Online Supplemental Appendix

A Additional Tables and Figures

Table A.1 shows summary statistics for the main products in our estimation sample - average prices, average / 25th percentile / median / 75th percentile shares, and the share of markets (month-DMA) in which each product is offered. Panel A presents these statistics for the 10 most popular products in the demand estimation sample. Panel B presents these statistics for the next 5 most popular Coca Cola products, while Panel C does the same for Energy Brands products. Note that some Coca Cola products have perfect coverage during this period - naturally, this will affect estimation of any bounds where the deviation requires adding a product. We discuss this issue in detail in Section 8.1.

		Avg	Shares (in %)				Share of
Firms	Products	price	Avg	p25	p50	p75	markets
Panel A: Top 10 Products							
PepsiCo	Aquafina	5.6	2.28	1.21	1.88	2.85	99.78
Coca-Cola	Dasani	5.9	1.44	0.68	1.12	1.78	99.08
Nestle SA	Arrowhead	4.9	2.41	1.44	2.26	3.25	22.55
Private 1	Independent	4.2	2.59	0.37	1.39	3.99	66.83
Nestle SA	Deer Park	5.5	1.60	0.06	0.94	2.39	33.77
Nestle SA	Poland Spring	5.9	0.58	0.01	0.04	0.19	45.01
PepsiCo	Aquafina 20 Oz	1.3	0.19	0.11	0.16	0.23	100.00
Coca-Cola	Dasani 20 Oz	1.3	0.21	0.11	0.16	0.24	100.00
Private 2	Independent	4.8	1.21	0.06	0.38	1.67	75.91
PepsiCo	Gatorade 32 Oz	1.2	0.14	0.07	0.11	0.18	99.97
Panel B: Next 5 Coca Cola Products							
	Dasani 12 Oz	4.1	0.15	0.08	0.13	0.19	99.49
	Simply Lemonade	2.7	0.05	0.03	0.04	0.07	85.34
Coca-Cola	Dasani 33.8 Oz	1.3	0.08	0.03	0.05	0.09	100.00
	Minute Maid	3.2	0.07	0.03	0.06	0.09	85.64
	Powerade Ion4	1.0	0.14	0.06	0.10	0.17	64.82
Panel C: Next 5 Energy Brands Products							
	Dragonfruit	1.3	0.04	0.01	0.03	0.05	97.05
Energy Brands	Fruit Punch	1.3	0.03	0.01	0.02	0.04	96.37
(VitaminWater)	Tropical Citrus	1.3	0.03	0.01	0.02	0.04	95.85
	Power C	1.3	0.02	0.01	0.02	0.03	93.22
	Essential	1.3	0.02	0.01	0.02	0.03	92.57

Table A.1: Summary statistics for the 10 best-selling UPCs in our sample and for the next 5 top-selling Coca Cola and Energy Brands products.

B Additional Details on Implementation

B.1 Demand Estimation

We estimate demand on pre-merger data using a nested logit model. This is an admittedly simple demand estimation approach but it helps us devote attention to the issues that are more closely related to inference with moment inequalities. Improving the demand estimation step would not affect the construction of confidence regions for the parameters determining sunk costs, though it would certainly affect the actual numbers. Since our goal here is mostly pedagogical, we keep this initial step simple.

Throughout the appendix we define a market i as a combination of DMA a and time t, so that $i = (a, t)$ and interchangeably index variables by i or by (a, t) depending on the context. That is, for any random variable X, $X_{i,j}$ and $X_{a,t,j}$ denote the same random variable.

