abd + (\overline{y}_{-iN} - \widehat{y}_{-in})a \\
&= a^2 d\bigg(\frac{1}{n_i^{(1)}n_i^{(2)}} \sum_{j \in n_i^{(1)}} W_{ij}u_j \sum_{j \in n_i^{(2)}} W_{ij}u_j + \left(1 - \frac{n_i^{(1)}}{N_i}\right)\eta_{iN}^{(1)}\left(1 - \frac{n_i^{(2)}}{N_i}\right)\eta_{iN}^{(2)}\bigg) \\
&\quad + a^2 d\bigg(\left(1 - \frac{n_i^{(1)}}{N_i}\right)\eta_{iN}^{(1)}\frac{1}{n_i^{(2)}} \sum_{j \in n_i^{(2)}} W_{ij}u_j + \left(1 - \frac{n_i^{(2)}}{N_i}\right)\eta_{iN}^{(2)}\frac{1}{n_i^{(1)}} \sum_{j \in n_i^{(1)}} W_{ij}u_j\bigg) \\
&\quad + d\bigg(\frac{1}{n_i^{(1)}} \sum_{j \in n_i^{(1)}} W_{ij}u_j + \frac{1}{n_i^{(2)}} \sum_{j \in n_i^{(2)}} W_{ij}u_j - \left(1 - \frac{n_i^{(1)}}{N_i}\right)\eta_{iN}^{(1)} \\
&\quad - \left(1 - \frac{n_i^{(2)}}{N_i}\right)\eta_{iN}^{(2)}\bigg)\big(x_i ab + a^2\overline{y}_{-iN}\big) \\
&\quad + \bigg(\frac{1}{n_i} \sum_{j \in n} W_{ij}u_j - \left(1 - \frac{n_i}{N_i}\right)\eta_{iN}\bigg)a
\end{aligned}$$

We can then show that

$$E(u_i + \widetilde{\varepsilon}_i \mid x_i, x_j) = 0, \tag{A28}$$

where $x_j$ are from those of $i$'s observed friends. With these moments, we can now construct instruments as before for GMM estimation.

## Appendix B: Preliminary data analyses

### B.1 *Generic model estimates*

In other, nondemand settings, the generic peer effects model of Section 3 may be more appropriate than the structural demand model. We implemented this model in Section 4.2, but in this section describe the results in more detail.

TABLE A1. Food spending as a function of group spending, generic model estimates.

| | RE | | FE | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| $a$ (peer mean expenditure) | 0.142 | 0.131 | −1.024 | −1.077 |
| | (0.047) | (0.046) | (0.428) | (0.442) |
| $b$ (own expenditure | 0.413 | 0.415 | 0.462 | 0.456 |
| | (0.011) | (0.011) | (0.019) | (0.018) |
| $d$ (curvature) | −0.181 | −0.182 | −0.099 | −0.067 |
| | (0.010) | (0.010) | (0.017) | (0.012) |
| $-a/b$ | −0.344 | −0.315 | 2.214 | 2.361 |
| | (0.118) | (0.115) | (0.928) | (0.975) |
| $p$ for $-a/b=1$ | 0.000 | 0.000 | 0.191 | 0.163 |
| Hausman for $a$ | | | 7.506 | 7.536 |
| P-value | | | 0.006 | 0.006 |
| Individual controls | No | Yes | No | Yes |
| Number of pairs | 128,640 | 128,640 | 128,640 | 128,640 |
| Number of groups | 4599 | 4599 | 4599 | 4599 |

*Note*: 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.

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 A1.

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 nonfood 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.

TABLE A2. Subjective well-being summary statistics.

| | Mean | SD | Min | Max |
|---|---|---|---|---|
| Income group 2 (= 1) | 0.4 | 0.49 | 0 | 1 |
| Income group 3 (= 1) | 0.21 | 0.4 | 0 | 1 |
| Income group 4 (= 1) | 0.087 | 0.28 | 0 | 1 |
| Income group 5 (= 1) | 0.041 | 0.2 | 0 | 1 |
| Group expenditure (1000 rupees) | 5.6 | 2.6 | 2.8 | 18 |
| Age | 0.34 | 0.12 | 0.15 | 0.77 |
| Sex | 1.4 | 0.49 | 1 | 2 |
| Household size | 0.32 | 0.19 | 0 | 0.9 |
| Married (= 1) | 0.84 | 0.36 | 0 | 1 |
| Primary education (= 1) | 0.095 | 0.29 | 0 | 1 |
| Secondary education (= 1) | 0.13 | 0.34 | 0 | 1 |
| Observations | 5084 | | | |

*Note*: Life satisfaction variable from the 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.

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, for example, 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 A2 summarizes the 4th (2001), 5th (2006), and 6th (2014) waves of the World ValuesSurvey. 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 5th 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 10-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

TABLE A3. 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 | 5084 | 5084 | 5084 | 5084 | 5084 | 5084 |

*Note*: 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.

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 A3 presents estimates of well-being as a function of both own total expenditures and group total expenditures, specified as

$$U_i = \sum_{s=2}^{5} \beta_g 1[I_i = s] + \pi \overline{\overline{x}}_{gt} + X_{igt}\alpha + \gamma_g + \phi_t + \varepsilon_{igt}, \tag{B1}$$

where $U_i$ is the z-normalized well-being indicator, $1[g_i = s]$ is an indicator for individual $i$ belonging to income group $s$, $\overline{\overline{x}}_{gt}$ is imputed group expenditures, $X_{igt}$ is vector of individual level controls, $\gamma_g$ is a group level fixed effect (groups are defined within states, so this effectively includes a state fixed effect as well), and $\phi_t$ is a year fixed effect. Identification of $\pi$ 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 A3 imply that satisfaction is increasing over the entire range of individual expenditures, but that a 1000 rupee increase in peer expenditure $\overline{\overline{x}}_{gt}$ *decreases* satisfaction by 0.15 standard deviations. Other specifications in Table A3 give similar results. The signs of these effects are consistent with our model of peer expenditures as negative consumption externalities. The magnitudes are also

relative large (average peer expenditure is 5554, with a standard deviation of 2580), consistent with our structural results.

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 A3). 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.

## References

Lee, Lung-Fei (2007), "Identification and estimation of econometric models with group interactions, contextual factors and fixed effects." *Journal of Econometrics*, 140 (2), 333–374. [1]

Tamer, Elie (2003), "Incomplete simultaneous discrete response model with multiple equilibria." *Review of Economic Studies*, 70 (1), 147–165. [2, 3, 15]

---

Co-editor Christopher Taber handled this manuscript.

Manuscript received 31 October, 2020; final version accepted 5 October, 2021; available online 19 January, 2022.