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

James H. Bisbee, James R. Hollyer, B. Peter Rosendorff, and James Raymond Vreeland

Abstract Precise international metrics and assessments may induce governments to alter policies in pursuit of more favorable assessments according to these metrics. In this paper, we explore a secondary effect of global performance indicators (GPIs). Insofar as governments have finite resources and make trade-offs in public goods investments, a GPI that precisely targets the provision of a particular public good may cause governments to substitute away from the provision of other, related, public goods. We argue that both the main effect of the GPI (on the measured public good) and this substitution effect vary systematically based on the domestic political institutions and informational environments of targeted states. Specifically, we contend that both the main and substitution effects of GPIs should be largest for governments that are least accountable (opaque and nondemocratic) and should be smallest for those that are most accountable. We illustrate the logic of these arguments using a formal model and test these claims using data on primary and secondary enrollment rates across 114 countries. We find that countries substitute toward primary education enrollment rates (which is targeted by the Millennium Development Goals) and away from secondary (which is not), and that these effects are mitigated as accountability rises.

Date received: March 22, 2018; Date accepted: December 9, 2018

> Not everything that can be counted counts, and not everything that counts can be counted.
>
> —William Bruce Cameron *Informal Sociology: A Casual Introduction to Sociological Thinking* (1963, 13)

Quantifiable targets shape incentives. The idea that global performance indicators (GPIs) represent attempts to influence governments is a central tenet of this symposium. When economic development organizations and the foreign aid ministries of states deploy performance-assessment metrics, they stimulate the governments of developing countries to meet and surpass those quantified goals.[1]

Other studies in this symposium corroborate this hypothesis. Doshi, Kelley, and Simmons document the *reactivity* of states to the World Bank's *Ease of Doing Business Indicators* as they take policy steps to attempt to climb in the rankings.[2]

1. Kelley and Simmons 2019.
2. Kelley, Simmons, and Doshi 2019.

International Organization 73, Summer 2019, pp. 547–78

doi:10.1017/S0020818319000109

Measurable targets can also influence aid donors—Honig and Weaver demonstrate that even a small organization can influence aid agencies to publish the outcomes of development projects simply by promising to report on the publication.[3] Failure to achieve measurement targets can also influence market actors. Julia Morse finds that being labeled as noncompliant by the Financial Action Task Force induces private sector investors to avoid a given country.[4]

The upshot of these studies is that GPIs can serve as effective incentives for governments by either directly influencing elite behavior or by indirectly affecting third parties. Precisely because of their effectiveness, however, we explore the unintended consequences of quantified targets. As suggested by the epigraph, when an organization sets a quantified target, it explicitly informs developing countries that some outcomes count—but it also implicitly conveys that other outcomes do not count. As a result, governments have an incentive to substitute resources and effort toward outcomes that are counted and away from those that are not. This unintended effect of assessment is implicitly noted by others in this symposium. In our study of the Millennium Development Goals, we formalize the mechanism by which assessments can lead to unintended outcomes and find empirical evidence consistent with our model.

The Millennium Development Goals (MDGs) were the product of the Millennium Summit of the UN General Assembly in September of 2000. At this meeting, members committed to "spare no effort to free our fellow men, women, and children from the abject and dehumanizing conditions of extreme poverty, to which more than a billion of them are currently subjected."[5] To this end, the declaration advanced resolutions to achieve key numeric targets that evolved into the Millennium Development Goals.[6] Specific targets applied to eight areas: the eradication of extreme poverty and hunger, the achievement of universal primary education, the promotion of gender equality, the reduction of child mortality, the improvement of maternal health, combating HIV and malaria, the promotion of environmental sustainability, and the creation of a global partnership for development.[7]

We focus on the key educational measures of primary and secondary enrollment rates and how they respond to GPIs in equilibrium. Our GPI of interest is the goal targeting universal primary education. Note that secondary education is not targeted by the MDGs, but may be influenced by them in unintended or unanticipated ways.

A rich literature on the relative importance of these public goods to long-run development outcomes motivates our focus on the relative provision of primary and secondary education. Writ large, education's impact has been charted directly (skill

3. Honig and Weaver 2019.
4. Morse 2019.
5. United Nations Millennium Declaration, A/RES/55/2, §11, 18 September 2000. Retrieved 18 April 2016 from <http://www.un.org/millennium/declaration/ares552e.pdf>.
6. Hulme 2009.
7. See Millennium Development Goals and Beyond 2015, retrieved 18 April 2016 from <http://www.un.org/millenniumgoals/bkgd.shtml>.

development, social network access, etc.) and indirectly (character traits such as motivation and self-discipline) which are crucial for social and economic outcomes.[8]

Evidence of unintended consequences of assessment in education reform at the micro level is ample. In the United States, policies such as No Child Left Behind (NCLB) and North Carolina's ABCs programs have been shown to diminish education outcomes for groups not included in assessment measures. Fuller and Ladd find that elementary schools transferred higher-quality teachers from lower (nonevaluated) grades to higher (evaluated) grades in response to NCLB to improve their accountability metrics.[9] Similarly, Hugh Macartney identifies a ratchet effect when accountability reforms are conditioned on prior student test scores, as teachers reduce contemporary effort to ensure room for improvement in the following year.[10]

Studies have also found more general evidence of unintended consequences of quantifiable goals. In sociology, the connection between evaluation and perverse outcomes is so widespread as to be labeled a "law."[11] Le and Malesky demonstrate that provincial governments respond to performance ratings. They also find that these governments "teach to the test."[12] Doshi, Kelley, and Simmons recount interviews with World Bank staff who describe the strategic responses to the Ease of Doing Business indicator, specifically optimizing their response to "score well" as evidenced by the negative relationship between regulatory "distance" and the number of reforms pursued.[13] In this case, governments devote effort to reforms that lead to the greatest changes in rankings, for minimal effort, rather than those that necessarily have the largest economic impact. The literature on government accountability has documented related behavior in which democratic governments focus their efforts on the provision of readily observable public goods in an attempt to claim credit,[14] a tendency that has been particularly documented in the provision of education.[15]

In this paper, we formalize our intuition with a simple model of government accountability in the absence and presence of a GPI. This framework generates empirical propositions that we take to the data, exploring whether the tendency to focus on achieving measured targets induces governments to substitute effort, attention, and resources away from the provision of unmeasured public goods. Specifically, we examine cross-national patterns in the provision of (assessed) primary and (unassessed) secondary education enrollment rates following the passage of the MDGs.

8. See Heckman, Stixrud, and Urzua 2006.
9. Fuller and Ladd 2013.
10. Macartney 2016.
11. See Campbell 1979. Academia itself is replete with examples of suboptimal equilibria as PhD students and tenure-track professors react to evaluation systems such as publications and citations. Edwards and Roy 2017.
12. Le and Malesky 2016.
13. Kelley, Simmons, and Doshi 2019.
14. Mani and Mukand 2007.
15. See Harding and Stasavage 2014. For a model that documents such distortions with regard to a very different policy area, see Bueno de Mesquita 2007.

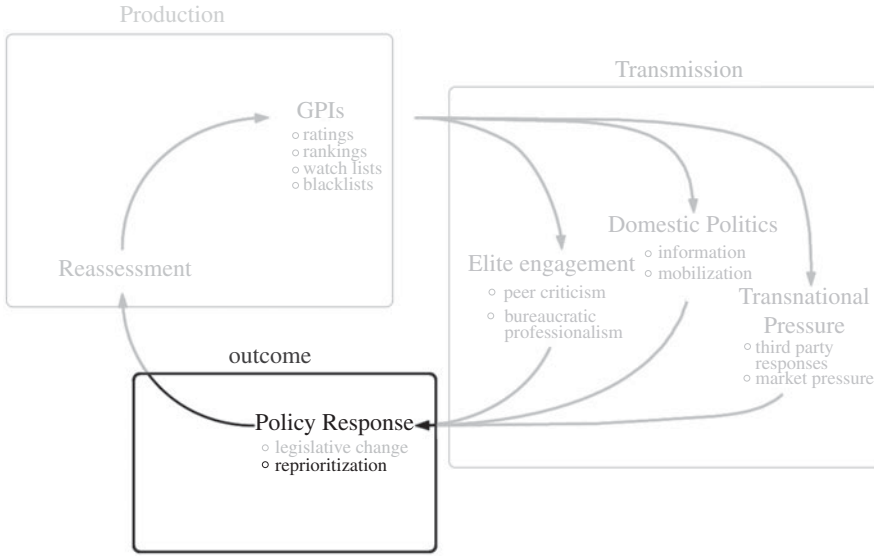
The most direct mechanism connecting GPIs to substitution of resources is via directed aid flows. The Millennium Development Goals and the rich metrics associated with them are used by international actors to inform where to donate aid. Kijima and Lipscy show that participation in cross-national assessments of education influences flows of education aid.[16] These flows of aid can yield substitution effects through the expansion of state budgets or via the incentives they carry for policymakers, although their results suggest that foreign aid inflows are not meaningfully associated with higher net secondary enrollment. Other channels of influence theorized to matter in global assessment elsewhere in this symposium (i.e., Hawthorn effects, activated civil society) can also pertain. We test competing channels of influence via mediation and moderation analyses, finding no evidence of a foreign aid channel and suggestive support for a role played by domestic politics.

We further argue that both the direct effect of international assessment (the promotion of primary education) and the indirect effect on the relationship between primary and secondary education will be moderated by the characteristics of targeted states. To be more precise, both effects should diminish as domestic accountability rises. In states with accountable governments—that is, in democracies with high levels of government transparency—the composition of public goods should reflect the desires of the electorate, at least to a greater extent than in autocracies or opaque states. Accountable states thus have both less need, and less ability, to alter their menu of service provision in response to international pressure.

Our analysis stimulates what we believe are interesting questions about not just the MDGs, but GPIs more generally. On the one hand, we find evidence suggesting that states with healthier civil societies and greater levels of transparency are less susceptible to the substitution incentives associated with monitoring and assessment. However, we also find that these countries are less responsive to the stated goal of the GPI and the explicit incentives are also weaker. Our contribution is located at the end of the theoretical framework encompassing this symposium, reproduced in Figure 1. Our formal model is agnostic about the specific mechanisms by which GPIs induce policy change. We use mediation analysis on our data to test whether foreign aid plays a meaningful role in the context of MDG adoption.

