Causal Inference without Ignorability: Identification with Nonrandom Assignment and Missing Treatment Data, Supplemental Material*

Walter R. Mebane, Jr.† Paul Poast‡

November 27, 2012

*Thanks to Peter Aronow for many important contributions. Comments from Don Green, Arthur Spirling, Allison Sovey, Rory Truex and the participants of the 2011 MPSA Annual National Conference and the 2011 Annual Meeting of the Society for Political Methodology are greatly appreciated.

†Department of Political Science and Department of Statistics, University of Michigan, 5700 Haven Hall, Ann Arbor MI, 48109. (E-mail: wmebane@umich.edu)

‡Department of Political Science, Rutgers University, Hickman Hall, New Brunswick NJ, 08901. (E-mail: paul.poast@rutgers.edu)

Supplemental Material Appendices

A Formal Description of the Manski Framework

Let y^j denote a variable measuring unit j's outcomes of interest. We consider binary outcomes, $y^j \in \{0, 1\}$. These outcomes are affected by treatments, measured by variable t, so that $y^j(t^j)$ represents the assumption that the outcome experienced by j is a function of the treatment received by j. Only one of many possible treatments, denoted z, is realized for each unit, so the realized and observable outcome is $y^j(z^j)$. The outcomes in the set $\{y^j(t^j), z^j \neq t^j\}$ are not observed. This unobservability is problematic because we would like to make inferences about the function $y^j(t^j)$. Instead, the observed data allow us to measure only the probability distributions $\text{P}[y|z = t]$ and $\text{P}[z = t]$. Manski (2011, 9) shows that if the unit's treatment response is a function only of the value of its own treatment and not of the treatment realized for any other unit—the Individualistic Treatment Response (ITR) assumption—then the probability distributions for the outcome functions can be identified in the region

$$H\{\text{P}[y[t^J]]\} = [\text{P}[y|z = t]\text{P}[z = t] + \delta\text{P}[z \neq t],\ \delta \in \Delta_Y],$$

where t^J denotes the vector of treatment assignments over the entire population J and Δ_Y denotes the space of all probability distributions on y. The term $\delta\text{P}[z \neq t]$, $\delta \in \Delta_Y$, represents that the outcome is unknown whenever the realized treatment is not the same as the potential treatment. With binary outcomes the identification region for the potential outcome distribution under the treatment t is an interval, $[\text{P}[y|z = t]\text{P}[z = t], \text{P}[y|z = t]\text{P}[z = t] + \text{P}[z \neq t]]$. ITR is a close analogue to the stable unit treatment value assumption in the Rubin (1978, 1991) Causal Model.

B Identification Bounds with Missing Treatment Data

Molinari (2010) extends Manski bounds to develop a procedure for identifying treatment effects when treatment data are missing.[1] The notation is the same as above, except Molinari introduces a new variable, d, which is a binary variable equal to 1 if the treatment received by an unit is observed, 0 otherwise. The analyst knows two distributions: $\mathrm{P}[d = 1]$, the probability that the treatment is observed and $\mathrm{P}[z|y, d = 1]$, the conditional distribution of realized treatments given realized outcomes and the observability of the realized treatment.

Given this setup, what can be learned about the ATE, $\mathrm{P}[y_1 = 1] - \mathrm{P}[y_0 = 1]$? Using the law of total probability, the ATE can be decomposed as

$$\begin{aligned}\mathrm{P}[y_1 = 1] - \mathrm{P}[y_0 = 1] = &(\mathrm{P}[y_1 = 1|d = 1] - \mathrm{P}[y_0 = 1|d = 1])\,\mathrm{P}[d = 1] \\ &+ (\mathrm{P}[y_1 = 1|d = 0] - \mathrm{P}[y_0 = 1|d = 0])\,\mathrm{P}[d = 0]. \qquad (1)\end{aligned}$$

