SCHOOL FOR
INTERNATIONAL STUDIES

Are Development Statistics Manipulable?

Andrew Kerner, Morten Jerven and Alison Beatty

Simons Papers in Security and Development

No. 37/2014 | June 2014

The **Simons Papers in Security and Development** are edited and published at the School for International Studies, Simon Fraser University. The papers serve to disseminate research work in progress by the School's faculty and associated and visiting scholars. Our aim is to encourage the exchange of ideas and academic debate. Inclusion of a paper in the series should not limit subsequent publication in any other venue. All papers can be downloaded free of charge from our website, www.sfu.ca/internationalstudies.

The series is supported by the Simons Foundation.

Series editor: Jeffrey T. Checkel
Managing editor: Martha Snodgrass

Kerner, Andrew, Morten Jerven and Alison Beatty, *Are Development Statistics Manipulable?* Simons Papers in Security and Development, No. 37/2014, School for International Studies, Simon Fraser University, Vancouver, June 2014.

ISSN 1922-5725

School for International Studies
Simon Fraser University
Suite 7200 - 515 West Hastings Street
Vancouver, BC Canada V6B 5K3

SCHOOL FOR INTERNATIONAL STUDIES

SFU SIMON FRASER UNIVERSITY ENGAGING THE WORLD

Are Development Statistics Manipulable?

Simons Papers in Security and Development
No. 37/2014 | June 2014

Abstract:
Coordinating foreign aid distribution to the poorest countries requires classifying them into developmental cohorts. In principle these designations are objective and immune from manipulation by aid-seeking countries. The objectivity and reliability of these data are important for aid distribution as well as for the use of these data in social scientific applications. We ask whether there are indications that these data are being influenced by aid-seeking manipulation. To do so we examine the distribution of GNIs per capita around the eligibility threshold for World Bank's International Development Association (IDA). We examine the data as whole and separately for countries that are plausibly more motivated to aid-seek by virtue of their aid-dependence or more capable of doing so by virtue of being perceived as trustworthy. We show that the distribution of GNIs per capita from aid-dependent countries displays indications of aid-seeking data manipulation. This finding is robust to a variety of model specifications, but somewhat sensitive to the exclusion of individual countries from the sample. As such, these findings are more suggestive than definitive, but they do lend credence to the idea of data generation as a strategic process and suggest the need for more research in this area.

About the authors:
Andrew Kerner is an assistant professor of political science at the University of Michigan. His research interests include the politics of finance and FDI, corporate governance and the role of institutional investors (particularly pension funds) in emerging market policymaking. His MA and PhD are from Emory University.

Morten Jerven is an economic historian and an associate professor at Simon Fraser University. His publications include Poor Numbers: *How We Are Misled by African Development Statistics and What to Do about It* (Cornell University Press, 2013). His second book, *Economic Growth and Measurement Reconsidered in Botswana, Kenya, Tanzania, and Zambia, 1965–1995*, appeared with Oxford University Press in 2014.

Alison Beatty is a PhD candidate in public policy and political science at the University of Michigan. Her research addresses interstate competition for FDI and the developing role of US states in the global economy. Before her doctoral studies Beatty interned with the US State Department's Office of Central Europe.

About the publisher:
The School for International Studies (SIS) fosters innovative interdisciplinary research and teaching programs concerned with a range of global issues, but with a particular emphasis on international development, and on global governance and security. The School aims to link theory, practice and engagement with other societies and cultures, while offering students a challenging and multi-faceted learning experience. SIS is located within the Faculty of Arts and Social Sciences at Simon Fraser University. Our website is www.sfu.ca/internationalstudies.

Are Development Statistics Manipulable?

Common understandings of which countries are poor and which countries are not are crucial to both the study and conduct of international politics. For the most part these common understanding come from development statistics such as GDP or GNI. The accuracy of these data and the understandings that follow from them are crucially important. To take an example at the heart of this paper, coordinating aid distribution to the poorest countries requires a common understanding of which countries are poorer than others, and this understanding is typically drawn from development statistics. The study of development's political and economic causes and consequences also requires common understandings of which countries are poor and that the data from which these understandings are drawn be reasonably accurate. Social scientific analyses also require that these data be unbiased, which is to say that whatever inevitable errors exist in the data be orthogonal to the political and economic attributes whose relationship to development social scientists seek to uncover.

