The Millennium Development Goals and Education: Accountability and Substitution in Global Assessment

James H. Bisbee[1], James R. Hollyer[2], B. Peter Rosendorff[3], and James Raymond Vreeland[4]

[1]New York University

[2]University of Minnesota

[3]New York University

A Formal Appendix

A.1 Proof of Lemma 1

Proof. We first characterize the equilibrium to this game. C must induce a defection by the representative member of the inning coalition to displace L. C's strategy amounts to maximizing the utility of a representative member of the winning coalition subject to the budget constraint. (Selectors outside of the winning coalition are irrelevant to displacing L since they receive no individual transfers.) The budget constraint is then $p_C^* + h_C^* + \omega t_C^* \leq B$. The conditions imposed on $U(\cdot)$ imply that an interior solution to this problem must exist, in which $U_p'(p_C^*) = U_h'(h_C^*) = \frac{U_t'(t_C^*)}{\omega}$.

Leader: The game is symmetric in each period of play, implying that any leader always chooses identical values of p, h and t. The continuation value for the selector is therefore contingent only on whether she is in the winning coalition. Let V_I denote this value for being in the coalition, and V_O for out of the coalition. Recall that ω is the probability that an individual's support is needed by the leader to remain in office – so the present value of unseating L to an agent that is in the current winning coalition is:

$$\begin{aligned} U_p(p_C^*) + U_h(h_C^*) + U_t(t_C^*) + \delta[\omega V_I + (1-\omega)V_O] \\ = U_C^* + \delta[\omega V_I + (1-\omega)V_O] \end{aligned}$$

using Definition 1 above. The problem for the incumbent is therefore to maximize the residual $B-p-h-\omega t$ subject to the retention constraint, $V_I \geq U_C^* + \delta[\omega V_I + (1-\omega)V_O]$. Notice further that V_I and V_O differ only insofar as coalition members receive private transfers (public goods are enjoyed by all). So, $V_O = V_I - \frac{1}{1-\delta}U_t(t^*)$. Substituting, $\delta[\omega V_I + (1-\omega)V_O] = \delta V_I - \frac{\delta}{1-\delta}(1-\omega)U_t(t^*)$, the retention constraint can be expressed

$$(1-\delta)V_I \geq U_C^* - \frac{\delta}{1-\delta}(1-\omega)U_t(t^*)$$

Since L's promises are credible, $V_I = \frac{1}{1-\delta}[U_p(p^*) + U_h(h^*) + U_t(t^*)]$ where p^* and h^* are the equilibrium offers of primary and secondary education by L. Substituting into the above, we have:

$$U_p(p^*) + U_h(h^*) + U_t(t^*)[1 + \frac{\delta}{1-\delta}(1-\omega)] \geq U_C^*$$

Which we call the "retention constraint". The leader's problem can be expressed

$$\max_{p,h,t} B - p - h - \omega t \text{ subject to } U_p(p^*) + U_h(h^*) + U_t(t^*)[1 + \frac{\delta}{1-\delta}(1-\omega)] \geq U_C^*$$

The first order condition implicitly gives us (p^*, t^*):

$$\omega U_p'(p^*) = [1 + \frac{\delta}{1-\delta}(1-\omega)]U_t'(t^*) \tag{1}$$

and h^* is available by symmetry with $p^* = h^*$. Rewriting equation 1 we have $\frac{U_p'(p^*)}{U_t'(t^*)} = \frac{\frac{1}{\omega}-\delta}{1-\delta}$, and as ω rises, the the ratio of t^* to p^* falls (from the concavity of $U(\cdot)$) – implying a substitution away from private towards public goods. $\square$

A.2 Proof of Lemma 2

Proof. Setting up the Lagrangian for the optimization problem facing the leader:

$$\mathcal{L} = B - p - h - \omega t + \lambda\left[U_p(p^*) + U_h(h^*) + U_t(t^*)[1 + \frac{\delta}{1-\delta}(1-\omega)] - U_C^*\right]$$

Taking first order conditions, setting up the Jacobian yields

$$\begin{bmatrix} \lambda U_p'' & 0 & 0 & U_p \\ 0 & \lambda U_h'' & 0 & U_h' \\ 0 & 0 & \lambda U_t''[1+\frac{\delta}{1-\delta}(1-\omega)] & U_t'[1+\frac{\delta}{1-\delta}(1-\omega)] \\ U_p' & U_h' & U_t'[1+\frac{\delta}{1-\delta}(1-\omega)] & 0 \end{bmatrix} \begin{bmatrix} \frac{dp^*}{d\omega} \\ \frac{dh^*}{d\omega} \\ \frac{dt^*}{d\omega} \\ \frac{d\lambda}{d\omega} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1+\lambda\frac{\delta}{1-\delta}U_t' \\ \frac{\delta}{1-\delta}U_t + dU_C^*/d\omega \end{bmatrix}$$

Applying Cramer's rule, and noticing that $\lambda = 1/U_p'$, and with the standard convexity assumptions on U, and $\delta, \omega \in (0,1)$, we get $\frac{dp^*}{d\omega} > 0$, $\frac{dh^*}{d\omega} > 0$, $\frac{dt^*}{d\omega} < 0$. $\square$

A.3 Proof of Lemma 3

Proof. Consider the triple (p^*, h^*, t^*). By definition 1, this triple satisfies the first two constraints, retention and residual. Substituting into the participation constraint, we have $h^* + \omega t^* \leq h^* + p^* - \bar{p} + \omega t^* + \beta$, or $\beta \geq \bar{p} - p^*$, a sufficient condition. $\square$

A.4 Proof of Proposition 1

Proof. The challenger has a dominant strategy as before, and follows the same strategy as the No MDG game. If $\bar{p} \leq p^*$, the incentives for L are unaffected by the MDGs. She will simply set $p = p^*$ as before and gain the international benefit β for doing so.