Notice that (1) has two components, an observability of treatment component, $[d = 1]$, and an unobservability of treatment component, $[d = 0]$. Using again the law of total probability, Molinari (citing Manski 1995, Chapter 2) identifies the sharp lower and upper bounds for the observability component:

$$LB_{TE}^{d=1} \leq \mathrm{P}[y_1 = 1|d = 1] - \mathrm{P}[y_0 = 1|d = 1] \leq UB_{TE}^{d=1}, \qquad (2)$$

[1] Molinari's method for treatment effect identification with missing treatment data assumes Individualistic Treatment Response (ITR): the unit's treatment response is a function only of the value of its own treatment and not of the treatment realized for any other unit, which is a close analogue to the stable-unit-treatment-value-assumption (SUTVA) in the Rubin (1978, 1991) Causal Model. Molinari assumes that each member j of a population has "a specific response function $y_j(\cdot) : T \to Y$ mapping treatments $t \in T$ into outcomes" (Molinari, 2010, 83) and so focuses on individualistic effects. Molinari's method and the sensitivity tests we introduce may still be useful in the absence of ITR if the data can be arranged so that an assessment of individualistic treatment effects is meaningful. In this case all of $\mathrm{P}[y(t^J)]$ will not be identifiable.

where

$$
\begin{aligned}
LB_{TE}^{d=1} &= \mathrm{P}[y=1|d=1,z=1]\mathrm{P}[z=1|d=1] \\
&\quad - \mathrm{P}[y=1|d=1,z=0]\mathrm{P}[z=0|d=1] - \mathrm{P}[z=1|d=1]\,, \\
UB_{TE}^{d=1} &= \mathrm{P}[y=1|d=1,z=1]\mathrm{P}[z=1|d=1] \\
&\quad + \mathrm{P}[z=0|d=1] - \mathrm{P}[y=1|d=1,z=0]\mathrm{P}[z=0|d=1].
\end{aligned}
$$

These are simply Manski bounds for the data for which the treatment is observed.

What Manski (1995) does not derive is the upper and lower bounds for the unobservable component. The key probability is $\mathrm{P}[z=1|d=0]$, which cannot be derived from the data. Without knowledge of $\mathrm{P}[z=1|d=0]$, Molinari states that the sharp bounds for $\mathrm{P}[y_1=1|d=0]-\mathrm{P}[y_0=1|d=0]$ are $-1 \leq \mathrm{P}[y_1=1|d=0]-\mathrm{P}[y_0=1|d=0] \leq 1$, which are not informative. Using this result with equation (1), the sharp bounds for the treatment effect in the absence of knowledge of $\mathrm{P}[z=1|d=0]$ are

$$
LB_{TE} \leq \mathrm{P}[y_1=1] - \mathrm{P}[y_0=1] \leq UB_{TE},
$$

where

$$
\begin{aligned}
LB_{TE} &= -\mathrm{P}[d=0] - \mathrm{P}[d=1]\,(\mathrm{P}[z=1|d=1] \\
&\quad + \mathrm{P}[y=1|d=1,\ z=0]\mathrm{P}[z=0|d=1] - \mathrm{P}[y=1|d=1,\ z=1]\mathrm{P}[z=1|d=1])\,, \\
UB_{TE} &= \mathrm{P}[d=0] + \mathrm{P}[d=1]\,(\mathrm{P}[z=0|d=1] \\
&\quad + \mathrm{P}[y=1|d=1,\ z=1]\mathrm{P}[z=1|d=1] - \mathrm{P}[y=1|d=1,\ z=0]\mathrm{P}[z=0|d=1])\,.
\end{aligned}
$$

C A Simple Example

To provide clear intuition for how Manski and Molinari bounds are calculated, consider Table 1. For each unit A, B, C, D, and E, Table 1 reports its outcome and whether it received the treatment, denoted 0 for control and 1 for treatment. Notice that units A and B received the treatment, units C and D did not receive the treatment, and we are uncertain of the treatment status of unit E. Table 2 presents the potential outcomes associated with Table 1. As Table 2 shows, we cannot observe the control outcome for the treated units and we cannot observe the treated outcome for the control units.[2] Missing treatment data are especially pernicious: we are unable to assign the outcome for unit E to either the treatment or control case.