This paper considers a threat to the reliability and unbiasedness of development statistics in many social science applications.[1] The threat to reliability and unbiasedness is as follows. Poorer countries are typically eligible for and receive more aid than richer countries. However, the statistics used to establish which countries are poorer than others represent economic concepts that cannot be observed directly. They must be estimated and these estimates are based on data that originates in national statistical offices. Aid-seeking countries may therefore prefer to report estimates that give them the appearance of being poor.[2] Doing so need not entail reporting data that is *per se* "wrong." The range of plausible estimates can be quite large, and national statistical agencies from aid-seeking countries may simply choose from the large set of plausible estimates those that are likely to generate more aid.[3] We'll refer to such a dynamic as

1 When we refer to the "biasedness" of data we are referring to the suitability of data in a particular context, rather than any intrinsic quality. We are particularly concerned about the use of data to answer questions such as "Does democracy promote development?" or "Does foreign aid promote development?" To be unbiased in these contexts requires that errors in the data are not correlated either democracy or aid-dependence.

2 Aid-seeking is surely not the only motive a government might have to control their data. Other governments likely want to appear to be richer, larger, or growing faster than they really are. Argentina, for example, was recently censured by the IMF for providing inaccurately rosy inflation and GDP growth data (Economist 2013).

3 For example, recent rebasing of Nigerian GDP data suggests that prior estimates understated Nigerian GDP by as much as 60%. http://www.economist.com/blogs/baobab/2014/01/nigerias-economy-will-soon-overtake-south-

"data manipulation," though we reiterate that in this context data manipulation does not imply inaccurate or unjustifiable estimates, but simply estimates that are purposefully chosen in order to further a country's aid-seeking interests.

The potential for aid-seeking data manipulation is mitigated by the central role that international organizations play in the oversight, production, and dissemination of data. International organizations' role in this process is important and intentional (Barnett and Finnemore 2004, 2; Abbott and Snidal 1998, 8; Keohane 2005 92; Koremenos, Lipson and Snidal 2001, 766). The authority of data released by international organizations reflects users' trust that the data represent that organization's best approximation of reality, uncolored by individual countries' political or economic interests (Abbott and Snidal 1998, 20). In practice, however, the disseminators of international data rely exclusively on the official estimates provided by member countries. The World Bank, for example, typically adjusts nationally reported figures, but if there are problems with the national data, the international datasets may inherit them (Jerven 2013a).[4]

This paper asks whether the macroeconomic datasets commonly used by researchers evince any evidence of aid seeking through data manipulation. The possibility that there exists a "political economy of data" is notable in itself and these sorts of questions are taking on increasing prominence in the academic literature (on this topic, see, for example, Ward 2004; Finnemore 2013; Hollyer, Rosendorff, and Vreeland 2011; Wallace 2014; De Castro, Perez, and Vives 2011; Alt, Lassen, and Wehner 2014). Our secondary purpose in asking this question refers to the use of development statistics in social scientific contexts. Aid-seeking data manipulation would complicate empirical studies based on the relationship between aid and development as the countries that receive more aid would also be more likely to manipulate their data in ways that would dampen the appearance of economic growth. These problems would be compounded if certain kinds of countries were systematically more willing to report low estimates of their development levels and/or more able to have those low estimates reflected in

africas

[4] As Jerven (2013b) notes, the process by which the data reported by national statistical offices gets translated into the data reported by the World Bank is proprietary to the World Bank.

the data released for social scientists use. To that end we ask whether there exists evidence of data manipulation in general and whether such evidence is more apparent in the data from countries with more incentive or opportunity to manipulate their data.