To estimate the nested logit model, we divide all products into 3 mutually exclusive groups: branded water, private label water, non-water, and the outside option. We assume a consumer c that lives in DMA a in period t and purchases product j belonging to group g receives indirect utility $U_{c,a,t,j}$ according to

$$U_{c,a,t,j} = \gamma_0 + \gamma_1 p_{a,t,j} + \xi_j + \xi_a + \xi_{j,a} + \xi_{a,t,j} + \zeta_{c,g} + (1-\rho)\epsilon_{c,a,t,j} \ , \tag{A.1}$$

where $p_{a,t,j}$ is price, ξ_j, ξ_a, $\xi_{j,a}$ and $\xi_{a,t,j}$ are product attributes that are unobserved to the econometrician, $\epsilon_{c,a,t,j}$ is identically and independently distributed Gumbel error term, $\zeta_{c,g}$ is a group-level error term, and ρ is the nesting parameter. The distribution of $\zeta_{c,g}$ is the unique distribution such that $\zeta_{c,g} + (1-\rho)\cdot\epsilon_{c,a,t,j}$ is also distributed Gumbel (Cardell, 1997). Estimating this model requires calculating market shares for each product and for the outside good. We do so by assuming that market size in DMA a, M_a, is equal to 1.5 times the maximum liters we observe being sold in a over time. With this assumption, one can compute the market share $\mathrm{s}_{a,t,j}$ of good j in market a in period t as the ratio between the number of liters of good j sold in a during t and M_a. The share of the outside good is one minus the sum of shares for each good j being offered in a during t. Having obtained shares, we invert the market share function (Berry, 1994) and estimate

$$\ln(\mathrm{s}_{a,t,j}) - \ln(\mathrm{s}_{a,t,0}) = \gamma_0 + \gamma_1 p_{a,t,j} + \rho \ln(\bar{s}^g_{a,t,j}) + \xi_j + \xi_a + \xi_{j,a} + \xi_{a,t,j} \ , \tag{A.2}$$

where $\mathrm{s}_{a,t,j}$ and $\mathrm{s}_{a,t,0}$ are product j's share and the outside good's share in DMA a and period t, respectively, and $\bar{s}^g_{a,t,j}$ is product j's share within group g in DMA a and period t. We estimate the model in (A.2) by ordinary least squares (OLS) using product, DMA and product-DMA fixed effects. The main identification assumption is that unobserved product-DMA-time product attributes, captured by $\xi_{a,t,j}$, are orthogonal to prices after controlling for product-DMA fixed effects. The most important deviation relative to the frontier of the demand estimation literature is that we do not use random coefficients. This implies that within nest substitution patterns suffer from the issues highlighted in Berry et al. (1995). Again, since our goal is to illustrate how to conduct inference using moment inequalities in this setting, instead of how to best estimate demand for bottled water and juice drinks, we are comfortable making these assumptions. The results are in Table B.1. We find large price elasticities, particularly for Coca Cola products.

The mean price elasticity is 9.2, and the median is 7.2. Given that we are estimating demand for bottled water products at the UPC level, we do not find these magnitudes implausible. We also find that demand for Energy Brands' products is less price-sensitive. This is consistent with our prior, as these products aim to be more differentiated than standard bottled water.

(a) Demand estimation		(b) Price elasticity weighted by shares			
	$\ln(\mathrm{s}_{a,t,j}/\mathrm{s}_{a,t,0})$		All firms	Coca-Cola	Energy Brands
price	-0.331	mean	9.20	7.98	4.34
	[-0.334, -0.327]	q25	4.37	4.13	4.01
$\ln(\bar{\mathrm{s}}^g_{a,t,j})$	0.899	q50	7.22	7.74	4.30
	[0.897, 0.900]	q75	12.54	10.83	4.63
Product FE	Y	Num. Prod.	212	Observ.	477,133
DMA FE	Y	Num. DMA	205	R^2	0.945

Table B.1: Demand estimation and price elasticity based on a nested-logit model using monthly pre-merge data.