For $\bar{p} > \hat{p}$, from Lemma 3 we know that there is no triple $(\bar{p}, h, t)$ in the feasible set Θ. That is the resources on offer to adopt the MDGs cannot satisfy all three (residual, retention and participation) constraints, and hence the leader rejects the benefits that go along with the MDGs. Instead the leader chooses the same optimal allocation as in the No MDG game. For $\bar{p} \in (p^*, \hat{p})$, we know from Lemma 3 that all three constraints are satisfied. Moreover, from the monotonicity and additive separability of the utility functions, there is no benefit from offering more public goods p than is minimally required to obtain the bonus β. Hence the leader is optimizing at $\bar{p}$; by the definition of $\bar{t}, \bar{h}$, these are optimal choices in the presence of the MDG goal, $\bar{p}$. □

A.5 Proof of Proposition 2

Proof. Proposition 1 holds that $p^{**} = p^*$ when $\bar{p} \notin (p^*, \hat{p})$ and equals $\bar{p}$ otherwise. Thus, $p^{**} - p^*$ is either strictly positive or equal to zero. From Lemma 2, we know that p^* is rising in ω. For any (exogenous) $\bar{p}$, $\bar{p} - p^*$ is falling in ω. □

A.6 Proof of Proposition 3

Proof. In both the no MDG and the MDG case the respective retention constraints are satisfied at equality:

$$U_h(h^*) + (\frac{1-\omega}{1-\delta})U_t(t^*) + U_p(p^*) - U_C^* = 0$$
$$U_h(h^{**}) + (\frac{1-\omega}{1-\delta})U_t(t^{**}) + U_p(p^{**}) - U_C^* = 0$$

Now when $\bar{p} > p^*$, then by concavity and additive seperability it must be that $h^{**} = \bar{h} < h^*$ and $t^{**} = \bar{t} < t^*$. Where $p^{**} = p^*$, the constraints are identical to each other; hence $h^{**} = h^*$. From Proposition 1, this is true for $\bar{p} \notin (p^*, \hat{p})$. Hence whenever $p^{**} > p^*$, we have $h^{**} < h^*$. □

A.7 Proof of Lemma 4

Proof. Define $\hat{\omega}$ implicitly as the value of ω such that $p^* = \bar{p}$. Lemma 2 establishes that p^* is monotonic and increasing in ω. Assumption 1 holds that $\bar{p} < \hat{p} \,\forall\, \omega$. Hence, $\hat{\omega}$ is well defined.

Proposition 1 establishes that $p^{**} = \bar{p}$ if $p^* < \bar{p}$ and $p^{**} = p^*$ otherwise. Given Lemma 2, $p^{**} = \bar{p}$ for $\omega < \hat{\omega}$ and $p^{**} = p^*$ otherwise. Proposition 3 then establishes that the substitution effect must exist for $\omega < \hat{\omega}$ and does not exist otherwise. □

A.8 Model Extension: Imperfect Commitment

Consider an infinite horizon discrete time model isomorphic to that described above. However, in this model extension, at the beginning of each period $t > 1$, a randomly selected set of winning coalition members of mass η are removed from the winning coalition and replaced by a randomly selected set of selectors not in the winning coalition of equal mass. Denote $\frac{\eta}{\omega} \equiv q_W$ and $\frac{\eta}{1-\omega} \equiv q_S$.

As before, the history of play is summarized by the state variable $\sigma_{i,t} \in \{0,1\}$ which assumes the value 1 if selector i is in the winning coalition. However, the transition probabilities for this variable are now different. Define these transition probabilities as follows: $Pr(\sigma_{i,t} = 1|\sigma_{i,t-1} = 1, retain_{t-1}) = 1 - q_W$; $Pr(\sigma_{i,t} = 0|\sigma_{i,t-1} = 1, retain_{t-1}) = q_W$; $Pr(\sigma_{i,t} = 1|\sigma_{i,t-1} = 0, retain_{t-1}) = q_S$; $Pr(\sigma_{i,t} = 0|\sigma_{i,t-1} = 0, retain_{t-1}) = 1 - q_S$; $Pr(\sigma_{i,t} = 1|remove_{t-1}) = \omega \; \forall \; \sigma_{i,t-1}$, where $\{retain, remove\}$ refers to the winning coalition's decision to remove or retain the incumbent L. Assume that $1 - \omega > q_W$ – i.e., that members of the winning coalition have a greater chance of retaining their current positions by sticking with the incumbent than when siding with the challenger. This discrepancy captures the credibility advantage of the incumbent, which diminishes continuously as $q_W \to 1 - \omega$.

As before, we search for a Markov perfect equilibrium:

Lemma 5. *An equilibrium to this game is the pair of triples,* (p^*_C, h^*_C, t^*_C) *for the challenger and* (p^*, h^*, t^*) *for the leader. Define* $U^*_C = U_p(p^*_C) + U_h(h^*_C) + U_t(t^*_C)$.

Proof. Notice, moreover, that the challenger's problem is unchanged relative to the above. The equilibrium strategy for the challenger (p^*_C, h^*_C, t^*_C) is thus unchanged, as is the value of U^*_C.

However, the continuation values for the selectors have been changed owing to the redefined transition probabilities, implying that L's best response will also change. For notational simplicity, define $U_p(p^*) + U_h(h^*) \equiv X$ and $U_t(t^*) \equiv Y$. We can now set up the following Bellman equations:

$$
\begin{aligned}
V_I &= X + Y + \delta[(1 - q_w)V_I + q_w V_O] \\
&= \frac{X + Y + \delta q_w V_O}{1 - \delta(1 - q_w)} \\
V_O &= X + \delta[(1 - q_S)V_O + q_S V_I] \\
&= \frac{X + \delta q_s V_I}{1 - \delta(1 - q_S)}
\end{aligned}
$$

Substituting:

$$
V_I = \frac{[1 - \delta + \delta(q_S + q_W)]X + (1 - \delta + \delta q_S)Y}{[1 - \delta(1 - q_W)][1 - \delta(1 - q_S)] - \delta^2 q_S q_W}
$$

Substituting again:

$$
V_O = \frac{[1 - \delta + \delta(q_S + q_W)]X + \delta q_S Y}{[1 - \delta(1 - q_W)][1 - \delta(1 - q_S)] - \delta^2 q_S q_W}
$$