Unit	Treatment	Outcome
A	1	0
B	1	1
C	0	0
D	0	1
E	?	0

Table 1: Simple Example

Unit	Treatment Status	Control Outcome	Treatment Outcome	Smallest Effect	Largest Effect
A	1	?	0	-1	0
B	1	?	1	0	1
C	0	0	?	0	1
D	0	1	?	-1	0
E	?	?	?	-1	1

Table 2: Simple Example: Potential Outcomes

Manski and Molinari bounds fill in the missing potential outcomes (each "?" in Table 2) with the most extreme possible values. For example, since unit A received the treatment and had an observed outcome of 0, then the largest positive effect the treatment can have on unit A is 0. However, it is possible that the treatment could have a negative effect (size $= -1$), meaning that

[2]This is known as "the fundamental problem of causal inference" (Holland, 1986)

in the control case unit A has an outcome of 1, but in case of treatment the outcome is 0. Since unit B is in the treatment group and has an observed outcome of 1, the treatment effect must be either 0 or 1. In other words, the treatment obviously does not eliminate the outcome of 1, but the outcome might have been 1 even without the treatment. Unit C is in the control group and the outcome is zero. It is possible that receiving the treatment will either cause the outcome to be 1 or it will not (the outcome will remain 0). Unit D is also in the control group and the outcome is 1. The treatment will either have no effect (as the outcome is already 1) or it will eliminate the effect (the outcome will become 0, meaning the effect is -1). For unit E, the treatment effect could be either -1, 0 or 1 since we have no information on the potential outcomes.

We may now compute both Manski and Molinari bounds. To compute bounds for the average treatment effect, we need only compute the averages of the smallest and largest possible effects. Manski bounds do so only for units for which the treatment is observed: in this case, units A through D. Referring to Table 2, the ATE for units with observed treatment data can be no smaller than $(-1+0+0-1)/4 = -1/2$ and no larger than $(0+1+1+0)/4 = 1/2$. Unlike Manski bounds, Molinari bounds are computed for the full population. Here the ATE can be no smaller than $(-1+0+0-1-1)/5 = -3/5$ and no larger than $(0+1+1+0+1)/5 = 3/5$. These constitute the lower and upper bounds on the ATE respectively. Molinari bounds thus permit us to draw inferences about causal effects for the full population.