B.2 Marginal Cost Estimation

Having recovered demand estimates, we now turn to estimating marginal costs. Recall that market i denotes a DMA-period (a, t) combination. Under the assumption that firms compete in prices, each firm's first order conditions in market i are

$$\mathrm{s}_{i,j}(p_i) + \sum_{j' \in \mathcal{J}_s} D_{i,j'}(p_{i,j'} - c_{i,j'}) \frac{\partial \mathrm{s}_{i,j'}(p_i)}{\partial p_{i,j}} = 0 \quad \{\forall j \in \mathcal{J}_s | D_{i,j} = 1\} \ . \tag{A.3}$$

We solve this system of equations and recover estimates $\hat{c}_{i,j}$, or $\hat{c}_{a,t,j}$ when being explicit about $i = (a, t)$. Further, we assume that marginal costs of each product j are constant in output and that marginal cost realizations satisfy

$$\ln(\hat{c}_{a,t,j}) = \omega_j + \omega_a + \omega_{j,a} + \omega_{a,t,j} \ , \tag{A.4}$$

where ω_j, ω_a, ω_{jd} and ω_{jdt} are product attributes that are unobservable to the econometrician. We estimate this model by OLS using product, DMA and product-DMA fixed effects.

B.3 Calculation of Estimated Revenue Differentials

This subsection discusses how to move from demand and marginal cost estimates to estimates of revenue differentials. First, note that we need estimated revenue differentials for two firms, Coca-Cola and Energy Brands, as we are only interested in computing sunk costs for them. These companies offer $J = 31$ products in our data, so in this section, and we some abuse of notation, we use $\mathcal{J}$ to denote products that are *only* offered by either Coca-Cola or Energy

Brands, without including products offered by other firms. Consistent with this change, each market i has an observed assortment vector $D_i \in \{0,1\}^J$ where $J = 31$ for the purposes of estimating revenue differentials.

Consider a given product portfolio $\tilde{D}_i \in \{0,1\}^J$, which could be an observed product portfolio or a counter-factual one. The algorithm to compute revenue differential given this particular product assortment is as follows:

1. For each product j that is offered in market $i = (a,t)$ given the product assortment, i.e. $\tilde{D}_{i,j} = 1$, we draw $B = 200$ values from the empirical distributions of $\hat{\xi}_{j,a} + \hat{\xi}_{a,t,j}$ and $\hat{\omega}_{j,a} + \hat{\omega}_{a,t,j}$ with replacement. These objects are the residuals from equations (A.2) and (A.4), respectively. We denote these values by $\{\hat{\xi}^{(b)}_{j,a} + \hat{\xi}^{(b)}_{a,t,j} : 1 \leq b \leq B\}$ and $\{\hat{\omega}^{(b)}_{j,a} + \hat{\omega}^{(b)}_{a,t,j} : 1 \leq b \leq B\}$.
2. We then add to each draw, $\hat{\xi}^{(b)}_{j,a} + \hat{\xi}^{(b)}_{a,t,j}$, the estimated values of ξ_j and ξ_a, and to every draw, $\hat{\omega}^{(b)}_{j,a} + \hat{\omega}^{(b)}_{a,t,j}$, the estimated values of ω_j and ω_a. For each b and for each j such that $\tilde{D}_{i,j} = 1$, this results in

$$\hat{\xi}^{(b)}_{i,j} = \hat{\xi}^{(b)}_{j,a} + \hat{\xi}^{(b)}_{a,t,j} + \hat{\xi}_j + \hat{\xi}_a \tag{A.5}$$

$$\hat{\omega}^{(b)}_{i,j} = \hat{\omega}^{(b)}_{j,a} + \hat{\omega}^{(b)}_{a,t,j} + \hat{\omega}_j + \hat{\omega}_a \ . \tag{A.6}$$