Our contribution also poses questions about the rarity of these mechanisms and the associated normative implications. While we believe unintended consequences are common, if not a foregone conclusion when dealing with strategic actors, it's unclear whether they always subvert the implicit goals of the GPI. All governments are resource constrained, and all allocate these resources in a manner to maximize political support at minimal cost.[17] Any increase in the political returns for the provision of a particular good will imply that the relative incentive to provide another will decline. The normative effects of such substitution, however, need not be negative. Governments' political incentives may induce them to focus on the provision of

16. Kijima and Lipscy 2016.
17. Bueno de Mesquita et al. 2003.



Note: Our paper focuses on the unintended consequences of reprioritization can have on state behavior (highlighted in black).

FIGURE 1. *The pathways of GPIs to policy change*

public goods valued most highly by a narrow elite, in the absence of external pressure. And spillovers in the production of public goods—for example, the expansion of primary education opens access to further schooling to a broader array of students—may imply that a shift in resources is positive for total social welfare. We bracket such critical welfare concerns to focus on the strategic incentives of governments facing GPIs. We acknowledge that such normative conclusions are crucial to the future research on GPIs, and are raised by other works in this symposium and in the ground-breaking research on this topic by Judith Kelley and Beth Simmons.[18]

The Millennium Development Goals

The MDGs consist of eight goals (see Table 1) which were promulgated by the UN, agreed to by such international financial institutions as the World Bank, and adopted by a variety of bilateral aid organizations. Each goal is mapped into one or more "targets," which, in turn, is mapped into several numerical indicators that serve as a basis for assessment of progress. For instance, the goal we examine here—(2) universal primary education—is mapped into the task of "ensur[ing] that, by 2015,

18. Kelley and Simmons 2015.

children everywhere … will be able to complete a full course of primary schooling." Progress toward this target is assessed by primary education enrollment and completion rates, as well as the literacy rate of fifteen-to-twenty-four year-olds, broken down by gender.[19]

TABLE 1. *Millennium development goals*

Goal 1. Eradicate extreme poverty and hunger
Goal 2. Achieve universal primary education
Goal 3. Promote gender equality and empower women
Goal 4. Reduce child mortality
Goal 5. Improve maternal health
Goal 6. Combat HIV/AIDS, malaria, and other diseases
Goal 7. Ensure environmental sustainability
Goal 8. Develop a global partnership for development

Many of these goals were drawn directly from the UN's Millennium Declaration, which the General Assembly issued and approved as the culmination of the Millennium Summit, held in September of 2000.[20] The more precise indicators, and several of the more general goals, were only agreed upon during a World Bank conference, "From Consensus to Action: a Seminar on the International Development Goals," held in 2001.[21] The agenda for both the World Bank conference and the UN Summit, however, was determined by a pre-existing set of development targets—the International Development Goals (IDGs).

The IDGs were originally promulgated by the Organisation for Economic Co-operation and Development (OECD) in 1996 as part of an effort to shape the aid policies of member states. In content, they were very nearly identical to the MDGs. The targets were set by rich-world donors, with fairly minimal input from aid recipients, and were constructed to appeal to domestic audiences in the developed world.[22] In particular, the goals were intended to focus on concrete measures that appealed to a broad audience (e.g., poverty reduction through economic growth, gender equality). Moreover, these goals were thought to advance a degree of *accountability* on the part of recipient states.[23]

The aim of such precise numeric targets was explicitly to increase accountability, particularly for states that received development aid. (Only indicators for goal 8, the "global partnership for development," included measures of donor generosity.) In

19. UN Official List of MDG Indicators, 15 January 2008. Retrieved 18 April 2016 from <http://mdgs.un.org/unsd/mdg/Host.aspx?Content=Indicators/OfficialList.htm>.
20. United Nations Millennium Declaration, 18 September 2000 <http://www.un.org/millennium/declaration/ares552e.pdf>.
21. Hulme 2009.
22. Ibid.
23. Ibid.

laying out these goals, the Secretary General's report noted: "In order to help focus national and international priority-setting, goals and targets should be limited in number, be stable over time and communicate clearly to a broad audience. Clear and stable numerical targets can help to trigger action and promote new alliances for development."[24]

We examine the effects of the MDGs through the lens of existing theories of political accountability. We focus on the goal pertaining to primary education, in part because the mapping between this goal and a numeric indicator was particularly precise, and in part because measures of closely related public goods (secondary and tertiary enrollment rates) were *not* targeted by the MDGs.

Global Assessment Power and Public Goods Production

The MDGs represented a public international commitment by governments to achieve gains in the provision of *specific* public goods. They increased governments' incentives to provide these goods. Governments, therefore, would be likely to devote greater personnel, material resources, and effort toward providing these specific public goods once the MDGs were in effect. However, these incentives also have a secondary effect—governments are likely to divert resources from the production of other public goods to achieve MDG goals.

First, consider the effect of the MDGs on government incentives. Kelley and Simmons argue that the dissemination of international standards may exert social pressure on governments to achieve certain policy outcomes, and that such standards are particularly likely to do so if "they are based on systematic monitoring, are comparative (and especially quantitative), are wielded by a respected actor or group/organization of actors, and are widely disseminated."[25] Perhaps no set of international standards meets these criteria as well as the MDGs. These standards were (1) explicitly linked to precise statistical targets;[26] (2) subject to annual monitoring by the UN and other donor organizations; (3) initially linked to the Millennium Declaration, and thus approved by the UN General Assembly and its constituent member states, and subsequently promulgated jointly by the World Bank, IMF, and OECD;[27] and (4) widely disseminated through public statements by these agencies as well as bilateral lenders.

As Figure 1 represents, international standards like the MDGs may affect government behavior through a variety of channels, including international actors, domestic

24. Annex to "Road Map Towards the Implementation of the United Nations Millennium Development Declaration," Report of the Secretary General. A/56/326, 6 September 2001, 56, section 3. Retrieved 18 April 2016 from <http://www.un.org/millenniumgoals/sgreport2001.pdf>.

25. Kelley and Simmons 2015, 55.

26. See official list of MDG indicators at <http://mdgs.un.org/unsd/mdg/Host.aspx?Content=Indicators/OfficialList.htm>.

27. Hulme 2009.

politics, and the personal or reputational concerns of individual leaders. Global standards may alter behaviors in the international or intergovernmental realm. The creation of a common set of international standards may facilitate decentralized enforcement through bilateral sanctioning or a withdrawal of bilateral benefits.[28] This was a particular risk with regard to the MDGs, given that the goals were explicitly endorsed by most bilateral development agencies soon after their promulgation —with the US as a notable holdout.[29] Alternatively, standards can be enforced through material sanctions (or withdrawal of benefits) by international bodies. Lebovic and Voeten, for instance, demonstrate that the World Bank enforced international standards of human rights, and was less willing to lend to countries that had been "shamed" by the UNCHR.[30] This danger, again, would seem particularly acute for governments that failed to make progress on the MDGs, given the central role both the World Bank and IMF played in the construction of the goals.

Global indicators may also serve to inform residents of monitored states, in a readily digestible manner, about the quality of their government's performance with respect to certain policy goals. Individuals within these states might use such information to discipline government behavior either through the ballot box[31] or through protest or collective action.[32] A well-developed literature finds that such information enables citizens to hold their leaders accountable for public goods provision,[33] which, in turn, induces leaders to more actively provide public goods.[34] Hence, government leaders should face greater domestic pressure to achieve specific public goods targets once the MDGs are promulgated.

Finally, international standards can alter government incentives by generating global norms of acceptable behavior.[35] In defining basic standards of "development," the MDGs may have altered leaders' notions of what constitutes success and failure. Such definitions may shape leaders' primitive preferences over public goods allocations—or shift their strategies for regime legitimation.[36] Focusing on the MDG targeting maternal health and its impact in Zambia, Alice Evans documents extensive evidence of norm adjustment among elites as the primary mechanism by which this particular GPI influenced policy.[37]

It is our contention that the MDGs had the effect of increasing the incentive to provide the specific public goods delineated in the goals. Any or all of these

28. Johns 2012; Milgrom and North 1990.
29. Hulme 2009.
30. Lebovic and Voeten 2009.
31. Besley and Burgess 2002.
32. Hollyer, Rosendorff, and Vreeland 2015.
33. See, for instance, Di Tella and Schargrodsky 2003; Ferraz and Finan 2008; Peters and Welch 1980.
34. See, for instance, Adserà, Boix, and Payne 2003; Besley and Burgess 2002; Glaeser and Goldin 2006; Hollyer, Rosendorff, and Vreeland 2014; Reinikka and Svensson 2003. However, Chong et al. 2010 introduce a cautionary note about these effects.
35. Risse and Sikkink 1999.
36. Kapuściński 1989 is instructive on Emperor Haile Selassie's obsession with a nebulously defined notion of "economic development."
37. Evans 2018.

mechanisms may be effective, and other, unlisted, mechanisms may also be at work. Although our main focus is on the equilibrium provision of targeted and nontargeted public goods, we are able to test some channels of influence, including foreign aid and domestic politics. Our findings relate to similar work by Helena Skagerlind who conducts a detailed process tracing of the effects of the MDG with regard to the empowerment of women.[38] She finds some evidence that all three of these causal channels were in operation with the strongest results for aid conditionality and the socialization of government elites. Our findings for education indicate alternative channels of influence, highlighting the heterogeneity of these mechanisms across targets.

While we contend that the MDGs had the intended effect of increasing government incentives to provide specific public goods, we further contend that they had the unintended effect of causing governments to substitute away from efforts *not* delineated by the UN. As Holmstsrom and Milgrom argue, when a given agent is faced with a multiplicity of tasks, some of which are monitored and others of which are not, the agent will rationally devote a disproportionate level of effort to those goals that are subject to monitoring.[39] We therefore focus on government's tendency to provide a public good listed in the MDGs—for example, primary education—relative to its tendency to provide a public good that is not listed—for example, secondary education. MDGs alter the relative provision of primary and secondary education, boosting primary relative to secondary. This tendency may be somewhat offset to the extent that various public goods are complements in production. For instance, increasing primary enrollment and completion rates perforce increases the pool of students eligible for secondary education. However, the substitution effect should dominate provided these complementarities in production are not too large.

Ironically, the very characteristics that make the MDGs valuable in terms of increasing accountability also render them particularly prone to this substitution effect. Because they focus on precisely defined quantitative metrics of performance, with regard to very specific public goods, they are particularly vulnerable to "gaming." More nebulous or qualitative goals, which are less susceptible to monitoring and less useful for the types of comparisons that give global indicators their power,[40] are also likely to be less prone to substitution effects. However, such weak assessments are also less likely to alter government behavior at all.