D Derivation of MTR Sensitivity Analysis

$$
\begin{aligned}
LB_{\upsilon} = \omega\{&-\mathbf{P}[d=0|t_{\text{MTR}}=0] - \mathbf{P}[d=1|t_{\text{MTR}}=0](\mathbf{P}[z=1|d=1,t_{\text{MTR}}=0] \\
&+ \mathbf{P}[y=1|d=1,\ z=0,t_{\text{MTR}}=0]\mathbf{P}[z=0|d=1,t_{\text{MTR}}=0] \\
&- \mathbf{P}[y=1|d=1,\ z=1,t_{\text{MTR}}=0]\mathbf{P}[z=1|d=1,t_{\text{MTR}}=0])\} \\
= &-\mathbf{P}[d=0|t_{\text{MTR}}=0]\omega - \mathbf{P}[d=1|t_{\text{MTR}}=0]\mathbf{P}[z=1|d=1,t_{\text{MTR}}=0]\omega \\
&- \mathbf{P}[d=1|t_{\text{MTR}}=0]\mathbf{P}[y=1|d=1,\ z=0,t_{\text{MTR}}=0]\mathbf{P}[z=0|d=1,t_{\text{MTR}}=0]\omega \\
&+ \mathbf{P}[d=1|t_{\text{MTR}}=0]\mathbf{P}[y=1|d=1,\ z=1,t_{\text{MTR}}=0]\mathbf{P}[z=1|d=1,t_{\text{MTR}}=0]\omega \\
= &-\mathbf{P}[d=0\cap t_{\text{MTR}}=0] - \mathbf{P}[z=1\cap d=1\cap t_{\text{MTR}}=0] \\
&- \mathbf{P}[y=1\cap d=1\cap z=0\cap t_{\text{MTR}}=0] + \mathbf{P}[y=1\cap d=1\cap z=1\cap t_{\text{MTR}}=0] \\
= &-\mathbf{P}[d=0\cap t_{\text{MTR}}=0] \\
&- \mathbf{P}[y=1\cap z=1\cap d=1\cap t_{\text{MTR}}=0] - \mathbf{P}[y=0\cap z=1\cap d=1\cap t_{\text{MTR}}=0] \\
&- \mathbf{P}[y=1\cap d=1\cap z=0\cap t_{\text{MTR}}=0] + \mathbf{P}[y=1\cap d=1\cap z=1\cap t_{\text{MTR}}=0] \\
= &-\mathbf{P}[d=0\cap t_{\text{MTR}}=0] \\
&- \mathbf{P}[y=0\cap z=1\cap d=1\cap t_{\text{MTR}}=0] - \mathbf{P}[y=1\cap d=1\cap z=0\cap t_{\text{MTR}}=0] \\
= &-(\upsilon_1\xi_1 + \upsilon_3\xi_3 + \upsilon_4\xi_4)\,,
\end{aligned}
$$

E Derivation of MTS Sensitivity Analysis

For parsimony of notation, we omit explicitly representing that all results are conditional on $\omega = 0$.

$$
\begin{aligned}
UB_\theta = \zeta\{&\mathbf{P}[d=0 \mid t_{\text{MTS}}=0] + (\mathbf{P}[z=0 \mid d=1, t_{\text{MTS}}=0] \\
&+ \mathbf{P}[y=1 \mid d=1,\ z=1, t_{\text{MTS}}=0]\mathbf{P}[z=1 \mid d=1, t_{\text{MTS}}=0] \\
&- \mathbf{P}[y=1 \mid d=1,\ z=0, t_{\text{MTS}}=0]\mathbf{P}[z=0 \mid d=1, t_{\text{MTS}}=0]) \\
&\times \mathbf{P}[d=1 \mid t_{\text{MTS}}=0]\} \\
&+ (1-\zeta)\{(\mathbf{P}[y=1 \mid d=1, z=1, t_{\text{MTS}}=1] \\
&- \mathbf{P}[y=1 \mid d=1, z=0, t_{\text{MTS}}=1])\mathbf{P}[d=1 \mid t_{\text{MTS}}=1] + \mathbf{P}[d=0 \mid t_{\text{MTS}}=1]\} \\
&= \kappa_1 + \kappa_2
\end{aligned}
$$

$$
\begin{aligned}
\kappa_1 = \mathbf{P}[t_{MTS}=0]\{&\mathbf{P}[d=0|t_{MTS}=0] + (\mathbf{P}[z=0|d=1, t_{MTS}=0] \\
&+ \mathbf{P}[y=1|d=1, z=1, t_{MTS}=0] \times \mathbf{P}[z=1|d=1, t_{MTS}=0] \\
&- \mathbf{P}[y=1|d=1, z=0, t_{MTS}=0] \times \mathbf{P}[z=0|d=1, t_{MTS}=0]) \\
&\times \mathbf{P}[d=1|t_{MTS}=0]\} \\
= \mathbf{P}[d=0 \cap t_{MTS}=0] &+ (\mathbf{P}[z=0|d=1, t_{MTS}=0] \\
&+ \mathbf{P}[y=1|d=1, z=1, t_{MTS}=0] \times \mathbf{P}[z=1|d=1, t_{MTS}=0] \\
&- \mathbf{P}[y=1|d=1, z=0, t_{MTS}=0] \times \mathbf{P}[z=0|d=1, t_{MTS}=0]) \\
&\times \mathbf{P}[d=1 \cap t_{MTS}=0] \\
= \mathbf{P}[d=0 \cap t_{MTS}=0] &+ \mathbf{P}[z=0 \cap d=1 \cap t_{MTS}=0] \\
&+ \mathbf{P}[y=1 \cap d=1 \cap z=1 \cap t_{MTS}=0] - \mathbf{P}[y=1 \cap d=1 \cap z=0 \cap t_{MTS}=0]
\end{aligned}
$$