3. For each b and for each j such that $\tilde{D}_{i,j} = 1$, we compute marginal costs, optimal prices, and market shares. First, we use equation (A.4) to solve for marginal costs using $\hat{\omega}^{(b)}_{i,j}$, i.e. $\hat{c}^{(b)}_{i,j} = \exp(\hat{\omega}^{(b)}_{i,j})$. Second, we solve for optimal prices by solving the non-linear system of equations in prices defined by equation (A.3), and denote them by $\hat{p}^{(b)}_{i,j}$. Finally, we compute nested logit market shares as,

$$\hat{s}^{(b)}_{a,t,j} = \frac{\exp\left(\frac{\hat{\delta}^{(b)}_{a,t,j}}{1-\hat{\rho}}\right)}{(\hat{\mathcal{D}}^{(b)}_g)^{\hat{\rho}}\left(\sum_{g'}(\hat{\mathcal{D}}^{(b)}_{g'})^{1-\hat{\rho}}\right)} \ , \tag{A.7}$$

where $\hat{\mathcal{D}}^{(b)}_g = \sum_{j\in\tilde{\mathcal{J}}_g} \exp\left(\frac{\hat{\delta}^{(b)}_{a,t,j}}{1-\hat{\rho}}\right)$, $\hat{\delta}^{(b)}_{a,t,j} = \hat{\gamma}_0 + \hat{\gamma}_1 p_{a,t,j} + \hat{\xi}_j + \hat{\xi}_a + \hat{\xi}^{(b)}_{j,a} + \hat{\xi}^{(b)}_{a,t,j}$, and $\tilde{\mathcal{J}}_g$ denotes the set of products belonging to group g.
4. For each b and each firm s, we compute revenue according to (51) in Assumption 7.1, and then compute $\hat{r}_{s,i}(O_i)$ by averaging across the B draws i.e.,

$$\hat{r}_{s,i}(O_i) = \frac{1}{B}\sum_{b=1}^{B}\sum_{j\in\mathcal{J}_s} M_i \tilde{D}_{i,j}(\hat{p}^{(b)}_{i,j} - \hat{c}^{(b)}_{i,j})\hat{\mathrm{s}}^{(b)}_{i,j} \ .$$

5. Revenue differentials are the difference in these averages between counter-factual product offerings and the observed portfolio, as defined in Section 3, where for each of these alternative assortments we repeat steps 1-4 above.

B.4 Inference without partitioning

As mentioned earlier, we are working with moment conditions that can be partitioned. In particular, each moment is a function of only one firm's parameter, which allows us to estimate the sunk cost of each firm separately. This is not always the case - in most settings it is not possible to separate problems in this fashion. We now discuss how to estimate problems without partitioning.

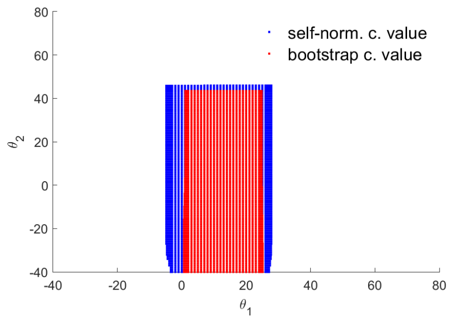
Without separating the problem, we have a two-dimensional vector of parameters of interest. We consider that in two dimensional cases the best course of action is still to perform a grid search, as this avoids the computational issues discussed in Section 6. As before, we operationalize the grid in two steps. In the first step, we use a grid between $-40,000$ and $100,000$ dollars in steps of $1,000$ for each dimension. We then refine the grid in steps of 100 dollars around the bounds obtained in the first step. Table B.2 presents results obtained using our main specification. Figure B.1 presents these results in graphical form. In this figure, blue points denote the 95% confidence set obtained using self-normalized critical values, while red points denote the set obtained using bootstrap critical values. These confidence regions are mostly rectangular, so while there are some differences, the results do not vary significantly between Table B.2 and Table 1. This is to be expected given the nature of our model. The most interesting result here is the fact that computing time is an order of magnitude higher than in table 1. This highlights that researchers working with problems that can be partitioned should do so.

