Appendix

Unreliable Inferences about Unobserved Processes

Carlisle Rainey[†]

Robert A. Jackson[‡]

1 Algorithm for Wrong Link Function Simulations

1. Choose n_x and n_z independently from a Poisson distribution with mean 1.5. If both n_x and n_z equal zero, then set n_z equal to one.
2. Choose $n_w - 1$ from a Poisson distribution with mean 0.5.
3. Assign a type (continuous or binary) to each of the $n_x + n_z + n_w$ variables with equal probability. To simplify the computation, continuous variables take on the values 0.0, 0.2, 0.4, 0.6, 0.8, or 1.0, and binary variables take on the values 0 or 1.
4. To simplify the computation, create a data set WXZ that contains each possible combination of all of the explanatory variables.
5. Choose the preliminary coefficients β^* and $\gamma*$ from a normal distribution. The intercept coefficients come from a normal distribution with mean 0 and standard deviation 2. The slope coefficients come from a normal distribution with mean 0 and standard deviation 1.
6. If the link function of the DGP is logit, then set β equal to β^* and γ equal to γ^*. If the link function is some other function g, then (1) simulate a large data set from the logit DGP, (2) estimate a model with link function g to that data set, and (3) set β and γ equal to the estimates $\hat{\beta}$ and $\hat{\gamma}$, respectively. This step scales the coefficients so that the quantities of interest have comparable magnitude regardless of the DGP.
7. Using the parameters β and γ, simulate 100,000,000 values of y_{obs} and d_{main} for each combination of the explanatory variables.
8. Fit a full observability model to the outcome variable d_{main}. Calculate the first difference as w_1 moves from 0 to 1. Store this large-sample estimate.
9. Fit a partial observability model to the observed outcome variable y_{obs}. Calculate the first difference as w_1 moves from 0 to 1. Store this large-sample estimate.

2 Algorithm for Monotonic Link Function Simulations

1. Choose the coefficients for the non-linear, interactive equations.

$$\Pr(d_{main}) = \beta_{cons} + \beta_w w + \beta_{w^2} w^2$$

and

$$\Pr(d_{nuisance}) = \gamma_{cons} + \gamma_w w + + \gamma_z z + \gamma_{w^2} w^2 + \gamma_{z^2} z^2 + \gamma_{wz} wz.$$

from a uniform distribution that ranges from -1 to 1.

[†]Carlisle Rainey is Assistant Professor of Political Science, Texas A&M University, 2010 Allen Building, College Station, TX, 77843 (crainey@tamu.edu).

[‡]Robert A. Jackson is Professor of Political Science, Florida State University, 531 Bellamy Building, Florida State University, Tallahassee, FL 32306 (rjackson@fsu.edu).

2. Check that this link function produces a monotonic relationship between the w and $\Pr(d_{main})$ and between w and z $\Pr(d_{nuisance})$.
3. Set w as either binary or continuous. For computational ease, continuous variables take on the values 0.0, 0.2, 0.4, 0.6, 0.8, or 1.0, and binary variables take on the values 0 or 1.
4. Create a data set WZ that contains each possible combination of w and z.
5. For each row in WZ, simulate 100,000,000 values of y_{obs} and d_{main}.
6. Fit a full observability model to the outcome variable d_{main}. Calculate the first difference as w_1 moves from 0 to 1. Store this large-sample estimate.
7. Fit a partial observability model to the outcome variable y_{obs}. Calculate the first difference as w_1 moves from 0 to 1. Store this large-sample estimate.