The Moderating Role of Information and Institutions

We expect that the intuition we describe may vary by context. Government institutions and the domestic informational environment moderate the effects of global

38. Skagerlind 2016.
39. Holmstrom and Milgrom 1991.
40. Kelley and Simmons 2019.

assessment—both the direct effects on the provision of specific services and the indirect substitution effects on the provision of unassessed public goods. To be more precise, both effects are weakened in instances where governments are highly accountable to domestic audiences. Highly transparent states and democracies are less prone to alter the menu of public goods delivered by the government in response to global indicators than opaque autocracies.

The rationale for this contention is twofold. First, and most directly, domestic and international accountability are substitutes. Recall that one of the mechanisms through which global indicators might induce policy change works through such indicators' effect on domestic politics. Global monitoring might serve to inform domestic interests of their government's performance with regard to a particular policy area, inducing domestic demands for improved performance.[41] Naturally, however, if the domestic information environment is relatively rich, such information is superfluous. Citizens are *already* aware of their government's performance and any pressure to change policy as a result is likely to have already made itself manifest. Moreover, if the issue area in question is one of high public salience, or has important implications for domestic welfare, governments are likely to be induced to disclose information in response to electoral pressures in democratic political systems.[42] Precisely because accountable governments face a strong domestic pressure to provide public goods, increased international demands are likely to have minor effects.

The second reason we expect global indicators to have a relatively minor effect on governments with a high level of domestic accountability is more subtle. Any shift in the composition of public goods in such regimes is likely to be politically costly. Governments in such states are *already* providing levels of public goods provision, in keeping with the demands of large swaths of the population, to stay in power. Any change in the composition of these public goods is likely to be politically costly, more so than for comparable autocracies who rely less on public goods provision to retain power. Moreover, because accountable states provide higher levels of public goods than autocratic states,[43] they will need to alter service provision less to satisfy international demands even in the instances in which they choose to so comply with the MDGs.

Model

To illustrate our claims formally, we adapt a variant of the selectorate model of Bueno de Mesquita and colleagues.[44] We compare two game forms: one with and one without the MDGs. The model demonstrates that (1) the MDGs (weakly) increase

41. Ibid.
42. Hollyer, Rosendorff, and Vreeland 2011.
43. Bueno de Mesquita et al. 2003.
44. Ibid.

the provision of primary education and (2) that wherever this increase is strictly positive there is a substitution away from secondary education provision. The formalization further demonstrates that the effects are more likely to be present in unaccountable (autocratic, opaque) than accountable (democratic, transparent) states.

Model without MDGs

Consider an infinite horizon discrete time interaction between a leader L and her selectorate. Normalize the size of the selectorate to one. In each period of play, L faces a challenger C.[45] To maintain her hold on power, L must maintain the support of a share $\omega \in \left(0, \frac{1}{2}\right]$ of her selectors—a "winning coalition." We conceive of ω as capturing the degree of accountability of L to the selectorate—the probability that L is accountable to (requires the support of) an individual selector randomly selected from the set of selectors.[46] We map the term ω into empirical measures of democracy, government transparency, and their interaction. All actors share a common discount factor, δ.

In each period, the leader and challenger make simultaneous offers of public goods—here consisting of primary (p) and higher/secondary (h) education; they may also offer private transfers (t) to each member of the (potential or actual) winning coalition. A representative member of the selectorate has preferences across $p,\ h,\ t$, according to the additively separable utility function $U(p,\ h,\ t) = U_p(p) + U_h(h) + U_t(t)$, with the standard assumptions: $U_k' > 0,\ U_k'' < 0,\ U_k(0) = 0,\ U_k'(0) = \infty,\ U_k'(\infty) = 0,\ k \in \{p, h, t\}$.

L and C make a decision about how to allocate a budget of size $B > 0$ between $p,\ h,$ and t. If in power, her utility is defined as the residual sum left over after this provision is made. If not in power, her utility is normalized to 0.

$$U_j(p, h, t) = \begin{cases} B - p - h - \omega t \text{ if in power} \\ 0 \text{ otherwise.} \end{cases} \quad \text{for } j \in \{L, C\}$$

At the start of each period of play, L and C make their offers of $p,\ h,$ and t. Then there is an election—a decision is made by members of the winning coalition whether to

45. Throughout, we adopt feminine pronouns for L and selectors. C is given a masculine pronoun.

46. Membership in L's winning coalition is fixed and common knowledge. Formally, denote $\sigma_{i,\tau} \in \{0, 1\}$ as an indicator function of such membership, where i indexes selectors and τ indexes time. L has lexicographic preferences for distributing private transfers to members of her winning coalition over any alternative selectors. In the event L is displaced, new values of $\sigma_{i,\tau+1}$ are drawn for all i, in which all selectors are equally likely to enter the new winning coalition. If L is retained in power, $\sigma_{i,\tau+1} = \sigma_{i,\tau}\ \forall\ i$. Values of $\sigma_{i,\tau}$ are common knowledge in each period τ and all actors condition their strategies on the realization of this value.

retain the leader L. To remain in power, L must retain the support of a share ω of the selectors.

Definition 1: Define an equilibrium to this game as the pair of triples as the pair of triples, (p_C^*, h_C^*, t_C^*) *for the challenger and* (p^*, h^*, t^*) *for the leader. Define* $U_C^* = U_p(p_C^*) + U_h(h_C^*) + U_t(t_C^*)$.

In equilibrium, C will tailor his offer to buy off members of the winning coalition—his entire budget will be devoted to this task. L will address this threat by offering a menu of public goods provision and transfers that guarantees members of the winning coalition a continuation value at least as high as that from backing C. This is feasible for the incumbent because she can credibly commit to provide private transfers t to members of the winning coalition into the future. The offers of the challenger lack future credibility, and are discounted accordingly. The incumbent thus always retains her office.

In the online appendix, we offer a model extension that relaxes the assumption that L's promises are perfectly credible. We assume that members of her winning coalition may be replaced by members of the wider selectorate with a fixed probability in each period of play. However the probability of being removed by the incumbent is less than the probability of losing one's place in the winning coalition if the incumbent is removed and the regime toppled. Our comparative statics are qualitatively unchanged by this model extension.

Lemma 1: In any (Markov Perfect) equilibrium to this (no MDG) game, the fraction $\frac{p^*}{t^*}$ *is rising in* ω.[47]

Lemma 2: Comparative statics: p^* *rises and* t^* *falls with* ω.

As is standard in these games,[48] greater domestic accountability leads to greater public goods provision. This is true both relative to the amount of private transfers provided and in absolute amounts. Having established these preliminaries and laid out the basic game form, we can now turn to the central question: How does the existence of the MDGs affect this domestic equilibrium?

Model with MDGs

To capture the role of the MDGs, we consider a model isomorphic to the one above. The only difference lies in L's utility function. We now assume that the international community offers the leader a benefit for exceeding a certain threshold of primary

47. All proofs are presented in the online appendix.
48. Bueno de Mesquita et al. 2003.

education, as the MDGs specify. (We remain agnostic about which of the theorized transmission mechanisms diagrammed in Figure 1 are at play. In our empirical results, we test the role of foreign aid via mediation analysis.) Let $\beta > 0$ denote this benefit and $\bar{p}$ denote the exogenously imposed threshold from the international community. We can then redefine the utility of the leader:

$$U_L(p,h,t) = \begin{cases} B - p - h - \omega t + \beta \mathcal{I}(p) \text{ if in power} \\ 0 \text{ otherwise.} \end{cases}$$

where

$$\mathcal{I}(p) = \begin{cases} 1 \text{ if } p \geq \bar{p} \\ 0 \text{ otherwise.} \end{cases}$$

As before, the incumbent leader will always successfully buy off her winning coalition and remain in power. Our focus is on her equilibrium offer of primary and secondary education levels, which we will denote p^{**} and h^{**}, respectively. Denote the equilibrium transfer to a given member of the winning coalition as t^{**}.

Before we can proceed further, we require some definitions. Recall that U_C^*, p^*, h^* and t^* are defined in Definition 1 and $\bar{p}$ is the exogenous MDG threshold. The choice of the triple (p, h, t) is endogenous.

Definition 2:

1. *Define the* retention constraint *as* $U_h(h) + \left[1 + \frac{\delta}{1-\delta}(1-\omega)\right] U_t(t) \geq U_C^* - U_p(p)$.
2. *Define the* residual constraint *as* $h + \omega t \leq B + \beta - p$.
3. *Define the* participation constraint as $h + \omega t \leq h^* + p^* - \bar{p} + \omega t^* + \beta$.
4. *Define the set* $\Theta(\omega, \bar{p}, \beta, B)$ *as the set of triples* (p, h, t) *that satisfy these three constraints. That is* $\Theta(\omega,\bar{p},\beta,B) = \{(p,h,t) \in \mathbb{R}^3_+ | U_h(h) + \left[1 + \frac{\delta}{1-\delta}(1-\omega)\right] U_t(t) \geq U_C^* - U_p(p)$ *and* $h + \omega t \leq B + \beta - p$ *and* $h + \omega t \leq h^* + p^* - \bar{p} + \omega t^* + \beta\}$. *Note that this set is parameterized by the exogenous variables, including* $\bar{p}$.
5. *Let* $\hat{p}$ *be the largest value of p among all the triples in* $\Theta(\omega, \bar{p}, \beta, B)$. *That is* $\hat{p} = \sup\{p | (p, h, t) \in \Theta(\omega, \bar{p}, \beta, B)\}$.

The retention constraint is, as its name suggests, a restriction on the values of the menu of goods offered to the selectorate which must be satisfied by the incumbent if she is to retain office. In particular the benefits to the selectorate must be larger than those on offer from the challenger. The residual constraint requires the leader to offer a menu that spends no more than the available budget. The participation constraint recognizes that the leader's decision to adopt the MDGs is voluntary, in the sense that the leader pockets a (weakly) larger residual after adopting the MDG $\bar{p}$ and receiving the bonus β than was the case without the MDG (and choosing p^*, h^*,

t^*). The set of all possible triples (p, h, t) that satisfy all three of these constraints is denoted $\Theta(\cdot)$, and is parameterized by the exogenous variables of the model.