For each member of the winning coalition, it is a best response to retain L iff:

$$V_I \geq U_C^* + \delta[\omega V_I + (1-\omega)V_O]$$

from the above:

$$\delta[\omega V_I + (1-\omega)V_O] = \delta V_I - \frac{\delta(1-\omega)(1-\delta)U_t(t^*)}{[1-\delta(1-q_W)][1-\delta(1-q_S)] - \delta^2 q_W q_S}$$

substituting:

$$(1-\delta)V_I \geq U_C^* - \frac{\delta(1-\omega)(1-\delta)U_t(t^*)}{[1-\delta(1-q_W)][1-\delta(1-q_S)] - \delta^2 q_W q_S}$$

simplifying:

$$U_p(p) + U_h(h) + (\frac{1+\delta(q_S-\omega)}{1-\delta+\delta(q_S+q_W)})U_t(t) \geq U_C^*$$

The equilibrium triple (p^*, h^*, t^*) is thus implicitly given by the values that solve:

$$\max_{p,h,t} B - p - h - \omega t \text{ subject to } U_p(p) + U_h(h) + (\frac{1+\delta(q_S-\omega)}{1-\delta+\delta(q_S+q_W)})U_t(t) \geq U_C^*$$

□

Lemma 6. *In any (Markov Perfect) equilibrium to this game, the fraction $\frac{p^*}{t^*}$ is rising in ω and in η.*

Proof. The first order conditions to L's maximization problem, which implicitly defines (p^*, h^*, t^*) in Lemma 5 imply that:

$$\max_{p,h,t} B - p - h - \omega t \text{ subject to } U_p(p) + U_h(h) + (\frac{1+\delta(q_S-\omega)}{1-\delta+\delta(q_S+q_W)})U_t(t) \geq U_C^*$$

Substituting $q_S = \frac{\eta}{1-\omega}$ and $q_W = \frac{\eta}{\omega}$ into the above yields:

$$\frac{U_p'(p^*)}{U_t'(t^*)} = \frac{(1-\delta\omega)(1-\omega)+\delta\eta}{(1-\delta)\omega(1-\omega)+\delta\eta}$$

The right-hand side of this expression is strictly falling in $\omega \in (0, \frac{1}{2}]$. Given the concavity of $U_k(\cdot)$, this implies that $\frac{p^*}{t^*}$ must be rising in ω over the admissible range of values. Analogously, the right hand side of this expression is strictly falling in η – implying that $\frac{p^*}{t^*}$ is rising in ω. □

Lemma 7. *Comparative statics: p^* rises and t^* falls with ω, η.*

Proof. The structure of this proof follows that of Lemma 2 directly, substituting $[\frac{(1-\delta\omega)\omega(1-\omega)+\delta\omega\eta}{(1-\delta)\omega(1-\omega)+\delta\eta}]$ for $[1+\frac{\delta}{1-\delta}(1-\omega)]$ in the Lagrangian. As this expression is falling in ω, η, and given that $\lambda = \frac{1}{U_p'}$, we have $\frac{\partial p^*}{\partial \omega} > 0$, $\frac{\partial h^*}{\partial \omega} > 0$, $\frac{\partial t^*}{\partial \omega} < 0$, $\frac{\partial p^*}{\partial \eta} > 0$, $\frac{\partial h^*}{\partial \eta} > 0$, and $\frac{\partial t^*}{\partial \eta} < 0$. □

All remaining proofs follow directly from the above. The qualitative conclusions of our comparative statics are unchanged in the extended model.

B Empirical Appendix

B.1 Outliers

In the following section, we present additional results from our preliminary analysis. We begin with an analysis of outliers that potentially drive our main results. We look for evidence of outliers with two related techniques. For the primary and secondary enrollment data, we randomly drop 10% of the dataset over 1,000 simulations and re-estimate model (4) from Tables 3 and 4. We extract the estimated coefficients on the Millennium Declaration, as well as its interaction with transparency and democracy measures. We then plot these coefficients as densities where the x-axis indicates the coefficient estimate. Figure 1 shows these robustness results for primary and secondary enrollment.

[FIGURE 1 ABOUT HERE]

We conduct a very similar analysis for the substitution specification except instead of dropping 10% of the data at random, we drop each country in turn. Figure 2 plots the densities for the Millennium Declaration as well as its interaction with transparency and democracy measures. As above, we are reassured by the robustness of our results to outliers, suggesting that our findings are not driven by a particular country.

[FIGURE 2 ABOUT HERE]

B.2 VAR Analysis

The analysis summarized in the main text uses multi-level models to estimate the relationship between enrollment rates and MDG adoption. While multi-level models are more flexible, these specifications still rely on potentially heroic parametric assumptions.

We relax these assumptions by estimating a highly flexible model of the relationship between primary and secondary enrollment rates and the MDGs via vector autoregression (VAR). This model estimates a series of two equations with, respectively, primary and secondary enrollment rates as the outcome variables. These terms are regressed on (four of) their own lags, (four of) each other's lags, an indicator for the promulgation of the MDGs, and controls (transparency, democracy, and per capita GDP). We thus estimate models of the following form:

$$
\begin{aligned}
P_{i,t} &= \sum_{L=1}^{4}(\beta_L P_{i,t-L} + \gamma_L S_{i,t-L}) + \delta I(year \geq 2000) + \epsilon_{i,t} \\
S_{i,t} &= \sum_{L=1}^{4}(\lambda S_{i,t-L} + \kappa_L P_{i,t-L}) + \mu I(year \geq 2000) + \nu_{i,t}. \qquad (2)
\end{aligned}
$$

We present the estimates from equation 2 graphically as a series of impulse response functions, each of which presents estimates on a 'shock' to the MDG indicator over time. The figures present estimates of the relation between the MDGs and primary and secondary enrollment rates respectively. As can be seen in Figure 3, these highly flexible specifications indicate that primary enrollment rates rise rapidly after the promulgation of the MDGs, while secondary enrollment rates fall. Writ large, the VAR results support our theorized concerns that assessments can prompt unexpected responses. While primary enrollment spikes following MDG adoption, secondary enrollment either declines outright or increases with a muted effect.