	Crit. Value	θ_1: Coca-Cola	θ_2: Energy Brands	Comp. Time
$\bar{V}$=500	self-norm	[-4.8 , 27.8]	[-40.0 , 45.9]	16.8
	bootstrap	[0.7 , 25.1]	[-40.0 , 43.4]	49.4
$\bar{V}$=1000	self-norm	[-33.9 , 33.1]	[-40.0 , 79.0]	19.3
	bootstrap	[-26.6 , 30.0]	[-40.0 , 74.5]	58.7

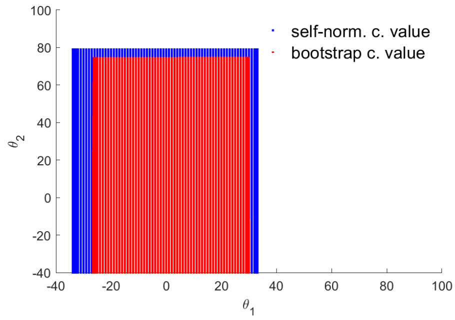
Table B.2: 95%-confidence region for the parameter vector (θ_1, θ_2). The self-normalized and bootstrap critical values are defined in (41) and in Section 5.2, respectively. The parameter space is $[-40, 100]^2$ where units are in thousands of US dollars. Computational time is presented in seconds.

B.5 Details on Counter-factuals

Following the discussion in Section 9, assume that the counter-factual of interest is predicting pricing and product variety decisions for Coca Cola post-merger. In this counter-factual, Coca Cola (which now includes Energy Brands' products) and all of its competitors will be allowed to adjust pricing and product variety. To perform this analysis, one would need sunk cost estimates θ_s for Coca Cola and *all* of its competitors. This is important since, as we discussed in Section 8, by virtue of the model being separable in θ, it is possible and convenient to conduct inference on θ_s for $s \in \{\text{CocaCola}, \text{EnergyBrands}\}$ without accounting for other competitors.



(a) $\bar{V} = 500$



(b) $\bar{V} = 1000$

Figure B.1: 95%-confidence regions for the parameter vector (θ_1, θ_2), where θ_1 and θ_2 are the entry costs of Coca-Cola and Energy Brands, respectively.

The methodology, however, can easily be expanded to recover sunk cost estimates for other firms. We proceed under the assumption that we have computed confidence regions for sunk costs for *all* firms in the market and so $\mathcal{J}$ and $\mathcal{S}$ now include information on all the firms we used to estimate demand. The algorithm we would propose is the following:

Algorithm B.1. Let $\mathcal{D}_i \equiv \{0,1\}^J$ denote the set of all possible product assortment profiles in market i and consider the following algorithm to compute counter-factuals.

Step 1 Compute expected equilibrium variable profits for each firm $s \in \mathcal{S}$ and for each $D_i \in \mathcal{D}_i$ and denote them by $\hat{r}_{s,i}(D_i)$. In Section 7.2 this is done by solving for equilibrium prices following Appendix B.3. Store the values of objects of interest of the counter-factual - prices, consumer surplus, etc.

Step 2 Let $\underline{\theta}_s$ denote the lowest value of θ in firm s confidence region. For each firm $s \in \mathcal{S}$ and each product portfolio $D_i \in \mathcal{D}_i$, consider the inequality

$$\hat{r}_{s,i}(D_i) - \sum_{j \in \mathcal{J}_s} D_{i,j}(\underline{\theta}_s - \bar{V}) \geq 0 \ . \tag{A.8}$$

Denote by $\mathcal{D}_i^2 \subseteq \mathcal{D}_i$ the set of product offerings D_i for which (A.8) holds for all $s \in \mathcal{S}$.

Step 3 For each firm $s \in \mathcal{S}$ and $D_i \in \mathcal{D}_i^2$, let $\mathcal{D}_s(D_i)$ be the subset of $\mathcal{D}_i^2$ that keeps the assortment decisions of s's competitors fixed relative to D_i, i.e., $\mathcal{D}_s(D_i) \equiv \Lambda_s(D_i) \cap \mathcal{D}_i^2$, where $\Lambda_s(D_i)$ is defined in (12). Then, for each $D_i' \in \mathcal{D}_s(D_i)$ check the condition

$$\hat{r}_{s,i}(D_i) - \hat{r}_{s,i}(D_i') - N_s^{\text{in}}(\underline{\theta}_s - \bar{V}) + N_s^{\text{out}}(\bar{\theta}_s + \bar{V}) \geq 0 \ , \tag{A.9}$$

where N_s^{in} is the number of products offered by s in D_i but not in D_i', and N_s^{out} is the number of products offered by s in D_i' but not in D_i. Let $\mathcal{D}_i^3 \subseteq \mathcal{D}_i^2$ denote the set of product portfolios D_i for which (A.9) holds for all $s \in \mathcal{S}$ and all $D_i' \in \mathcal{D}_s(D_i)$.