If the benefits the international community offers for attaining the MDG goals are not too large, there is a maximal target the international community can set and expect compliance. If this target for primary education provision $\bar{p}$ is set too high, L will recognize that she cannot possibly satisfy international demands, her residual and participation constraints, and the need to stay in power. We define this value of the MDG goal as $\hat{p}$.

This of course holds for "modest" values of the international reward β. If the international reward is large enough, then the leader can buy off the winning coalition, and satisfy the international community and have plenty left over. To ensure that the equilibrium to MDG game is not degenerate—in the sense that the parameter space where the MDGs can potentially affect behavior and is not empty, we adopt the following restriction:

Lemma 3: A sufficient condition for $\Theta(\omega, \bar{p}, \beta, B)$ *to be non-empty is* $\beta \geq \bar{p} - p^*$.

Given that the feasible set $\Theta(\cdot)$ is not empty (and convex), we can find the element in this set (the optimal choice of t and h) that maximizes the leader's utility in the case that the leader chooses to adopt the MDG, $\bar{p}$:

Definition 3: For any exogenous $\bar{p}$, *define the triple* $(\bar{p}, \bar{h}, \bar{t})$ $\arg \max_{(\bar{p},h,t)\in\Theta(\omega,\bar{p},\beta,B)} B + \beta - \bar{p} - \bar{h} - \omega t$

We are now ready to characterize the equilibrium to the MDG game. Let the equilibrium to the MDG game be denoted $\{(p_C^{**}, h_C^{**}, t_C^{**}); (p^{**}, h^{**}, t^{**})\}$. Recall the equilibrium to the No MDG game is denoted $\{(p_C^*, h_C^*, t_C^*); (p^*, h^*, t^*)\}$.

Proposition 1: Given an exogenous MDG goal of $\bar{p}$, *the equilibrium to the MDG game is: For C,* $(p_C^{**}, h_C^{**}, t_C^{**}) = (p_C^*, h_C^*, t_C^*)$. *For L,*

$$(p^{**}, h^{**}, t^{**}) = \begin{cases} (p^*, h^*, t^*) & \text{if} \quad \bar{p} \leq p^* \\ (\bar{p}, \bar{h}, \bar{t}) & \text{if} \quad \bar{p} \in (p^*, \hat{p}) \\ (p^*, h^*, t^*) & \text{if} \quad \bar{p} \geq \hat{p} \end{cases}$$

The promulgation of the MDGs thus can affect primary education provision, but will do so for only certain configurations of parameter values. Straightforwardly, if the government were already meeting the international community's demand that primary education exceed a certain threshold, the MDGs will have no effect on state behavior. Compliance was given ex ante. By contrast, if international standards are too high, the government will find it impossible to comply while simultaneously meeting her domestic constraints. Since failing to satisfy these domestic considerations will entail a loss of office, the government will disregard the international target.

But for intermediate values of $\bar{p}$—those that exceed the prior provision of primary education but are not so onerous that they inhibit the government's ability to satisfy domestic audiences—the MDGs induce a change in behavior. The government will meet the international community's demand and set $p^{**} = \bar{p}$.

Comparison Across the Two Cases

We are now in a position to compare the equilibrium provision of primary and secondary education under the two different equilibria, which we interpret as the effect of the MDGs. Specifically, we will examine how this effect is moderated by the level of domestic accountability—that is, how this effect varies with ω.

While the MDG level $\bar{p}$ is exogenous with respect to the game form we described, we should bear in mind that the level is in fact set by the international community. It would make little sense for the community to set an MDG level that is in fact out of reach of the very countries and leaders that it is designed to incentivize. That is, for the low ω countries, it makes little practical sense to consider any $\bar{p} > \hat{p}$. In what follows we make the following assumption:

Assumption 1: $\bar{p} \leq \hat{p} \; \forall \; \omega$.

Intuitively, this means that the MDGs are attainable for all countries (given the size of the reward, β) without causing failure in office.[49] For any country, if β is large enough, Assumption 1 is satisfied.

Let us begin by looking at the direct effect of the MDGs—the effect on primary education provision. Formally, this is defined as the difference in the equilibrium primary education provision across the two models p^{**} and p^{*}.

Proposition 2: The MDGs weakly increase the provision of primary education $p^{**} - p^{*} \geq 0$. *The size of this effect is (weakly) decreasing in domestic accountability* ω.

The provision of primary education is weakly higher when the MDGs are present relative to when they are absent. Some polities increase their provision up to the standard set by the international community ($\bar{p}$). Others leave their level of primary education unchanged.

49. Substantively, we believe this restriction on the parameter space to be reasonable. While the goals ($\bar{p}$) specified by the MDGs (universal primary education) apply to all states, the material benefits β used to achieve these goals vary across countries. Given the backing of major multilateral and bilateral development agencies, resources could be expected to be allocated such that the MDGs were achievable for all states. Theoretically, we put this restriction in place because the sign of $\frac{\partial \hat{p}}{\partial \omega}$ is ambiguous, limiting our ability to make predictions outside the portion of the parameter space defined by Assumption 1. More definitively pinning down predictions for this term would require stricter functional form restrictions on $U_k(\cdot)$, $k \in \{p, h, t\}$.

Systematically, the effect of the introduction of the MDGs is smaller in more accountable (higher ω) states. Primary education provision is higher in states with greater domestic accountability ex ante. Some such states will meet international standards with or without the MDGs, and so the introduction of global performance indicators will have no effect. Other accountable governments may fall a little short of international standards, but they will do so by less than their less accountable contemporaries. Meeting the goals set by the international community requires less of such states.

We can now turn our attention to the substitution effect. We begin by documenting that, whenever the effect of the MDGs on primary education is strictly positive (i.e., where the effect is non-zero), their effect on the provision of secondary education is negative. When governments respond to the MDGs they do so by increasing primary, and reducing secondary, education provision. Everywhere else, the MDGs have no effect on public goods provision.

Proposition 3: In all instances where the MDGs increase primary education ($p^{**} > p^{*}$), equilibrium levels of secondary education provision fall ($h^{**} < h^{*}$). We call this the *substitution effect.*

We can go further and characterize the circumstances under which the substitution effect is likely to emerge. Systematically, less accountable states are more likely to experience substitution than more accountable states.[50] Recall from Proposition 2 that polities with higher levels of domestic accountability are less responsive to the MDGs with regard to primary education provision. When there is no direct effect of the MDGs on primary education, there can be no substitution effect on secondary education. By contrast, less accountable states have greater leeway to respond to international incentives. Because their provision of primary education is more responsive to the MDGs, they are more inclined to engage in substitution.

Lemma 4: For any value of $\bar{p}$ that satisfies assumption 1, there exists a corresponding threshold $\bar{\omega}$ such that the substitution effect is present for $\omega < \bar{\omega}$ and absent for values of $\omega \geq \bar{\omega}$.

In what follows, we empirically test propositions 2 and 3 and lemma 4. We anticipate that MDG adoption increased the provision of primary education, and this effect

50. This is not equivalent to stating that everywhere the magnitude of the substitution effect $h^{*} - h^{**}$ is falling in ω. There are two competing effects here. First, less accountable states must boost primary education provision by more to meet international standards. This would tend to increase the tendency toward substitution. Second, and contrastingly, secondary education provision is lower ex ante in less accountable states than in more. The concavity of $U_h(\cdot)$ then implies that cutting secondary education provision reduces the utility of members of the winning coalition by more as ω falls, making substitution more costly. We loosely say that an increase in ω diminishes the size of the substitution effect so long as $U_h(\cdot)$ is not "too concave." More formally, if we let $U_h(h) = h^{\alpha}$, $\alpha \in (0, 1)$, then in the limit as $\alpha \to 1$ the substitution effect is falling in ω for all parameter values.

was most evident in the least domestically accountable states (autocracies with low levels of transparency). This effect should be smallest in more accountable polities (democracies with high government transparency). We further test the existence of a substitution effect—a decline in the relative provision of secondary education following the promulgation of the MDGs. This substitution effect should likewise be most evident in the least accountable states.

Empirics

We test our formal intuition on a cross-national panel of country-level primary and secondary enrollment rates measured between 1980 and 2010. Our empirical tests of the relationship between MDG adoption and enrollment use several methods to reduce bias and spurious associations, including country and year random effects, county-specific time trends, and first-differenced and lagged measures of our dependent variables. Nevertheless, our empirical approach relies on documenting evidence consistent with our theoretical predictions and not making causal claims about MDG adoption.

Data

We draw our definitions of primary and secondary enrollment rates from the UNESCO Institute of Statistics.[51] Primary enrollment rates are defined as the total number of children of official primary-school age who are enrolled in primary education divided by the total population of official primary-school age. Secondary and tertiary enrollment rates are defined analogously and are combined in our data (and our analysis). The UNESCO data contain a nontrivial number of missing observations. We drop country panels with fewer than ten years of data and impute missing values for the remaining country panels using the Amelia II package for R. We implement multiple imputation techniques to account for uncertainty across imputed data sets.[52]

After dropping countries with fewer than ten years of data, we are left with 114 countries with data from 1980 to 2010. We merge this data set with the HRV index of transparency from Hollyer, Rosendorff, and Vreeland[53] which includes additional information—importantly the annual change in real GDP—drawn originally from the Penn World Table version 6.3[54]—and a measure of democracy—drawn

51. See "Data for the Sustainable Development Goals," retrieved 1 March 2016 from <http://www.uis.unesco.org/Pages/default.aspx>.
52. Honaker, King, and Blackwell 2011.
53. Hollyer, Rosendorff, and Vreeland 2014.
54. Heston, Summers, and Aten 2009.

originally from Cheibub, Gandhi, and Vreeland.[55] The democracy measure extends only through 2008 so we impute the values for 2009 and 2010 from this value for each country. We also include a measure of per capita GDP (current USD) from the World Bank[56] as well as three different measures of aid: total aid, aid targeting primary education, and aid targeting secondary education. All measures of aid are transformed to reflect logged per capita estimates. Table 2 presents summary statistics of our key variables, separated by the full sample, countries with GDP per capita in 2000 under $1,000, and those with GDP per capita in 2000 under $13,000.