$$
\begin{aligned}
&= \mathbf{P}[d = 0 \cap t_{MTS} = 0] \\
&\quad + \mathbf{P}[y = 1 \cap d = 1 \cap z = 0 \cap t_{MTS} = 0] + \mathbf{P}[y = 0 \cap d = 1 \cap z = 0 \cap t_{MTS} = 0] \\
&\quad + \mathbf{P}[y = 1 \cap d = 1 \cap z = 1 \cap t_{MTS} = 0] - \mathbf{P}[y = 1 \cap d = 1 \cap z = 0 \cap t_{MTS} = 0] \\
&= \theta_1\xi_1 + \theta_5\xi_5 + \theta_2\xi_2.
\end{aligned}
$$

$$
\begin{aligned}
\kappa_2 &= \mathbf{P}[t_{MTS} = 1] \times \{(\mathbf{P}[y = 1|d = 1, z = 1, t_{MTS} = 1] \\
&\quad - \mathbf{P}[y = 1|d = 1, z = 0, t_{MTS} = 1]) \times \mathbf{P}[d = 1|t_{MTS} = 1] + \mathbf{P}[d = 0|t_{MTS} = 1]\} \\
&= \mathbf{P}[d = 0 \cap t_{MTS} = 1] + \mathbf{P}[t_{MTS} = 1]\{\mathbf{P}[y = 1|d = 1, z = 1, t_{MTS} = 1]\mathbf{P}[d = 1|t_{MTS} = 1] \\
&\quad - \mathbf{P}[y = 1|d = 1, z = 0, t_{MTS} = 1]\mathbf{P}[d = 1|t_{MTS} = 1]\} \\
&= \mathbf{P}[d = 0 \cap t_{MTS} = 1] + \mathbf{P}[y = 1|d = 1, z = 1, t_{MTS} = 1]\mathbf{P}[d = 1|t_{MTS} = 1]\mathbf{P}[t_{MTS} = 1] \\
&\quad - \mathbf{P}[y = 1|d = 1, z = 0, t_{MTS} = 1]\mathbf{P}[d = 1|t_{MTS} = 1]\mathbf{P}[t_{MTS} = 1]
\end{aligned}
$$

Assume $\mathbf{P}[z = 1|d = 1, t_{MTS} = 1], \mathbf{P}[z = 0|d = 1, t_{MTS} = 1] > 0$.

$$
\begin{aligned}
\kappa_2 &= \mathbf{P}[d = 0 \cap t_{MTS} = 1] \\
&\quad + \mathbf{P}[y = 1 \cap d = 1 \cap z = 1 \cap t_{MTS} = 1]/\mathbf{P}[z = 1|d = 1, t_{MTS} = 1] \\
&\quad - \mathbf{P}[y = 1 \cap d = 1 \cap z = 0 \cap t_{MTS} = 1]/\mathbf{P}[z = 0|d = 1, t_{MTS} = 1] \\
&= (1 - \theta_1) \times \xi_1 + (1 - \theta_2) \times \xi_2/\mathbf{P}[z = 1|d = 1, t_{MTS} = 1] \\
&\quad - (1 - \theta_4) \times \xi_4/\mathbf{P}[z = 0|d = 1, t_{MTS} = 1]
\end{aligned}
$$