Step 4 Construct bounds on the objects of interest (average prices, consumer surplus, etc.) by finding the maximum and minimum values of those objects across all portfolios D_i in $\mathcal{D}_i^3$.

Step 1 is the most computationally demanding step, as it requires computing Nash Equilibrium prices for all possible product offerings. We see this step as the main limitation when performing counter-factual analysis in this class of discrete games. Researchers often will need to restrict strategy space in order to make this feasible. For example, instead of modelling the action space of the product variety game as the decision to offer each of the $J = 212$ products in our market separately, which would require solving for $2^{212} - 1$ Nash Equilibria, one can assume that subsets of products are always offered together, or that they will always be offered. Beyond this, parallelization of step 1 is useful, and a benefit of this algorithm is that step 1 only needs to be done once.

Step 2 checks whether a set of product offerings delivers positive profits to all firms in the most advantageous case possible - when sunk costs are the lowest possible value. Since $e_{i,j} \geq \max[\underline{\theta}_s - \overline{V}, 0]$, D_i cannot be an equilibrium assortment profile if the condition in step 2 does not hold for every firm. This check is straightforward once step 1 is computed and it potentially removes irrelevant assortments to be considered in the next step, which is computationally more intense.

Step 3 checks whether there are unilateral incentives to deviate to any other assortments in the case where deviation incentives are the weakest - when sunk costs for offered products are at the lowest possible value and sunk costs for non offered products are at their highest. It is straightforward to show if D_i is a Nash Equilibrium, it must belong to $\mathcal{D}_i^3$. The converse is not true, however, as in equation (A.9) θ_s differs for offered and un-offered products, and our behavioral model posits a fixed θ_s across products produced by firm s. Therefore, $\mathcal{D}_i^3$ contains all Nash Equilibria, and can be thought of as a conservative approximation to the strategies that are a Nash Equilibrium in the counter-factual.

Claim B.1. Suppose $e_{i,j} \in \left[\underline{\theta}_s - \bar{V}, \overline{\theta}_s + \bar{V}\right]$. If D_i is a Nash Equilibrium, then $D_i \in \mathcal{D}_i^3$.

Proof of Claim B.1: If D_i is a Nash Equilibrium, then any firm s must have positive profits

$$\hat{r}_{s,i}\left(D_i\right) - \sum_{j \in \mathcal{J}_s} D_{i,j} e_{i,j} \geq 0 \ , \tag{A.10}$$

and any firm s has no unilateral incentives to deviate

$$\hat{r}_{s,i}\left(D_i\right) - \hat{r}_{s,i}\left(D_i'\right) - \sum_{j \in \mathcal{J}_s} \left(D_{i,j} - D_{i,j}'\right) e_{i,j} \geq 0 \tag{A.11}$$

for any $D_i' \in \Lambda_s(D_i)$. Since $e_{i,j} \in [\underline{\theta}_s - \bar{V}, \overline{\theta}_s + \bar{V}]$, we have that

$$\hat{r}_{s,i}\left(D_i\right) - \sum_{j \in \mathcal{J}_s} D_{i,j} e_{i,j} \leq \hat{r}_{s,i}\left(D_i\right) - \sum_{j \in \mathcal{J}_s} D_{i,j}\left(\underline{\theta}_s - \bar{V}\right) \ ,$$