TABLE 2. *Summary statistics for pre- and post-millennium declaration by sample*

	Full Sample		*$GDP_{2000} \leq \$1{,}000$*		*$GDP_{2000} \leq \$13{,}000$*	
	Pre-MDG	*Post-MDG*	*Pre-MDG*	*Post-MDG*	*Pre-MDG*	*Post-MDG*
PRIMARY ENROLLMENT	96.09	102.64	80.66	95.54	94.95	102.35
(% OF ELIGIBLE)	(22.80)	(15.48)	(29.40)	(21.71)	(24.88)	(16.93)
SEC. / TER. ENROLLMENT	58.74	72.24	31.70	43.43	50.55	64.72
(% OF ELIGIBLE)	(33.01)	(31.14)	(27.67)	(26.74)	(30.17)	(28.95)
GDP	4,789	9,853	1,233	2,701	2,411	5,199
(2000 USD PC)	(7,278)	(13,029)	(2,963)	(4,471)	(3,525)	(5,400)
TRANSPARENCY INDEX	0.86	2.01	−0.04	0.55	0.54	1.61
	(1.73)	(2.20)	(1.22)	(1.40)	(1.53)	(1.99)
DEMOCRACY	0.46	0.58	0.24	0.42	0.38	0.52
	(0.44)	(0.43)	(0.37)	(0.45)	(0.42)	(0.44)
TOTAL AID	48.91	53.18	51.74	53.39	53.74	62.11
(2011 USD PC)	(109.03)	(131.93)	(70.25)	(63.03)	(81.95)	(121.33)
PRIMARY EDUC. AID	0.48	0.79	0.66	1.27	0.59	0.97
(2011 USD PC)	(3.40)	(2.80)	(4.04)	(3.41)	(3.76)	(3.08)
SECONDARY EDUC. AID	0.24	0.20	0.28	0.19	0.26	0.25
(2011 USD PC)	(1.91)	(1.46)	(1.92)	(0.63)	(1.99)	(1.62)
N	3,300	1,815	1,120	616	2,680	1,474

One issue with both the primary and secondary enrollment rate data is that they are subject to ceiling effects. While neither enrollment figure is strictly bounded above by 100 percent—if, for instance, children enroll in primary education at ages below six years, the primary enrollment figure may exceed 100 percent—in practice, neither figure can substantially exceed 100 percent. To diminish the risk that ceiling effects may produce nonlinearities in our relationships of interest, we first scale all enrollment figures relative to their maximum values (such that all fall on the [0, 1] interval) and then apply a logistic transformation to these data.[57] We apply this transformation to all enrollment figures in our data set.

55. Cheibub, Gandhi, and Vreeland 2010.

56. See "GDP per Capita (current US$)," World Bank Data, retrieved 14 April 2016 from <http://data.worldbank.org/indicator/NY.GDP.PCAP.CD>.

57. Denote $p \in [0, 1]$ as the rescaled enrollment figure. We transform this measure by $ln\left(\frac{p}{1-p}\right)$

To measure government accountability, we rely on two measures. The first of these captures political institutions, namely democracy. We draw our measure of democracy from the *Democracy and Development Revisited* data set,[58] which codes a binary indicator $democracy \in \{0, 1\}$ equal to 1 if both legislative and executive posts are filled by meaningful elections.

Our second measure of accountability measures the informational environment of a given country-year. Our measure of this concept is the HRV index of government transparency.[59] The HRV index is a continuous measure, derived from an item-response model of data missingness. It captures the government's tendency to disclose credible information, pertinent to economics and citizen welfare, to the public.[60]

Estimating the Direct Effect of the MDGs

We begin by estimating the direct effect of the MDGs on primary enrollment rates. That is, we start by parsing the evidence that the MDGs increased the growth rate of primary enrollment across countries. Proposition 2 contends that the effect of the MDGs should be positive, and should be larger in unaccountable states. To conduct this estimation, we estimate a varying intercepts multilevel model of the following form:

$$\begin{aligned}\Delta P_{i,t} = {} & \alpha_i + \beta_1 Transparency_{i,t} + \beta_2 Democracy_{i,t} \\ & + I(t \geq 2000)[\gamma_1 + \gamma_2 Transparency_{i,t} + \gamma_3 Democracy_{i,t}] \\ & + \beta_3 \Delta rGDP_{i,t} + \beta_4 Aid_{i,t-1} + \tau_t + \epsilon_{i,t} \\ & \text{where} \\ & \alpha_i \sim \mathcal{N}(\mu_a, \sigma_a) \\ & \tau_t \sim \mathcal{N}(\mu_\tau, \sigma_\tau)\end{aligned} \tag{1}$$

where Δ denotes the difference operator, i indexes country, and t indexes year. By first differencing the outcome term, we eliminate country-specific factors that drive the level of primary education as potential confounds from the specification. In controlling for country (α_i) and year (τ_t) random effects, we help to adjust for random variations in the growth rate in enrollment figures across countries and time. Our results are robust to the inclusion of linear and quadratic time trends.

Our hypotheses are reflected in the γ_j coefficients. Proposition 2 holds that $\gamma_1 > 0$. Our moderating hypotheses from proposition 2, predicting that any such effect of the

58. Cheibub, Gandhi, and Vreeland 2010.

59. Hollyer, Rosendorff, and Vreeland 2014.

60. We use a narrow measure of a specific facet of transparency: the dissemination of data. We view this measure as most appropriate for this study, given that MDGs are measured by some of the very same variables used to construct the HRV index. That said, we acknowledge that in other theoretical settings, measures that capture alternative facets of transparency may be more appropriate. See, for example, Adserà, Boix, and Payne 2003; Berliner 2014; Berliner and Erlich 2015; Besley and Burgess 2002; Bremmer 2006; Broz 2002; Brunetti and Weder 2003; Copelovitch, Gandrud, and Hallerberg 2015; Djankov et al. 2003; Grief 2006; Habyarimana et al. 2009; Islam 2006; Kosack and Fung 2014.

MDGs should diminish in transparent and democratic polities, imply that γ_2, $\gamma_3 < 0$. Results from the model specified in equation 1 are presented in Table 3.

TABLE 3. *The relationship between the millennium declaration and the primary school enrollment rate.*

	Dependent variable:				
	Δ *Primary Enrollment*				
	(1)	*(2)*	*(3)*	*(4)*	*(5)*
MDG	0.069***	0.073***	0.094***	0.067***	0.084***
	(0.013)	(0.014)	(0.017)	(0.011)	(0.016)
TRANS		−0.025***	−0.024***	−0.014*	−0.018**
		(0.007)	(0.007)	(0.008)	(0.009)
DEM		0.055***	0.077***	0.071***	0.083***
		(0.015)	(0.019)	(0.017)	(0.019)
TOTAL AID		−0.003	−0.003	−0.017*	−0.017*
		(0.009)	(0.009)	(0.009)	(0.009)
AID PRIMARY		0.008*	0.008*	0.009*	0.009*
		(0.005)	(0.005)	(0.005)	(0.005)
AID SECONDARY		0.001	0.001	0.0003	0.0003
		(0.004)	(0.004)	(0.004)	(0.004)
MDG*TRANS				−0.030***	−0.023**
				(0.009)	(0.010)
MDG*DEM			−0.045**		−0.034
			(0.021)		(0.023)
CONSTANT	−0.074***	−0.111***	−0.124***	−0.111***	−0.118***
	(0.018)	(0.020)	(0.021)	(0.018)	(0.019)
Observations	4,950	4,950	4,950	4,950	4,950

Notes: Multilevel model analysis allowing for random effects by country and year. Lagged primary enrollment rate and country-specific time trends not shown. Standard errors presented in parentheses. *$p < .10$; **$p < .05$; ***$p < .01$.

Table 3 reveals that coefficients on the MDG indicator are consistently positive and statistically significant at conventional levels. Descriptively, these results indicate that the change in primary enrollment rates was, year on year, consistently more positive following the Millennium Declaration, conditioning on countries, years, and country-specific cubic trends. However, our motivating theory would posit that, failing to control for the confounding effects of transparency and electoral accountability, our estimating equation is inaccurate.

Column 2 updates the specification with these controls. Specifically, we control for transparency and accountability linearly, along with the change in real GDP per capita (not shown) as well as measures of aid, including total aid received, and aid targeting primary and secondary education. The resulting estimate on the MDG indicator maintains its significance and magnitude. However, we still are not properly specifying a linear model in line with our theoretical motivation. To accurately reflect our intuition, we turn to interaction effects, as displayed in columns 3 through 5.

Here we see suggestive evidence supporting our theoretical framework. Namely, across specifications the MDG indicator is significant and positive. We further note that the conditional effect of MDGs declines with both accountability and transparency. Figure 2 displays the marginal effects associated with the relationship between the Millennium Declaration and primary enrollment rates over different values of transparency and among democracies and nondemocracies. Column 5 combines the specifications and finds statistically significant relationships for only transparency.

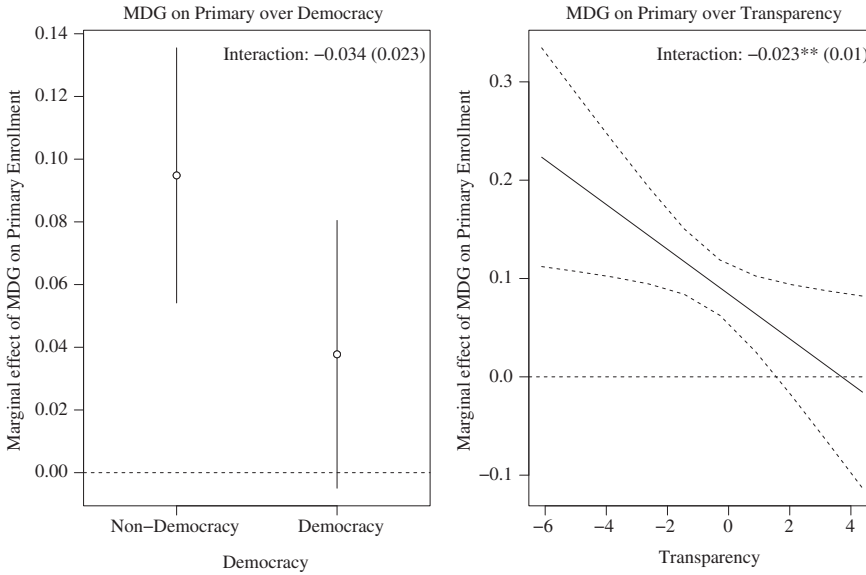


FIGURE 2. *Linear marginal effects plots characterizing heterogeneity in effect of MDG on primary enrollment rates over values of democracy (left) and transparency index (right)*

Results from equation 1 are thus consistent with our proposition 2: (1) The MDGs are associated with an increase in primary enrollment rates, and (2) this effect is diminished in accountable (and particularly in transparent) states.