Now we derive $\mathrm{P}[z = 1|d = 1, t_{MTS} = 1]$.

$$\begin{aligned}
\mathrm{P}[z = 1|d = 1, t_{MTS} = 1] &= \mathrm{P}[z = 1 \cap d = 1 \cap t_{MTS} = 1]/[\mathrm{P}[d = 1 \cap t_{MTS} = 1]] \\
&= (\mathrm{P}[y = 1 \cap z = 1 \cap d = 1 \cap t_{MTS} = 1] + \mathrm{P}[y = 0 \cap z = 1 \cap d = 1 \cap t_{MTS} = 1])/ \\
&\quad (\mathrm{P}[y = 1 \cap z = 1 \cap d = 1 \cap t_{MTS} = 1] + \mathrm{P}[y = 1 \cap z = 0 \cap d = 1 \cap t_{MTS} = 1] \\
&\quad + \mathrm{P}[y = 0 \cap z = 1 \cap d = 1 \cap t_{MTS} = 1] + \mathrm{P}[y = 0 \cap z = 0 \cap d = 1 \cap t_{MTS} = 1]) \\
&= [(1 - \theta_2)\xi_2 + (1 - \theta_3)\xi_3]/[(1 - \theta_2)\xi_2 + (1 - \theta_3)\xi_3 + (1 - \theta_4)\xi_4 + (1 - \theta_5)\xi_5]
\end{aligned}$$

Similarly,

$$\begin{aligned}
\mathrm{P}[z = 0|d = 1, t_{MTS} = 1] &= [(1 - \theta_4)\xi_4 + (1 - \theta_5)\xi_5]/ \\
&\qquad [(1 - \theta_2)\xi_2 + (1 - \theta_3)\xi_3 + (1 - \theta_4)\xi_4 + (1 - \theta_5)\xi_5]
\end{aligned}$$

Combining,

$$\begin{aligned}
\kappa_2 = (1 - \theta_1) \times \xi_1 + \left[\frac{(1 - \theta_2) \times \xi_2}{[(1 - \theta_2)\xi_2 + (1 - \theta_3)\xi_3]} - \frac{(1 - \theta_4) \times \xi_4}{[(1 - \theta_4)\xi_4 + (1 - \theta_5)\xi_5]} \right] \\
\times [(1 - \theta_2)\xi_2 + (1 - \theta_3)\xi_3 + (1 - \theta_4)\xi_4 + (1 - \theta_5)\xi_5].
\end{aligned}$$

The derivation is complete.