[FIGURE 3 ABOUT HERE]

B.3 Sensitivity Analysis

We also explore the extent to which our results are sensitive to confounding. As discussed in the conclusion, the Millennium Development Goals were tied to aid. While we believe the omission of this control makes our estimates conservative, we run a simulation to understand the level of correlation between our treatment (MDG) and outcome (substitution) that would undermine our results. We generate a confounding variable that is correlated with both treatment and outcome and re-estimate the coefficient of interest.[1] Figure 4 presents the results for the Millennium Declaration variable. We note that confounding would have to be very strong (correlations with the substitution effect and MDG more than 10 times as strong as per capita GDP) in order to yield estimates outside of our original 95% confidence interval.

[FIGURE 4 ABOUT HERE]

We re-run this analysis focusing on the interaction coefficients and find even more reassuring results (Figure 5) We note that our coefficient on the interacted relationship for the MDG and transparency is even less sensitive to confounds and the estimate the democracy interaction is totally insensitive. The latter result makes sense given that we do not find a statistically significant relationship in the data for the interacted relationship with democracy and MDGs after controlling for transparency. Taken together, these results constitute reassuring evidence that our main results are not sensitive to omitted variables.

[FIGURE 5 ABOUT HERE]

[1] Imbens 2003

B.4 Causality

In our main results, we find a positive relationship between the Millennium Declaration and primary enrollment, suggesting countries responded to the MDGs. We can frame this result as a causal effect if we appeal to the Conditional Independence Assumption (CIA). This claim relies on the assumption that, after controlling for both observed and unobserved potential confounds via the random effects in the multilevel model, the increase in primary school enrollment rates can be attributed to the release of the MDGs.

There are ample reasons to question this claim. We therefore let the data determine where discontinuous changes (if any) exist in primary and secondary enrollment rates. We do this in two ways: first, we appeal to the intuition behind regression discontinuity designs to see whether the primary enrollment rate exhibits a discontinuous shift at the year 2000, corresponding to the Millennium Declaration. Figure 6 plots the predicted values and 90% confidence intervals estimated using a local polynomial regression of primary enrollment on year where 2000 is the cut-point. As illustrated, we find only weak evidence of a discontinuous shift in the overall data but we do note stronger evidence for countries in the lowest quartiles of transparency and those with the initial primary enrollment rates at the beginning of our period of analysis.

[FIGURE 6 ABOUT HERE]

However, it may be that these discontinuities merely reflect the overall trend during the period of analysis. We augment our approach with Bayesian Change Point analysis to let the data determine when the primary and secondary school enrollment rates exhibited a discontinuous shift. This analysis is applied to each country in isolation, yielding Figure 7 which plots the density of change points over time. We see two separate periods of discontinuous change, one during the late 1980s / early 1990s and a second around the time of the Millennium Summit.

[FIGURE 7 ABOUT HERE]

The suggestion of a shift in primary school enrollment rates corresponding to the period in which the initial conversations regarding what would eventually become the MDGs is interesting and highlights the problematic assertion that the release of the MDGs “caused” changes in enrollment rates. Nevertheless, the appearance of a cluster of change points around the time of the MDGs suggests that the Millennium Declaration corresponded to a change in enrollment rates, at least for a subset of the countries. Notably, this clustering is driven primarily by countries with low levels of primary enrollment (below 70). This suggests that MDGs were particularly impactful on the very countries they were designed to target (i.e., those with low primary enrollment rates).

B.5 Subsets of the Data

Finally, we re-test our main analyses using subsets of the data. We focus only on countries with per capita GDP measures below \$1,000 and \$13,000. Our results are robust to these restrictions in terms of coefficient sign and magnitude although we suffer from small samples in the most restrictive dataset. Nevertheless, we find support for our main findings among the low and middle income countries. The tables on the following pages summarize these findings, along with similar simulations to those presented above.

[FIGURE 8 ABOUT HERE]

References

Imbens, Guido W. 2003. Sensitivity to exogeneity assumptions in program evaluation. *The American Economic Review* 93 (2):126–132.

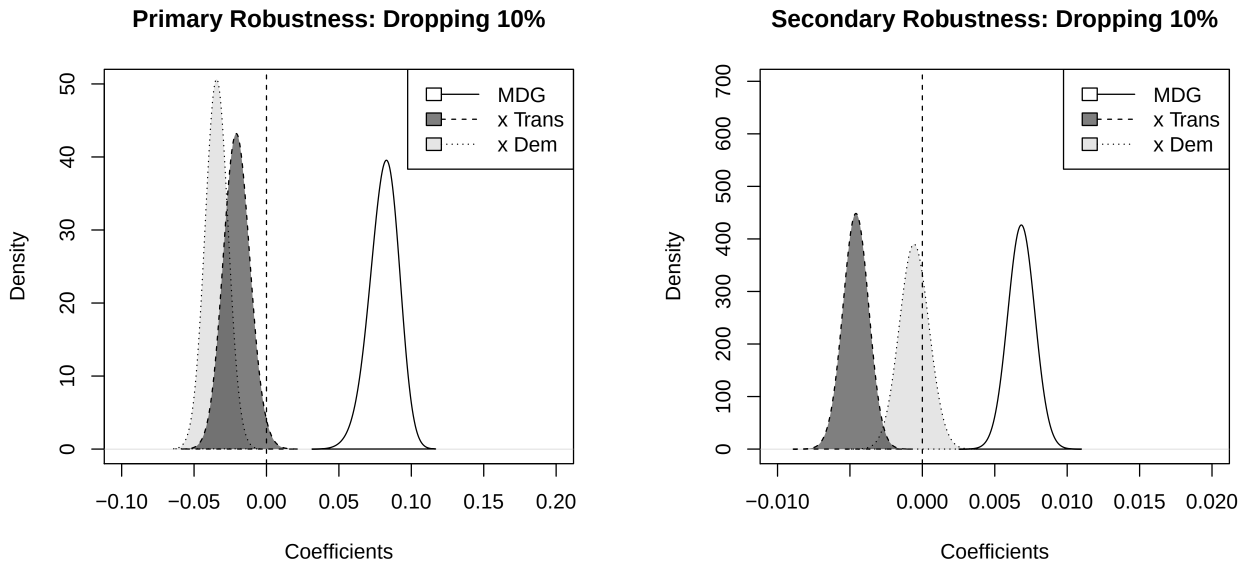

Figure 1: Robustness to outliers. Plots indicate the distribution of coefficient estimates over 1,000 simulations in which 10% of the dataset is dropped at random.