As an additional test, we replicate the model specified in equation 1, substituting secondary and tertiary enrollment rates for primary. These results offer some preliminary evidence for proposition 3 and lemma 4. Table 4 presents the results, with no evidence of a relationship between enrollment rates at higher levels of education and MDG adoption after controlling for aid flows.

The coefficient on MDGs is noisily estimated but commensurate in magnitude to the coefficient from the primary enrollment rate regression. It is interesting that we still see negative coefficients on the interaction terms with transparency although these are marginally significant and only half the size of the primary enrollment results.

TABLE 4. *The relationship between the millennium declaration and the secondary school enrollment rate.*

	Dependent variable:				
	Δ *Secondary Enrollment*				
	(1)	*(2)*	*(3)*	*(4)*	*(5)*
MDG	0.005**	0.007	0.010	0.011	0.011
	(0.002)	(0.010)	(0.010)	(0.010)	(0.010)
TRANS		0.002	0.003	0.005**	0.005**
		(0.002)	(0.002)	(0.002)	(0.002)
DEM		0.006**	0.008***	0.006**	0.007*
		(0.003)	(0.003)	(0.003)	(0.003)
TOTAL AID		−0.003	−0.003	−0.004	−0.004
		(0.003)	(0.003)	(0.004)	(0.004)
AID PRIMARY		0.00003	−0.00001	−0.00003	−0.00003
		(0.001)	(0.001)	(0.001)	(0.001)
AID SECONDARY		0.0003	0.0003	0.0003	0.0003
		(0.001)	(0.001)	(0.001)	(0.001)
MDG*TRANS				−0.005**	−0.005*
				(0.002)	(0.003)
MDG*DEM			−0.006		−0.0004
			(0.004)		(0.005)
CONSTANT	0.020***	0.039	0.038	0.054	0.054
	(0.003)	(0.043)	(0.042)	(0.048)	(0.048)
Observations	4,950	4,950	4,950	4,950	4,950

Notes: Multilevel model analysis allowing for random effects by country and year. Lagged secondary enrollment rate and country-specific time trends not shown. Standard errors presented in parentheses. $^{*}p < .10$; $^{**}p < .05$; $^{***}p < .01$.

lthough these results are consistent with our theoretical expectations, they do not demonstrate that the *same* countries that increase their primary enrollment rates post-2000 also increase their secondary enrollment rates. To test this relationship, we turn to characterizing the country-level substitution from secondary enrollment to primary in our data.

Estimating the Substitution Effect of the MDGs

Method 1: Estimating the Difference between Primary and Secondary Enrollment Rates

We begin our analysis of substitution by employing a simple measure of the relationship between primary and secondary education within a given country-year: the difference between secondary and primary enrollment rates ($P_{i,t} - S_{i,t}$). If our substitution claim is correct, this measure should increase following the promulgation of the MDGs—that is, the primary enrollment rate should increase relative to the secondary enrollment rate.

We thus estimate a model of the form:

$$(P_{i,t} - S_{i,t}) = \alpha_i + I(year \geq 2000) [\gamma_1 + \gamma_2 Transparency_{i,t} + \gamma_3 Democracy_{i,t}] + \mathbf{X_{i,t}}\boldsymbol{\beta} + \epsilon_{i,t} \quad (2)$$

where i denotes country, t year, α_i is a country random effect, and $\mathbf{X_{i,t}}\ \beta$ is a data vector and associated coefficients. Included as controls are the transparency and democracy measures, GDP per capita, and the earlier measures of aid.

Our hypotheses hold that $\gamma_1 > 0$, and γ_2, $\gamma_3 < 0$. We present results from the model specified in equation 2 in Table 5. Consistent with our theoretical expectations, γ_1 is positive across all specifications and significant at conventional thresholds, indicating that MDG adoption is associated with an increase in the enrollment rates separating primary and secondary students.

TABLE 5. *Coefficients relating the millennium declaration to the difference between primary and secondary enrollment rates across countries.*

	Dependent variable:					
	Primary - Secondary					
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
MDG	1.882**	2.517***	2.120***	2.512***	4.433***	4.785***
	(0.743)	(0.700)	(0.710)	(0.720)	(0.951)	(0.871)
TRANS		−2.995***	−3.092***	−2.273***	−2.993***	−3.109***
		(0.613)	(0.624)	(0.747)	(0.622)	(0.406)
DEM		4.224***	3.700***	3.712***	5.144***	5.312***
		(0.951)	(1.001)	(0.992)	(1.044)	(0.816)
TOTAL AID			2.647***	2.543***	2.519***	2.623***
			(0.554)	(0.552)	(0.551)	(0.535)
AID PRIMARY			1.154***	1.087***	1.138***	1.124***
			(0.231)	(0.231)	(0.230)	(0.218)
AID SECONDARY			0.123	0.117	0.123	0.153
			(0.208)	(0.208)	(0.208)	(0.193)
MDG*TRANS				−1.536***		−0.723
				(0.557)		(0.465)
MDG*DEM					−4.161***	−4.046***
					(1.035)	(1.042)
CONSTANT	49.995***	46.114***	45.823***	45.687***	44.808***	44.127***
	(1.911)	(2.000)	(1.972)	(1.987)	(2.023)	(1.947)
Observations	4,950	4,950	4,950	4,950	4,950	4,950

Notes: Multilevel model analysis allowing for random effects by country and year. Change in real GDP per capita and country-specific time trends not shown. Standard errors presented in parentheses. $^{*}p < .10$; $^{**}p < .05$; $^{***}p < .01$.

Unlike the results presented for primary enrollment in isolation, we note that the transparency index interaction is no longer significant while democracy appears as a highly significant moderator when estimated in the fully interacted specification

(column 6). Aid for primary education is also significant and positive while the estimate for secondary education aid is a precisely estimated 0.

However, the use of the difference between primary and secondary enrollment as an outcome measure induces considerable loss of information. We are unable to tell, for instance, whether the changes in the measure are driven by an absolute decline in secondary rates, by a surge in primary enrollment rates, or some combination of the two factors.

Given these concerns, we turn to two alternative specifications: the first is an error correction model and the second is a vector autoregression discussed in our appendix.

Method 2: Error Correction Specification

First, we turn to an error correction model (ECM). Developed to deal with cointegrated time series, ECMs postulate a long-term equilibrium relationship between covariates—here primary and secondary education.[61] We are interested in changes in this long-term equilibrium. Specifically, we estimate a model of the form:

$$\Delta P_{i,M,t} = \gamma[P_{i,M,t-1} - \beta_{i,M} S_{i,M,t-1}] + \delta_{i,M}\Delta S_{i,M,t-1} + \epsilon_{i,M,t} \quad (3)$$

where, as in equation 1, Δ is the difference operator, i denotes country, t denotes year, P is the primary enrollment rate, and S the secondary. $M \in \{0, 1\}$ is an indicator taking the value 1 after the year 2000.

The expression $P_{i,M,t-1} - \beta_{i,M}\, S_{i,M,t-1}$ denotes the equilibrium relationship between primary and secondary education. When $P_{i,M,t-1} = \beta_{i,M}\, S_{i,M,t\text{-}1}$, the system is in equilibrium, and primary enrollment rates will not tend to adjust. Given $\gamma < 0$ (which is the case in all estimates), when primary education exceeds its equilibrium target ($P_{i,M,t-1} > \beta_{i,M}\, S_{i,M,t-1}$) it will tend to decline. By contrast, when primary education falls below its equilibrium level ($P_{i,M,t-1} < \beta_{i,M}\, S_{i,M,t-1}$), primary enrollment rates will tend to rise.

Since we are interested in changes in the equilibrium relationship between primary and secondary education, we examine changes in the $\beta_{i,M}$ parameter before and after the introduction of the MDGs, and variation in this term across countries, as given by:[62]

61. Beck and Katz 2011.

62. We do so via a two-stage process. First, we estimate an untransformed version of equation 3 given by:

$$\begin{aligned} \Delta P_{i,M,t} &= \gamma P_{i,M,t-1} + \zeta_{i,M} S_{i,M,t-1} + \delta_{i,M}\Delta S_{i,M,t-1} + \epsilon_{i,M,t} \\ \zeta_{i,M} &= \omega_{0,i} + \omega_{1,i} M_i + E_i \\ &\text{where,} \\ \omega_{0,i} &\sim \mathcal{N}(\alpha_{0,i}, U_{0,i}) \\ \omega_{1,i} &\sim \mathcal{N}(\alpha_{1,i}, U_{1,i}) \end{aligned} \quad (4)$$

which gives us estimated values of $\beta_{i,M}$. We then regress these parameters as described in equation 5.

$$\begin{aligned}\beta_{i,M} &= \alpha + \pi_1 Trans_i + \pi_2 Dem_i + \pi_3 Trans_i \times Dem_i \\ &+ I(year \geq 2000)[\omega_1 + \omega_2 Trans_i + \omega_3 Dem_i] \\ &+ ln(GDPpc_{2000})_i + \eta_{i,M}\end{aligned} \tag{5}$$

Proposition 3 holds that $\omega_1 > 0$, implying that the equilibrium level of secondary education, relative to primary, declines once the MDGs go into effect. Lemma 4 holds that ω_j, $j \in \{2, 3\} < 0$—the substitution effect is diminished in accountable polities. Estimates from equation 5 are presented in Table 6.

TABLE 6. β *substitution tests for equilibrium levels of secondary and primary education.*

	Dependent variable:			
	$\beta_i M$			
	(1)	*(2)*	*(3)*	*(4)*
MDG	0.080***	0.130***	0.136***	0.145***
	(0.026)	(0.030)	(0.037)	(0.033)
TRANS			0.013*	0.011
			(0.007)	(0.008)
DEM		0.043**		0.014
		(0.021)		(0.025)
TOTAL AID	0.016	0.014	0.009	0.009
	(0.016)	(0.016)	(0.016)	(0.016)
AID PRIMARY	−0.014	−0.015	−0.009	−0.010
	(0.016)	(0.016)	(0.015)	(0.016)
AID SECONDARY	−0.004	−0.002	−0.002	−0.002
	(0.020)	(0.020)	(0.019)	(0.018)
MDG*TRANS			−0.028***	−0.026***
			(0.007)	(0.009)
MDG*DEM		−0.087***		−0.023
		(0.029)		(0.037)
CONSTANT	0.066	0.040	0.021	0.017
	(0.216)	(0.214)	(0.208)	(0.204)
Observations	228	228	228	228

Notes: Multilevel model analysis allowing for random effects by country and year. Change in real GDP per capita and country-specific time trends not shown. Standard errors presented in parentheses. $^{*}p < .10$; $^{**}p < .05$; $^{***}p < .01$.