F Sensitivity Analysis Figures with 100,000 Prior Draws

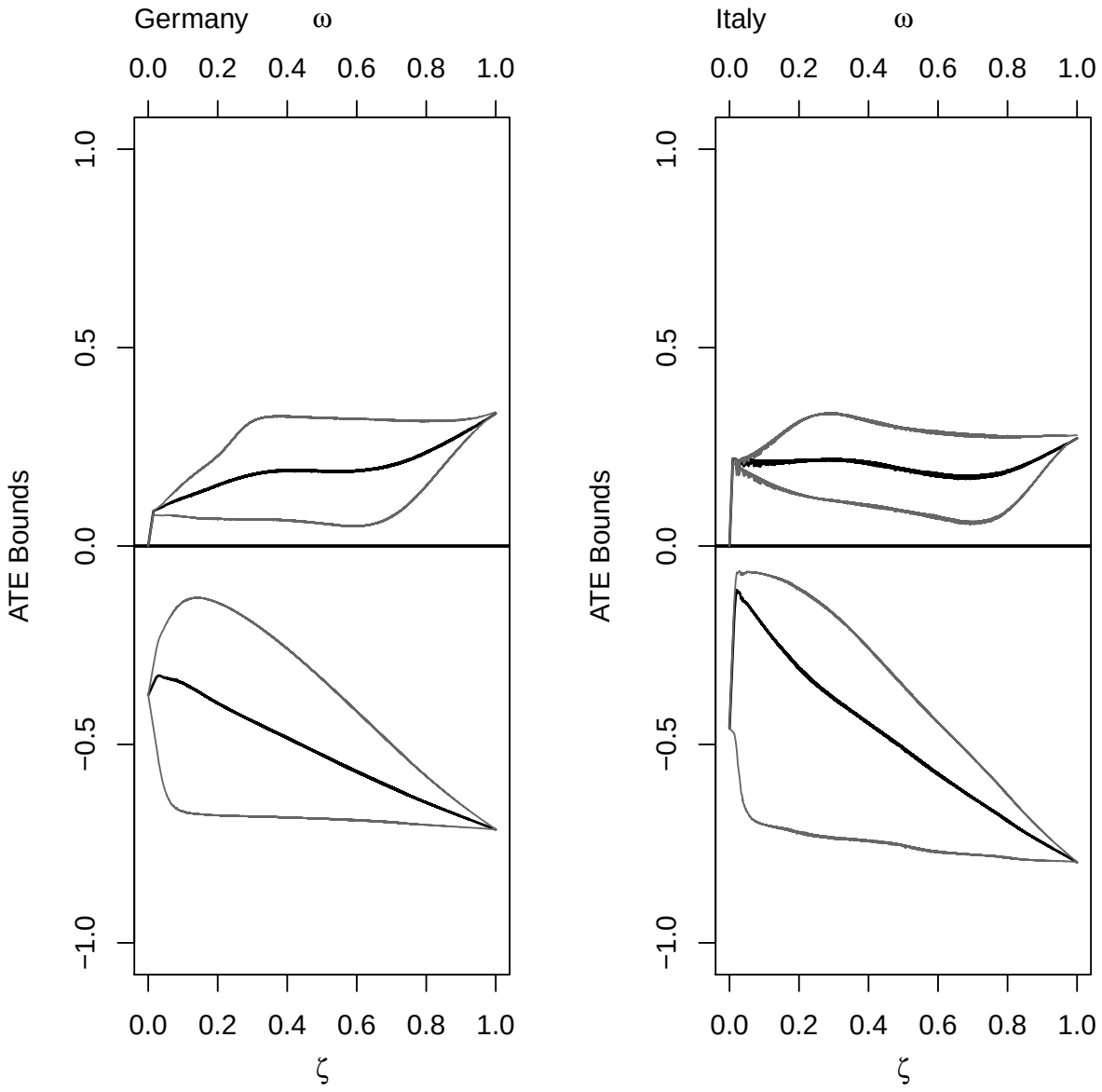

Figure 1: Sensitivity Analysis for Effect of Immigrant Parent on Anti-immigrant policies, German and Italian World Values Survey samples. Solid black lines represent boundaries of $100(1-\alpha)\%$ confidence regions for expected values of bounds assuming uniform priors. Gray lines represent boundaries of $100(1-\alpha)\%$ confidence regions for 95% posterior intervals assuming uniform priors. 100,000 draws from the prior used to simulate posterior distribution. 500 bootstrap replications used to estimate boundaries for $\alpha = .05$.

References

Holland, Paul W. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81:945–960.

Manski, Charles. 1995. *Identification Problems in the Social Sciences*. Cambridge: Harvard University Press.

Manski, Charles. 2011. "Identification of Treatment Response with Social Interactions." Working paper.

Molinari, Francesca. 2010. "Missing Treatments." *Journal of Business and Economic Statistics* 28:82–95.

Rubin, Donald B. 1978. "Bayesian Inference for Causal Effects: The Role of Randomization." *Annals of Statistics* 6:34–58.

Rubin, Donald B. 1991. "Practical Implications of Modes of Statistical Inference for Causal Effects and the Critical Role of the Assignment Mechanism." *Biometrics* 47:1213–1234.