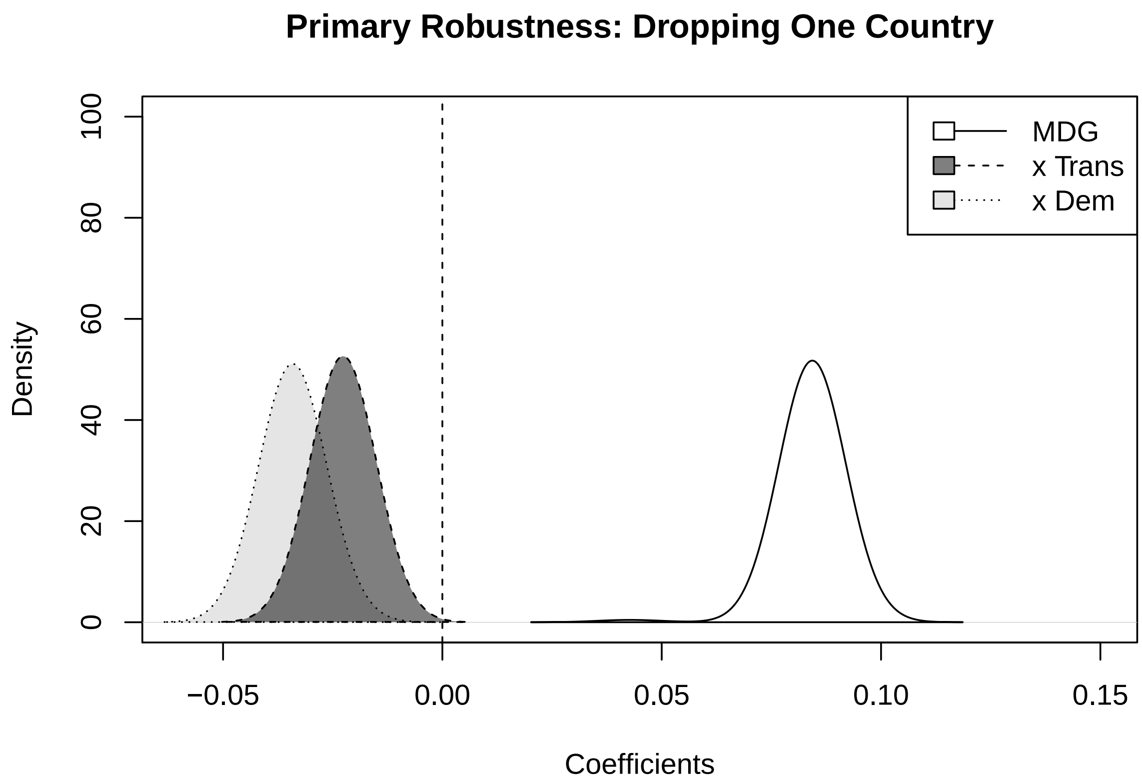

Figure 2: Robustness to outliers. Plots indicate the distribution of coefficient estimates re-estimated dropping one country at a time.