As suggested by Table 6, the Millennium Declaration corresponds to a divergence of primary and secondary enrollment rates, suggesting a substitution effect (as indicated by the statistically significant coefficient in column 1).

However, this substitution effect is mitigated by democratic accountability, with the interaction term in column 2 entering negative and marginally significant. A similar effect is documented with transparency, where the substitution effect is reduced in countries with higher levels of transparency. The coefficient on the transparency term is positive, indicating that, before the MDGs, transparent governments had a higher equilibrium level of primary relative to secondary education than

unaccountable states. This is consistent with existing claims, notably by David Stasavage.[63] Combining these specifications suggests that the moderating effect is driven primarily by transparency, not electoral accountability.

ECM models, however, impose strong parametric assumptions. We merely posit that primary and secondary enrollment rates exhibit a long-term structural relationship of the form documented by equation 3, and rely heavily on this equilibrium in our interpretation. In particular, Grant and Lebo argue that such general error correction models will be biased if the respective time series—here primary and secondary enrollment rates—are either not co-integrated, or are co-integrated of different orders.[64] The coefficient γ, in that instance, will be inconsistent—and so too will our parameters of interest β.

To address this concern, we first examine the co-integration of these two time series. While the series appear co-integrated in global tests, running these tests country by country supports co-integration in only roughly half our sample of states. We therefore re-estimate our models, restricting our sample to only those states where the evidence clearly indicates that primary and secondary enrollment are co-integrated. Results are substantively similar to those in the full sample.

Mechanisms

Thus far we have developed a model that predicts GPIs can lead to unintended substitution behavior and have presented empirical evidence consistent with our predictions. First, we have shown that MDG adoption prompted an increase in primary school enrollment as was its intention but that the magnitude of this relationship is strongest in the least accountable polities in our data. Second, this increase in primary enrollment was accompanied by a decline in secondary enrollment, consistent with our theoretical prediction for substitution effects. Again, evidence of this substitution effect is strongest in the least accountable polities.

However, our formal model is agnostic with respect to the specific mechanisms by which GPIs operate (see Figure 1). We are uncertain about whether these results arise from the budgetary impacts of foreign aid and investment which may respond to GPIs ("transnational pressures"), the activation of civil society ("domestic politics"), or from a direct effect on policymakers themselves ("elite shaming"). The moderating effects of accountability documented earlier suggest a possible channel via domestic politics. We therefore test whether a transnational pressure channel via foreign aid also operates.

To conduct this analysis, we treat foreign aid as a mediating variable and use a two-stage regression framework as in Baron and Kenney.[65] The first stage predicts changes in foreign aid flows as a function of MDG adoption, given in equation

63. Stasavage 2005.
64. Grant and Lebo 2016.
65. Baron and Kenny 1986.

6. We interact MDG adoption with the primary enrollment rate in the year 2000, reflecting the intuition that performance assessments like the Millennium Development Goals affect aid flows specifically for underperformers.

$$\begin{aligned}\Delta Aid_{i,t} = {} & \alpha_i + \beta_1 Primary_{i,2000} \\ & + I(t \geq 2000)[\rho_1 + \rho_2 Primary_{i,2000}] \\ & + \beta_1 \Delta rGDP_{i,t} + \beta_2 GDP_{i,2000} + \beta_3 Transparency_{i,t} \\ & + \beta_4 Democracy_{i,t} + \tau_t + \epsilon_{i,t} \\ & \text{where} \\ & \alpha_i \sim \mathcal{N}(\mu_a, \sigma_a) \\ & \tau_t \sim \mathcal{N}(\mu_\tau, \sigma_\tau)\end{aligned} \tag{6}$$

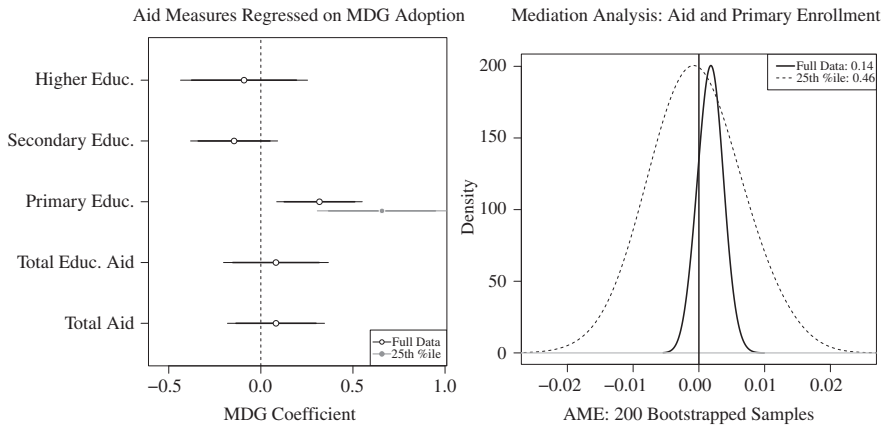
We save the coefficient on MDG adoption (ρ_1) for use in our second-stage analysis. We test a number of different measures of foreign aid, plotted in the left panel of Figure 3. As illustrated, aid targeting primary education increases following MDG adoption, particularly toward countries with low initial primary enrollment rates. This first-stage pattern is consistent with the theorized mechanism of "transnational pressures" discussed in the introduction to this symposium and illustrated in Figure 1—aid flows do respond to the MDGs.

In the second stage, we regress the change in primary enrollment on MDG adoption and the same aid measure, replicating the specification described in equation 1 and saving the β_4 coefficient. Under restrictive assumptions,[66] the product of these coefficients captures the average mediation effect between MDG adoption and primary enrollment rates that travels via changes in aid flows. To accommodate the multiply imputed data, we bootstrap these estimates and plot the distribution of the results in the right panel of Figure 3. Pseudo p-values are calculated as the share of bootstrapped estimates greater than 0.

As illustrated, mediation analysis offers no robust support for the theorized mechanism of foreign aid, despite the first-stage channel being active. While foreign aid targeting primary enrollment does respond to the assessments provided by the MDGs (left panel of Figure 3), there is no evidence that this is the mechanism by which policymakers reallocate resources. The magnitude of the average mediation effect (AME) is only a fraction of the overall coefficient and is only marginally significant for the full data. Our null finding is consistent with the results of Kijima and Lipscy who also find that, while assessment participation is associated with higher foreign aid inflows, these sources of aid are not meaningfully associated with higher net secondary enrollment.[67] As such, we conclude that, in the context of

66. Our empirical endeavor is purely observational and we make no causal claims that would necessitate more sophisticated causal mediation methods such as those pioneered by Imai, Keele, and Yamamoto 2010.

67. Kijima and Lipscy 2016.



Notes: The first-stage results (left panel) show that foreign aid for primary education increases following MDG adoption, particularly for the countries in the bottom quartile of primary enrollment. However, the right panel indicates that the aid channel represents only a small fraction of the overall relationship and is not statistically significant (pseudo p-values presented in the legend) suggesting that this mechanism is inactive.

FIGURE 3. *First stage (left panel) and overall (right panel) results from mediation analysis on the mechanism of foreign aid*

education and the MDGs, transnational pressures are not the primary mechanism by which this particular GPI induces change.

Conclusion

To what extent can international initiatives alter country behavior? This special issue adds to the literature by addressing the topic holistically, starting with how GPIs are produced, charting the mechanisms, and assessing their impact. Our contribution expands our understanding of how GPIs influence state behavior by exploring how these initiatives affect not just their specific target but the larger equilibrium levels of public goods provision. We argue that countries with finite resources must trade off between different public goods and that their ability to do so is constrained by their institutions. We develop a formal model expressing our argument and test the hypotheses it generates with data.

In our empirical context, we document evidence of a substitution effect between secondary and primary education in response to the Millennium Development Goals targeting universal primary enrollment. We further find that this substitution effect is stronger in less accountable countries. Our findings support the simple premise that resource-constrained governments (1) do respond to international initiatives but (2) may do so by reallocating away from other, potentially desirable, outcomes.

Our moderation results highlight an important role played by domestic society. We find that both the direct and substitution effects of MDGs are strongest in less accountable polities. Contrastingly, our mediation analysis suggests a limited role for international actors directly influencing government decision making through aid allocation. The robustness of our results using the HRV index of transparency, particularly when compared to the weak findings for democracy, suggest that GPIs like the Millennium Development Goals are most effective as information conduits.

Neither of these findings precludes other mechanisms, such as the reputational concerns documented in Zambia by Alice Evans or other types of international actors beyond donors.[68] Further research is required to identify the mechanisms connecting the MDGs to increasing primary enrollment rates and substitution between secondary and primary outcomes.

Our findings offer a mixed take-away for policymakers. On the one hand, we find evidence that global assessment can alter government behavior, and particularly increase the provision of public goods (i.e., primary education). Moreover, these effects are likely to be strongest in the countries that need it most—opaque autocracies that underprovide public goods in the absence of international pressure. However, we also find that global accountability can have unintended second-order effects. Resource-constrained leaders, intent on their own political survival, will diminish the provision of goods and services for which the international community does not hold them accountable, even as they increase those emphasized through global assessment programs.

We are careful, however, not to draw broader normative conclusions from our analysis. Although the counterfactual conclusion is that secondary enrollment would have been even higher in the absence of the MDGs, this ignores the positive spillovers between primary and secondary enrollment.

Our findings connect with the symposium by expanding the range of outcome behavior to include nontargeted public goods. By accounting for strategic trade-offs and equilibrium behavior, we present a more rounded view of the effects of GPIs, both intended and not. Our findings echo educational program evaluation elsewhere, in which targeted schools make trade-offs in order to attain desired metrics.[69] While our results do not identify the specific mechanisms by which these trade-offs occur, they highlight a fruitful area for research on GPIs writ large and MDGs specifically.

Supplementary Material

Supplementary material for this article is available at <https://doi.org/10.1017/S0020818319000109>.

68. Evans 2018.
69. See Fuller and Ladd 2013; Macartney 2016.

References

Adserà, Alícia, Carles Boix, and Mark Payne. 2003. Are You Being Served? Political Accountability and Quality of Government. *Journal of Law, Economics and Organization* 19 (2):445–90.