which implies $D_i \in \mathcal{D}_i^2$ using (A.10). Similarly, we obtain

$$\hat{r}_{s,i}\left(D_i\right) - \hat{r}_{s,i}\left(D_i'\right) - \sum_{j \in \mathcal{J}_s} \left(D_{i,j} - D_{i,j}'\right) e_{i,j} \leq \hat{r}_{s,i}\left(D_i\right) - \hat{r}_{s,i}\left(D_i'\right) - N_s^{\text{in}}(\underline{\theta}_s - \bar{V}) + N_s^{\text{out}}(\bar{\theta}_s + \bar{V}) \ ,$$

which implies $D_i \in \mathcal{D}_i^3$ using (A.11). ∎

C Misspecification Robust CS

Andrews and Kwon (2019) study inference in moment inequality models like those in (1) in settings where the model is misspecified and so $\Theta_0(P)$ in (2) is empty. A potential consequence of misspecification is spurious precision of standard confidence sets for θ, meaning that the coverage probability of the confidence set is less than its nominal level $1-\alpha$ for all parameter values, including any potential pseudo-true value.

Andrews and Kwon (2019) introduce a misspecification index that equals the maximum violation across moment inequalities (normalized by their standard deviations) evaluated at the parameter value that minimizes the maximum violation. In order to describe the method clearly, we introduce some additional notation. Let

$$\sigma_\ell(\theta) = \sqrt{\text{var}[m_\ell(W_i,\theta)]} \tag{A.12}$$

for $1 \le \ell \le k$ and define

$$m^*(W_i,\theta) = (m_1^*(W_i,\theta),\dots,m_k^*(W_i,\theta))' \text{ for } m_\ell^*(W_i,\theta) \equiv \frac{m_\ell(W_i,\theta)}{\sigma_\ell(\theta)} \ . \tag{A.13}$$

$$\Theta_0^*(P) = \{\theta\in\Theta : E_P[m(W_i,\theta)] - r^{\inf}1_k \le 0\} \ , \tag{A.14}$$

where 1_k is a k-dimensional vector of ones, and $r^{\inf}$ is a scalar given by

$$r^{\inf} = \inf_{\theta\in\Theta}\max_{1\le\ell\le k}\max\{E[m_\ell^*(W_i,\theta)],0\} \ . \tag{A.15}$$

Note that $r^{\inf}$ equals the maximum violation across moment inequalities (normalized by their standard deviations) evaluated at the parameter value that minimizes the maximum violation. The misspecification-robust identified set is non-empty even under model misspecification.

The SPUR1 test we use in Section 8.2 consists of the following X steps:

1. Compute the sample analog of $r^{\inf}$ as follows,

$$\hat{r}_n^{\inf} \equiv \inf_{\theta\in\Theta}\max_{1\le\ell\le k}\max\left\{\frac{\bar{m}_{n,\ell}(\theta)}{\hat{\sigma}_{n,\ell}(\theta)},0\right\} \ . \tag{A.16}$$

2. Modify the test statistic in (42) to account for $\hat{r}_n^{\inf}$ as follows,

$$T_n^*(\theta) \equiv T\Big(\sqrt{N}(\hat{D}_n^{-1}(\theta)\bar{m}_n(\theta) + \hat{r}_n^{\inf}1_k), \hat{\Omega}_n(\theta)\Big) \ . \tag{A.17}$$

3. Compute the SPUR1 bootstrap critical, denoted by $\hat{c}_n^{\text{spur}}(1-\alpha,\theta)$, as described in Andrews and Kwon (2019).

4. Reject whenever $T_n^*(\theta) > \hat{c}_n^{\text{spur}}(1-\alpha,\theta)$

Note that the critical value in Step 3, while similar in spirit to the one described in section 5.2, requires substantial modifications to account for the additional randomness introduced by $\hat{r}_n^{\inf}$. We also note that, when the test statistic in (42) equals the max test statistic in (38), $T_n^*(\theta)$ is equivalent to the re-centered version of $T_n(\theta)$ in (38). We refer the reader to Andrews and Kwon (2019) for details or to our companion R and Matlab packages.