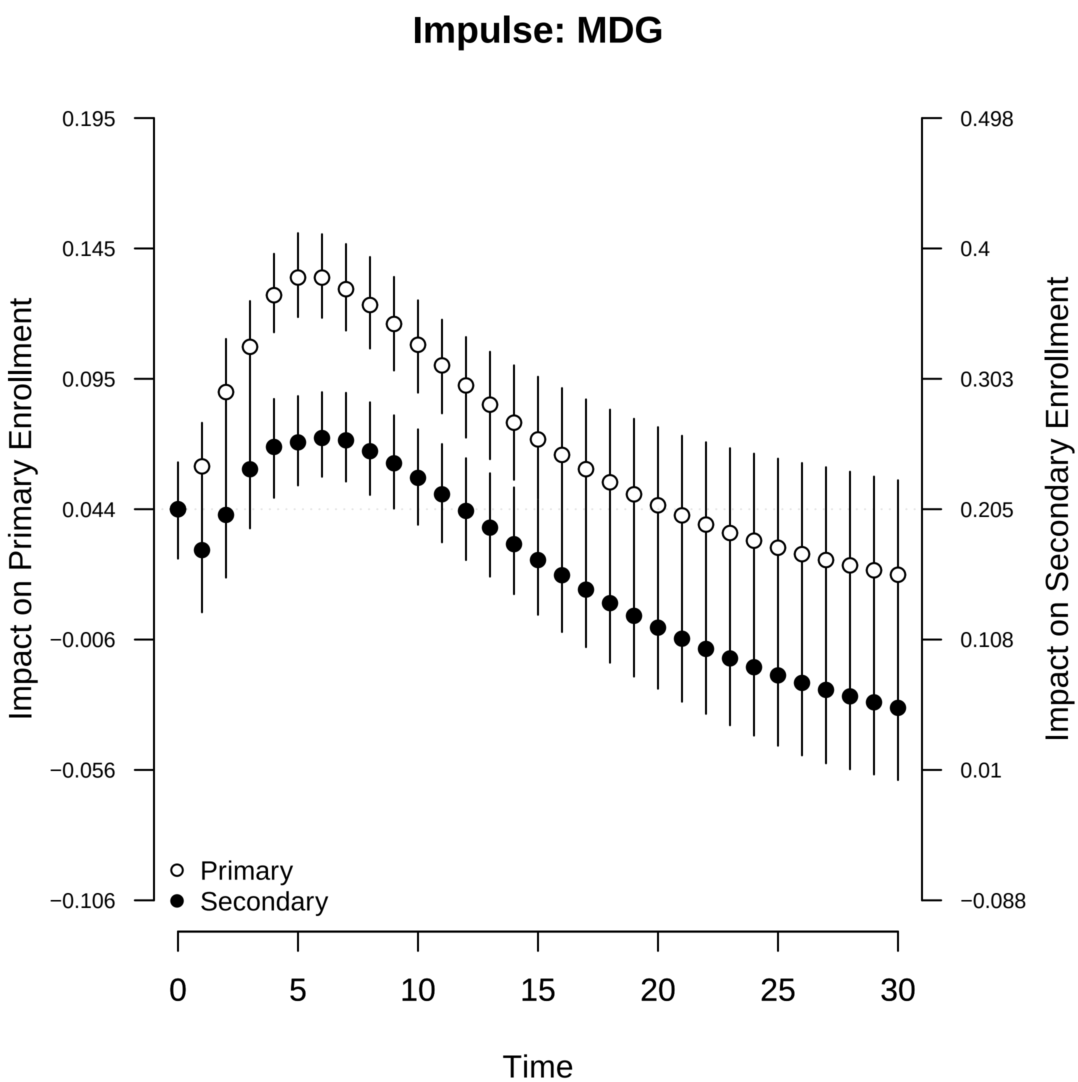

Figure 3: Impulse-response plots for primary (white circles - left axis) and secondary (black circles - right axis) against MDG adoption. Panel VAR estimates produced using Stata's `xtvar` function, controlling for democracy, transparency, and per capita GDP. The horizontal dotted line indicates the $t = 0$ response for reference.

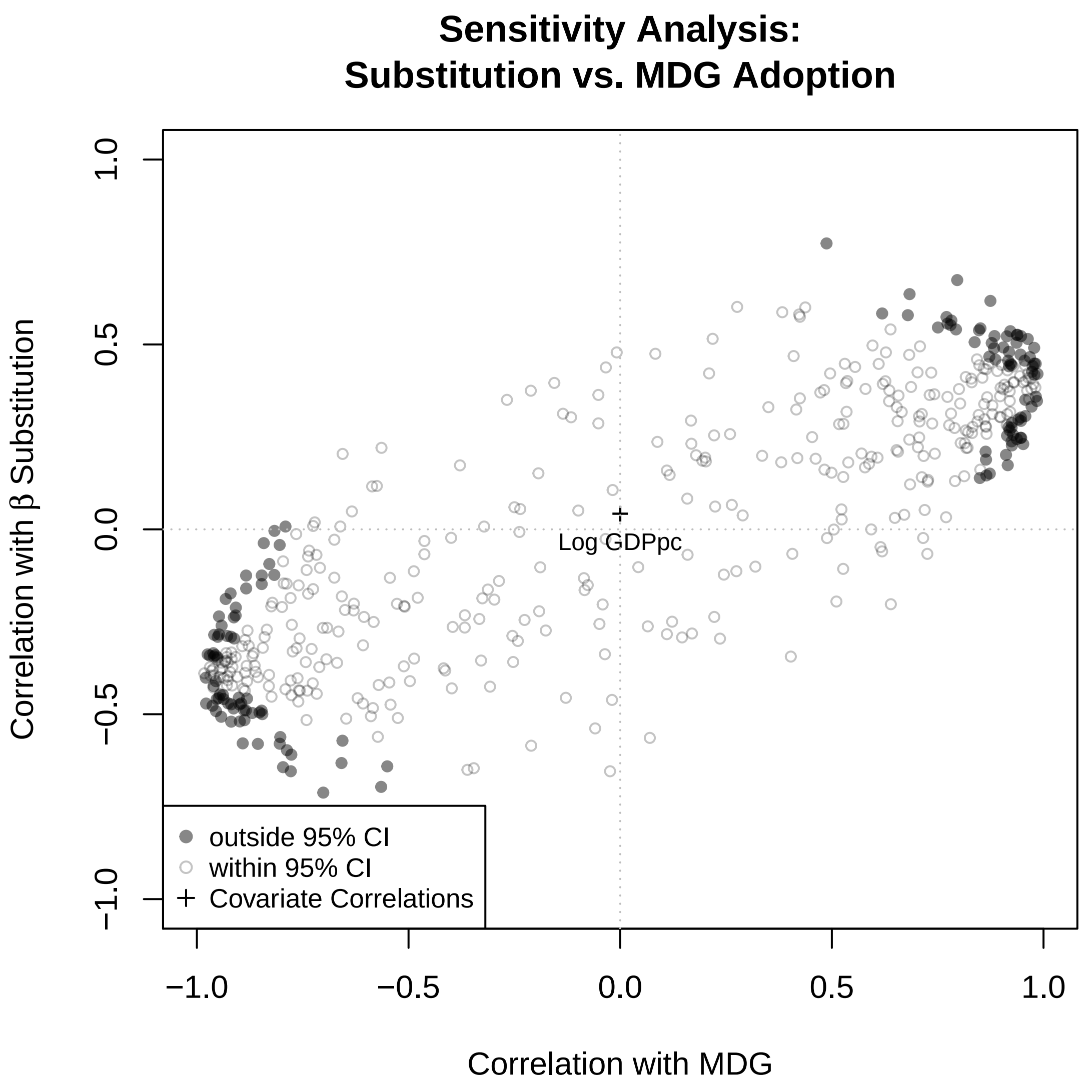

Figure 4: Correlations of simulated confound with treatment (x-axis) and outcome (y-axis). Hollow circles indicate that confounded estimates of the relationship between the MDG indicator and the β substitution measure are within the 95% confidence interval of the original estimate. Light gray circles indicate estimates that are in the bottom 2.5% of the t-distribution while dark gray circles represent those in the top 2.5% (above the 97.5th percentile). We include the control measure of per capita GDP for reference.