Baron, Reuben M., and David A. Kenny. 1986. The Moderator–Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations. *Journal of Personality and Social Psychology* 51 (6):1173–82.

Beck, Nathaniel, and Jonathan N. Katz. 2011. Modeling Dynamics in Time-Series-Cross-Section Political Economy Data. *Annual Review of Political Science* 14:331–52.

Berliner, Daniel. 2014. The Political Origins of Transparency. *Journal of Politics* 76 (2):479–91.

Berliner, Daniel, and Aaron Erlich. 2015. Competing for Transparency: Political Competition and Institutional Reform in Mexican States. *American Political Science Review* 109 (1):110–28.

Besley, Timothy, and Robin Burgess. 2002. The Political Economy of Government Responsiveness: Theory and Evidence from India. *Quarterly Journal of Economics* 117 (4):1415–51.

Bremmer, Ian. 2006. *The J Curve: A New Way to Understand Why Nations Rise and Fall*. Simon and Schuster.

Broz, J. Lawrence. 2002. Political System Transparency and Monetary Commitment Regimes. *International Organization* 56 (4):861–87.

Brunetti, Aymo, and Beatrice Weder. 2003. A Free Press Is Bad News for Corruption. *Journal of Public Economics* 87:1801–24.

Bueno de Mesquita, Bruce, Alastair Smith, Randolph M. Siverson, and James D. Morrow. 2003. *The Logic of Political Survival.* MIT Press.

Bueno de Mesquita, Ethan. 2007. Politics and the Suboptimal Provision of Counterterror. *International Organization* 61:9–36.

Cameron, William Bruce. 1963. *Informal Sociology: A Casual Introduction to Sociological Thinking*, vol. 21. Random House.

Campbell, Donald T. 1979. Assessing the Impact of Planned Social Change. *Evaluation and Program Planning* 2 (1):67–90.

Cheibub, José Antonio, Jennifer Gandhi, and James Raymond Vreeland. 2010. Democracy and Dictatorship Revisited. *Public Choice* 143 (1–2):67–101.

Chong, Alberto, Ana L. De La O, Dean Karlan, and Leonard Wantchekon. 2010. Information Dissemination and Local Governments' Electoral Returns: Evidence from a Field Experiment in Mexico. Paper presented at the Annual Meeting of the American Political Science Association, Seattle, March.

Copelovitch, Mark, Christopher Gandrud, and Mark Hallerberg. 2015. Financial Regulatory Transparency, International Institutions, and Borrowing Costs. Prepared for the 2015 Annual Political Economy of International Organizations Conference, February.

Di Tella, Rafael, and Ernesto Schargrodsky. 2003. The Role of Wages and Auditing During a Crackdown on Corruption in the City of Buenos Aires. *Journal of Law, Economics and Organization* 46 (1):269–92.

Djankov, Simeon, Caralee McLiesh, Tatiana Nenova, and Andrei Shleifer. 2003. Who Owns the Media? *Journal of Law and Economics* 46 (2):341–81.

Edwards, Marc A., and Siddhartha Roy. 2017. Academic Research in the Twenty-First Century: Maintaining Scientific Integrity in a Climate of Perverse Incentives and Hypercompetition. *Environmental Engineering Science* 34 (1):51–61.

Evans, Alice. 2018. Amplifying Accountability by Benchmarking Results at District and National Levels. *Development Policy Review* 36 (2):221–40.

Ferraz, Claudio, and Frederico Finan. 2008. Exposing Corrupt Politicians: The Effects of Brazil's Publically Released Audits on Electoral Outcomes. *Quarterly Journal of Economics* 123 (2):703–45.

Fuller, Sarah C., and Helen F. Ladd. 2013. School-Based Accountability and the Distribution of Teacher Quality Across Grades in Elementary School. *Education* 8 (4):528–59.

Glaeser, Edward L., and Claudia Goldin, eds. 2006. *Corruption and Reform: Lessons from America's Economic History*. University of Chicago Press.
Grant, Taylor, and Matthew J. Lebo. 2016. Error Correction Methods with Political Time Series. *Political Analysis* 24 (1):3–30.
Grief, Avner. 2006. *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade*. Cambridge University Press.
Habyarimana, James, Macartan Humphreys, Daniel N. Posner, and Jeremy M. Weinstein. 2009. *Coethnicity: Diversity and the Dilemmas of Collective Action*. Russell Sage Foundation.
Harding, Robin, and David Stasavage. 2014. What Democracy Does (and Doesn't Do) for Basic Services: School Fees, School Inputs, and African Elections. *Journal of Politics* 76 (1):229–45.
Heckman, James J., Jora Stixrud, and Sergio Urzua. 2006. The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics* 24 (3):411–82.
Heston, Alan, Robert Summers, and Bettina Aten. 2009. Penn World Table Version 6.3. Technical Report, Center of International Comparisons at the University of Pennsylvania.
Hollyer, James R., B. Peter Rosendorff, and James Raymond Vreeland. 2011. Democracy and Transparency. *Journal of Politics* 73 (4):1–15.
———. 2014. Measuring Transparency. *Political Analysis* 22 (4):413–34.
———. 2015. Transparency, Protest and Autocratic Instability. *American Political Science Review* 109 (4): 764–84.
Holmstrom, Bengt, and Paul Milgrom. 1991. Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. *Journal of Law, Economics and Organization* 7:24–52.
Honaker, James, Gary King, and Matthew Blackwell. 2011. Amelia II: A Program for Missing Data. *Journal of Statistical Software* 45 (7):1–47.
Honig, Dan, and Catherine Weaver. 2019. A Race to the Top? The Aid Transparency Index and the Normative Power of Global Performance Assessments. *International Organization* 73 (3). <https://doi.org/10.1017/S0020818319000122>
Hulme, David. 2009. The Millennium Development Goals (MDGs): A Short History of the World's Biggest Promise. Brooks World Poverty Institute Working Paper 100.
Imai, Kosuke, Luke Keele, and Teppei Yamamoto. 2010. Identification, Inference and Sensitivity Analysis for Causal Mediation Effects. *Statistical Science* 25 (1):51–71.
Islam, Roumeen. 2006. Does More Transparency Go Along with Better Governance? *Economics and Politics* 18 (2):121–67.
Johns, Leslie. 2012. Courts as Coordinators: Endogenous Enforcement and Jurisdiction in International Adjudication. *Journal of Conflict Resolution* 56 (2):257–89.
Kapuściński, Ryszard. 1989. *The Emperor: Downfall of an Autocrat*. Vintage.
Kelley, Judith G., and Beth A. Simmons. 2015. Politics by Number: Indicators as Social Pressure in International Relations. *American Journal of Political Science* 59 (1):55–70.
———. 2019. Global Assessment Power in the Twenty-First Century. *International Organization* 73 (3).
Kelley, Judith G., Beth A. Simmons, and Rush Doshi. 2019. The Power of Ranking: The Ease of Doing Business Indicator as a Form of Social Pressure. *International Organization* 73 (3). <https://doi.org/10.1017/S0020818319000158>
Kijima, Rie, and Phillip Y. Lipscy. 2016. The Politics of International Testing. Working paper presented at the Assessment Power in World Politics Workshop, Harvard University, May.
Kosack, Stephen, and Archon Fung. 2014. Does Transparency Improve Governance? *Annual Review of Political Science* 17:65–87.
Le, Ahn, and Edmund Malesky. 2016. Do Subnational Assessments (SPAs) Lead to Improved Governance? Evidence from a Field Experiment in Vietnam. Working paper presented at the Assessment Power in World Politics Workshop, Harvard University, May.
Lebovic, James H., and Erik Voeten. 2009. The Cost of Shame: International Organizations and Foreign Aid in the Punishing of Human Rights Violations. *Journal of Peace Research* 46 (1):79–97.
Macartney, Hugh. 2016. The Dynamic Effects of Educational Accountability. *Journal of Labor Economics* 34 (1):1–28.

Mani, Anandi, and Sharun Mukand. 2007. Democracy, Visibility and Public Good Provision. *Journal of Development Economics* 83 (2):506–29.
Milgrom, Paul R., and Douglass C. North. 1990. The Role of Institutions in the Revival of Trade: The Law Merchant, Private Judges, and the Champaign Fairs. *Economics and Politics* 2 (1):1–23.
Morse, Julia C. 2019. Blacklists, Market Enforcement, and the Global Regime to Combat Terrorist Financing. *International Organization* 73 (3). <https://doi.org/10.1017/S002081831900016X>
Peters, John G., and Susan Welch. 1980. The Effects of Charges of Corruption on Voting Behavior in Congressional Elections. *American Political Science Review* 74 (3):697–708.
Reinikka, Ritva, and Jakob Svensson. 2003. The Power of Information: Evidence from a Newspaper Campaign to Reduce Capture. Working Paper: Institute for International Economic Studies, Stockholm University.
Risse, Thomas, and Kathryn Sikkink. 1999. The Socialization of International Human Rights Norms into Domestic Practices. In *The Power of Human Rights: International Norms and Domestic Change*, 1–38. Cambridge University Press.
Skagerlind, Helena Hede. 2016. Assessment Power in Global Development Policy: The Millennium Development Goals. Working paper presented at the Assessment Power in World Politics Workshop, Harvard University, May.
Stasavage, David. 2005. Democracy and Education Spending in Africa. *American Journal of Political Science* 49 (2):343–58.

Authors

James H. Bisbee is a PhD candidate in the Department of Politics at New York University. He can be reached at james.bisbee@nyu.edu.
James R. Hollyer is a Benjamin Evans Lippincott Associate Professor in Political Economy at the University of Minnesota in Minneapolis. He can be reached at jhollyer@umn.edu.
B. Peter Rosendorff is Professor of Politics in the Department of Politics at New York University. He can be reached at peter.rosendorff@nyu.edu.
James Raymond Vreeland is Professor of Politics and International Affairs in the Woodrow Wilson School and Department of Politics at Princeton University. He can be reached at jrv@princeton.edu.

Acknowledgments

We are grateful to the attendees of the Global Assessment Power Workshop at Harvard University, in particular to our discussant, Karthik Ramanna, our quantitative discussants, Anton Strezhnev and Christopher Lucas, and the organizers, Beth Simmons and Judith Kelley. We are also grateful to John Freeman for feedback on our error correction model, and Mary Stegmaier and participants in the 2016 EPSA panel on "Economic Voting when it Matters Most."

Funding

James Hollyer thanks the Niehaus Center on Globalization and Governance and the Benjamin Lippincott foundation for research support.