Online Technical Appendix

Estimation of the Exposure Function

Define $x_{n,i}$ as one arbitrary characteristic i from those contained in x_n. The Simmons data provides estimates of the form $\widehat{\Pr}(w_{n,p,c} = 1|x_{n,i} = q)$. For example, this could be the percentage of very liberal viewers in the Simmons sample that watched *60 Minutes* on a given viewing occasion. We would like to compare these to their theoretical counterparts, which are given by,

$$\Pr(w_{n,p,c} = 1|x_{n,i} = q) = \frac{\sum_{x:x_i=q} \Pr(x)\Lambda(\gamma_p' x)}{\sum_{x:x_i=q} \Pr(x)}$$

Here, $\Pr(x)$ denotes the probability mass function of the demographic characteristics. This distribution is not directly observed in our television viewership data, but we observe a large sample of x in the NAES rolling cross section. Specifically, let x_n^{rcs} denote the demographic characteristics of an individual observed in the rolling cross section sample of the NAES and let N^{rcs} denote the sample size of the rolling cross section. We use the empirical distribution to obtain an estimate of the (population) distribution of x. That is, we can estimate,

$$\widehat{\Pr}(x) = \tfrac{1}{N^{rcs}} \sum_{n=1}^{N^{rcs}} 1\{x_n^{rcs} = x\}$$

We use this to obtain the desired estimate of the model predicted probabilities. Specifically, we estimate the probability that an individual with characteristic x_i at a value q watches program p on occasion c,

$$\Pr(w_{n,p,c} = 1|x_i = q) \approx \frac{\frac{1}{N^{rcs}} \sum_{x:x_i=q} \sum_{n=1}^{N^{rcs}} 1\{x_n^{rcs} = x\}\Lambda(\gamma_p' x)}{\frac{1}{N^{rcs}} \sum_{x:x_i=q} \sum_{n=1}^{N^{rcs}} 1\{x_n^{rcs} = x\}}$$

Based on this, we can form the following moment conditions,

$$\widehat{h}_{p,i,q}(\gamma_p) = \widehat{\Pr}(w_{n,p,c} = 1|x_i = q) - \frac{\frac{1}{N^{rcs}} \sum_{x:x_i=q} \sum_{n=1}^{N^{rcs}} 1\{x_n^{rcs} = x\}\Lambda(\gamma_p' x)}{\frac{1}{N^{rcs}} \sum_{x:x_i=q} \sum_{n=1}^{N^{rcs}} 1\{x_n^{rcs} = x\}}$$

Here $\widehat{\Pr}(w_{n,p,c} = 1|x_i = q)$ is observed in the Simmons data, i denotes the variable being conditioned on, and q denotes a value that the variable takes.

For each i, we select moments based on all the values q that x_i can take.[16] This corresponds to using the proportion of each subgroup that watches each show as moments in our estimation. Let $\widehat{h}_p(\gamma_p)$ denote the vector of moments for program p. By choosing these moments, we have that $\widehat{h}_p(\gamma_p) \overset{prob.}{\longrightarrow} 0$ if and only if $\gamma_p = \gamma_{p,0}$ (where $\gamma_{p,0}$ denotes the true parameter vector characterizing the data generating process for one program), so the moments will define a minimum distance estimator. We therefore have an exactly identified minimum distance estimator and we estimate $\gamma_{p,0}$ by solving the nonlinear system $\widehat{h}_p(\gamma_p) = 0$.

Estimation of the Effectiveness of Advertising

Based on the model we aggregate over individuals to get the turnout and Republican shares by

$$s^t_{j,m}(\theta) = \tfrac{1}{N^{rcs}_{j,m}} \sum_{n \in \mathcal{N}^{rcs}_{j,m}} \Phi(\xi^t_{j,m} + \beta'_t x^{rcs}_n + \alpha_t e^{rcs}_{n,T})$$

$$s^v_{j,m}(\theta) = \frac{\frac{1}{N^{rcs}_{j,m}} \sum_{n \in \mathcal{N}^{rcs}_{j,m}} \Phi(\xi^t_{j,m} + \beta'_t x^{rcs}_n + \alpha_t e^{rcs}_{n,T}) \Phi(\xi^v_{j,m} + \beta'_v x^{rcs}_n + \alpha_v (e^{rcs}_{n,R} - e^{rcs}_{n,D}))}{s^t_{j,m}(\theta)}$$

where $\mathcal{N}^{rcs}_{j,m}$ denotes the indices of individuals that reside in the jth congressional district and the mth media market and $N^{rcs}_{j,m} = |\mathcal{N}^{rcs}_{j,m}|$.

Our estimation therefore maximizes the log-likelihood function,

$$l(\theta) = \sum_{n=1}^{N} 1\{y^{ep}_n = 0\} \log \Pr(y^{ep}_n = 0 | z^{ep}_n; \theta)$$

$$+1\{y^{ep}_n = 1\} \log \Pr(y^{ep}_n = 1 | z^{ep}_n; \theta) + 1\{y^{ep}_n = 2\} \log \Pr(y^{ep}_n = 2 | z^{ep}_n; \theta)$$

subject to the constraints,

$$s^t_{j,m} - s^t_{j,m}(\theta) = 0 \tag{5}$$

[16] For example, if x_i denotes gender, then x_i can take on the values 1 (for females) and 0 (for males).

$$s^v_{j,m} - s^v_{j,m}(\theta) = 0 \quad (6)$$

In our implementation, we apply a nested fixed point approach (Berry, 1994; Berry, Levinsohn and Pakes, 1995). Each time we evaluate the likelihood at (β, α), we find the vector of fixed effects that are consistent with the constraints. For each j and m, we first solve Equation 5 for $\xi^t_{j,m}$ using a one-dimensional solver and then solve Equation 6 for $\xi^v_{j,m}$ again using a one-dimensional solver.

Our estimation problem is related to the estimation problems of Petrin (2002) and Berry, Levinsohn and Pakes (2004) in that we observe a combination of micro and macro level data and we want to invert aggregate data to recover unobserved characteristics (shocks). In Petrin (2002) and Berry, Levinsohn and Pakes (2004) the primary data to estimate the parameters of interest comes from aggregate level moments, so it is simpler for them to combine the two types of data by constructing aggregate level moments from the micro data. However, in our application, the primary data to estimate the parameters of interest is available at the individual level, so we form the likelihood of an individual observation and constrain the unobserved characteristics to match the aggregate level data. Hence, we apply constrained maximum likelihood estimation.