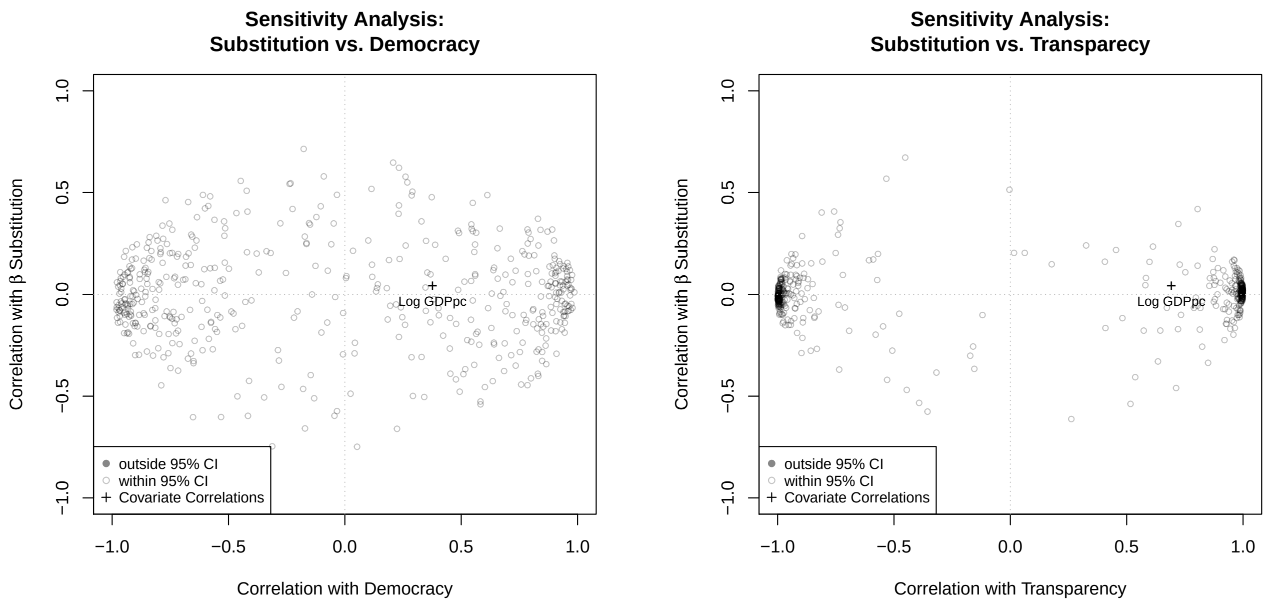

Figure 5: Correlations of simulated confound with treatment (x-axis) and outcome (y-axis). Hollow circles indicate that confounded estimates of the interacted relationship between the MDG indicator and either democracy (left) or transparency (right) and the β substitution measure are within the 95% confidence interval of the original estimate. Light gray circles indicate estimates that are in the bottom 2.5% of the t-distribution while dark gray circles represent those in the top 2.5% (above the 97.5th percentile). We include the control measure of per capita GDP for reference.

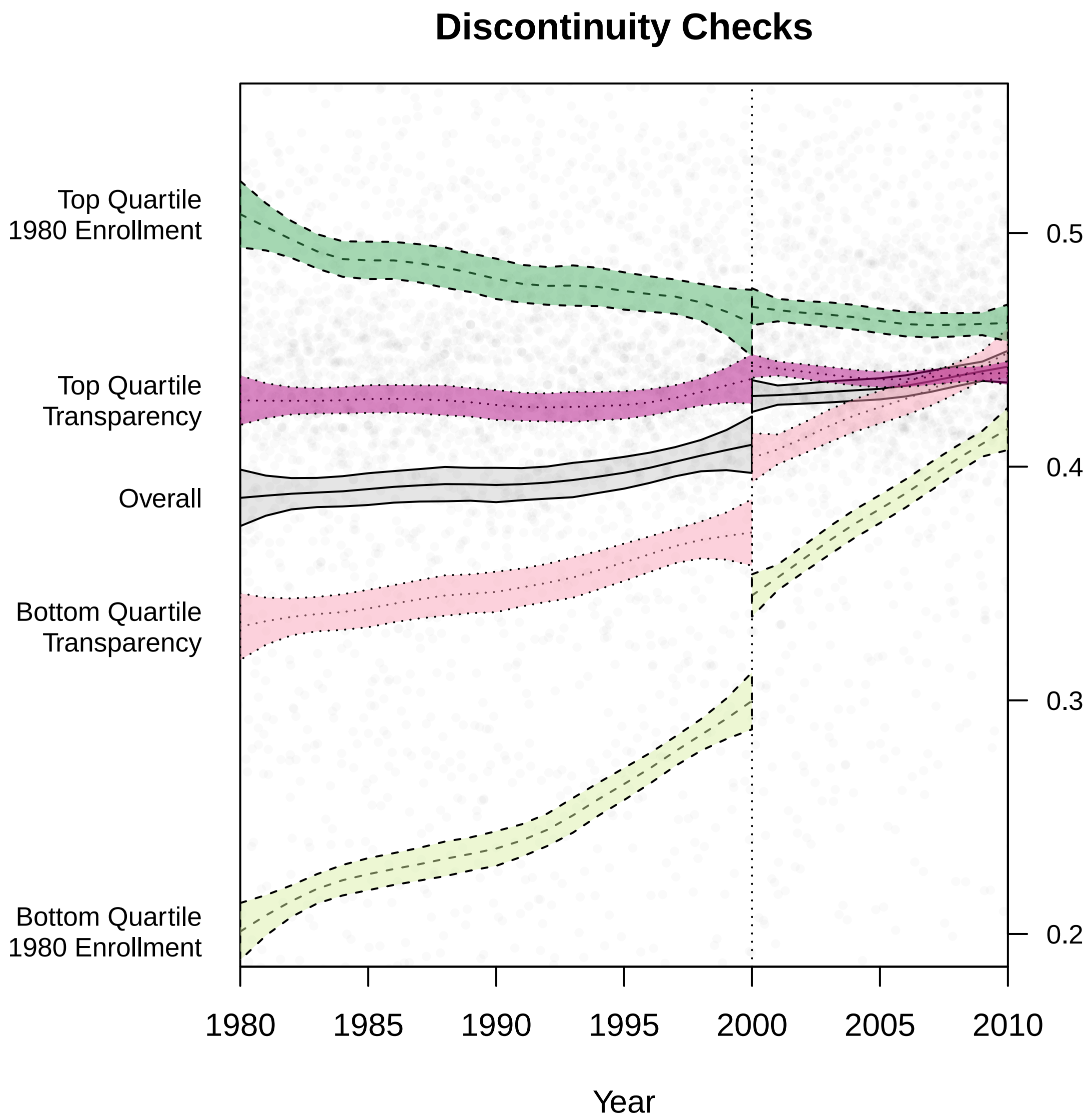

Figure 6: Local polynomial smoothers applied to pre- and post-MDG adoption for subsets of the data.

