A Supplementary Information

Appendix: Table of Contents

A.1 Proof

Model (1) and Model (4) in the main text are re-stated as follows:

$$Y = \mu + \eta X + \alpha D + \beta DX + \gamma Z + \epsilon; \qquad (1)$$

$$Y = \sum_{j=1}^{3} \{\mu_j + \alpha_j D + \eta_j (X - x_j) + \beta_j (X - x_j) D\} G_j + \gamma Z + \epsilon. \qquad (4)$$

It is to be proved that, if Model (1) is correct :

$$\hat{\alpha}_j - (\hat{\alpha} + \hat{\beta} x_j) \xrightarrow{p} 0, \quad j = 1, 2, 3,$$

in which $\hat{\alpha}$ and $\hat{\beta}$ are estimated from Model (1) and $\hat{\alpha}_j$ are estimated from Model (4).

Proof: First, rewrite Model (4) as:

$$Y = \sum_{j=1}^{3} \{(\mu_j - \eta x_j) + \eta_j X + (\alpha_j - \beta_j x_j) D + \beta_j DX\} G_j + \gamma Z + \epsilon \qquad (6)$$

and define $\underline{\alpha}_j = \alpha_j - \beta_j x_j$. When Model (1) is correct, if we regress Y on G_j, XG_j, DG_j, XDG_j $(j = 1, 2, 3)$ and Z, we have:

$$\underline{\hat{\alpha}}_j \xrightarrow{p} \alpha \text{ and } \hat{\beta}_j \xrightarrow{p} \beta, \quad j = 1, 2, 3.$$

Since $\underline{\hat{\alpha}}_j = \hat{\alpha}_j - \hat{\beta}_j x_j$, we have: $\hat{\alpha}_j \xrightarrow{p} \alpha - \beta x_j$. Because

$$\hat{\alpha} \xrightarrow{p} \alpha \text{ and } \hat{\beta} \xrightarrow{p} \beta$$

when Model (1) is correct, we have:

$$\hat{\alpha}_j - (\hat{\alpha} + \hat{\beta} x_j) \xrightarrow{p} 0 \quad j = 1, 2, 3.$$

Q.E.D.

A.2 Additional information on replication files

TABLE A1. REPLICATION RESULTS

Study	Journal	Not Rejecting Same Effect at Low vs. High	Severe Extra-polation	Rejecting Linear Model	Overall Score
Adams et al. (2006)	AJPS	0	1	0	1
Aklin and Urpelainen (2013)	AJPS	1	1	1	3
Aklin and Urpelainen (2013)	AJPS	1	1	1	3
Banks and Valentino (2012)	AJPS	0	0	0	0
Banks and Valentino (2012)	AJPS	1	1	0	2
Banks and Valentino (2012)	AJPS	1	1	0	2
Bodea and Hicks (2015*a*)	JOP	0	0	1	1
Bodea and Hicks (2015*a*)	JOP	0	1	1	2
Bodea and Hicks (2015*b*)	IO	1	1	0	2
Bodea and Hicks (2015*b*)	IO	1	0		
Bodea and Hicks (2015*b*)	IO	1	1		
Bodea and Hicks (2015*b*)	IO	1	0		
Carpenter and Moore (2014)	APSR	0	1	1	2
Chapman (2009)	IO		1		
Clark and Golder (2006)	CPS	1	0	1	2
Clark and Golder (2006)	CPS	0	0	1	1
Clark and Golder (2006)	CPS	0	0	1	1
Clark and Golder (2006)	CPS		1	1	
Clark and Leiter (2014)	CPS	1	1	0	2
Hellwig and Samuels (2007)	CPS	1	0	0	1
Hellwig and Samuels (2007)	CPS	1	1	0	2
Hicken and Simmons (2008)	AJPS	1	0	0	1
Huddy, Mason and Aarøe (2015)	APSR	0	0	0	0
Huddy, Mason and Aarøe (2015)	APSR	0	0	0	0
Kim and LeVeck (2013)	APSR	0	0	1	1
Kim and LeVeck (2013)	APSR	1	0	1	2
Kim and LeVeck (2013)	APSR	0	0	1	1
Malesky, Schuler and Tran (2012)	APSR	1	1	1	3
Malesky, Schuler and Tran (2012)	APSR	1	1	1	3
Malesky, Schuler and Tran (2012)	APSR	1	1	0	2
Malesky, Schuler and Tran (2012)	APSR	1	1	1	3
Neblo et al. (2010)	APSR	1	0	1	2
Pelc (2011)	IO	0	1	1	2
Pelc (2011)	IO	1	1	1	3
Petersen and Aarøe (2013)	APSR	1	0	0	1
Petersen and Aarøe (2013)	APSR	1	0	0	1
Somer-Topcu (2009)	JOP	1	0	0	1
Tavits (2008)	CPS	0	0	0	0
Truex (2014)	APSR	1	0	1	2
Truex (2014)	APSR	1	1	1	3
Truex (2014)	APSR	1	0	1	2
Truex (2014)	APSR	1	1	0	2
Vernby (2013)	AJPS	1	1	0	2
Vernby (2013)	AJPS	1	1	0	2
Williams (2011)	CPS	1	0	0	1
Williams (2011)	CPS	1	0	1	2

Note that missing values are due to restrictions in the data, such as lack of common support, that prevented the test from being conducted. In such cases an aggregate score was not computed.

A.3 GAM Plot

In cases where both D and X are continuous, an alternative to the scatterplot is to use a generalized additive model (GAM) to plot the surface that describes how the average Y changes across D and X. While the statistical theory underlying GAMs is a bit more involved (Hastie and Tibshirani 1986), the plots of the GAM surface can be easily constructed using canned routines in `R`. Figure A1 shows such a GAM plot for the simulated data from the second sample looking at the surface from four distinctive directions. Lighter color on the surface represents a higher value of Y.

Figure A1 has several features. First, it is obvious that holding X constant, Y is increasing in D and holding D constant, Y is increasing in X. Second, the slope of Y on D is larger with higher X than with lower X. Third, the surface of Y over D and X is fairly smooth, with a gentle curvature in the middle but devoid of drastic humps, wrinkles, or holes. In the Online Appendix, we will see that the GAM plots of examples that likely violate the linearity assumption look quite different from Figure A1.

FIGURE A1. GAM PLOT: SIMULATED SAMPLE WITH CONTINUOUS TREATMENT