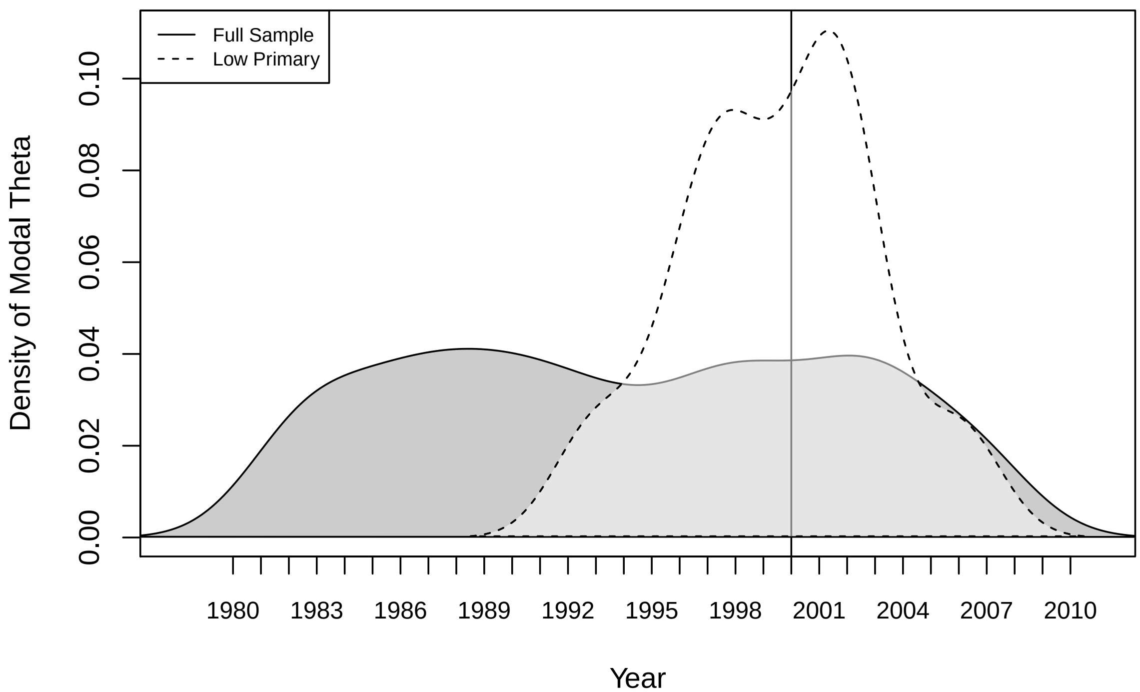

Figure 7: Distribution of change-points for full sample (gray density with solid border) and low primary enrollment only (white density with dashed border).

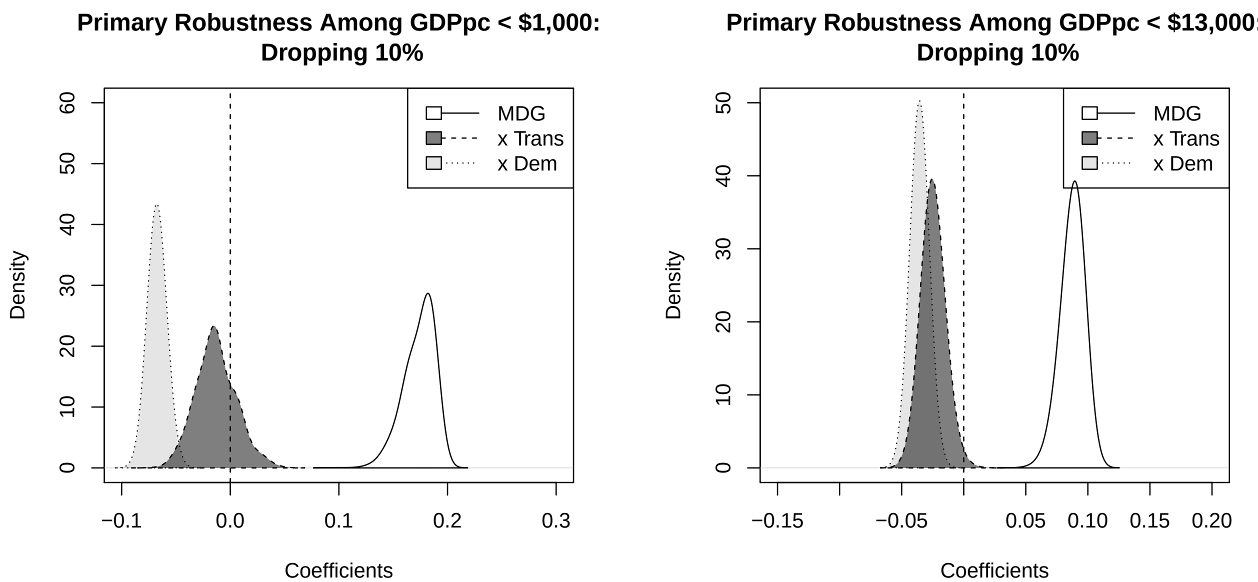

Figure 8: Robustness to outliers. Plots indicate the distribution of coefficient estimates over 1,000 simulations in which 10% of the dataset is dropped at random. The left panel uses only countries with a per capita GDP of less than $1,000 in the year 2000. The right panel uses only countries with a per capita GDP of less than $13,000 in the year 2000.