We consider two possibilities concerning which countries should be more likely to successfully aid-seek through data manipulation. Our central hypothesis considers the possibility that governments in relatively aid-dependent countries may be more willing to appear poor on paper if it facilitates continued aid flow.[5] We also consider the relationship between data manipulation and democracy. In particular, we consider the possibility that democratic governments are viewed as being trustworthy and that this perception might, perversely, allow their manipulated estimates to receive less scrutiny by international organizations and therefore be more likely to manifest in the released data.[6]

To explore whether aid-seeking data manipulation occurs we analyze the distribution of gross national income (GNI) per capita statistics around the eligibility threshold for grants and concessional loans from the World Bank's International Development Association (IDA). These data are well suited to our purposes for two reasons. First, IDA eligibility matters. The IDA is an important capital source for poor countries, and the eligibility threshold that it uses is shared by other multilateral aid providing organizations, including the IMF's Poverty Reduction and Growth Trust (PRGT), the World Bank's Heavily Indebted Poor Countries (HIPC) Initiative, and concessional lending programs administered by the African Development Fund, the Asian Development Bank, and the Inter-American Development Bank (IDA 2012c, Annex C, Table 1). Moreover, removal from IDA eligibility typically triggers a reduction in bilateral aid flows that

[5] We thank participants at the 2014 PEIO conference for suggesting this hypothesis. We proxy for aid dependence with the 10-year lag of aid per capita. See section III for a fuller justification of the need for a proxy variable in this application and the suitability of this particular proxy.

[6] We take some care in not labeling the proposed relationship between democracy and data manipulation a hypothesis. Our intuition about this relationship evolved deductively with our analysis of the data. Given that, readers should consider this portion of paper an exploration of the data more than a formal hypothesis test. Our hypothesis concerning the role of aid-dependence on data manipulation was developed and tested inductively and the empirical analysis associated with it can be properly considered a hypothesis test. This particular hypothesis was suggested to us by audience members at the 2014 PEIO conference who, presumably, had not previously looked at the data. At the point it was suggested to us we hadn't looked for this relationship in the data either.

Knack et al. (2013) estimate at 20 percent. As such, if countries are willing and able to aid-seeking data manipulation, establishing IDA eligibility is a very plausible goal for such efforts.

A second reason to focus on the IDA eligibility threshold is that IDA eligibility is determined in a way that is amenable to discerning patterns that would indicate data manipulation. The World Bank determines IDA eligibility according to whether a country's GNI per capita, translated into US dollars using the Atlas method, is above or below a predetermined threshold. In 2012, the IDA eligibility threshold was a GNI of $1,205 per capita. If countries are able to "lowball" their GNI estimates in order to sort themselves below the IDA threshold, we should expect to observe an over-representation of "just-barely-IDA-eligible" countries and a relative under-representation of "just-barely-IDA-ineligible" countries. If aid-seeking data manipulation were more common (or successful) among democratic and/or aid-dependent countries we would expect that abnormally high density of country-year observations with GNIs per capita just below the IDA eligibility threshold to be limited to or more pronounced in those subsamples.

We use two sets of data in our analysis. The first is the GNI per capita estimates downloadable from the World Development Indicators (WDI).[7] Because most social scientists use data that are directly downloaded from WDI or similar websites, the consequences of data manipulation to researchers would be most acute if it expressed itself in these data. However, current estimates of past macroeconomic activity are often revised versions of the estimates that were first published. Any evidence of aid-seeking data manipulation would, if it exists, presumably wash out over time in the revision process. For that reason we also test for sorting behavior in the distribution of GNI per capita data taken from back issues of the World Bank annual Atlas, which provide a better approximation of the data as it first appeared in the public record.[8]

[7] We downloaded these data in August 2013.

[8] Importantly, GNI per capita data are revised between the time that they are used to determine IDA eligibility and the time that they are released in the World Bank Atlas. A such, these data are better suited to explore possible biases in the data used by social scientists than in the data used to determine IDA eligibility determination at the World Bank.

We document the extent of this clustering through three methods: 1) visual inspection of histograms of the data, 2) using McCrary tests to detect breaks in the density function of GNI per capita observations at the IDA eligibility threshold, and 3) estimating regression discontinuity models on the data to determine if being IDA eligible appears to "cause" lagged polity scores or lagged values of aid-dependence. Because we know that placement above or below the IDA threshold in the current period cannot have a causal effect on polity scores or aid-dependence in the past, "evidence "of such a phantom causal relationship indicates sorting among countries in these groups (see Lee 2008: 690–91 Caughey and Sekhon 2011: 392–393 for similar applications).

Our analysis suggests a lot of good news about data reliability. In keeping with findings in Knack et al (2013, 2014) (see also Aronow, Carnegie and Samii 2014), we find no evidence that the data *as a whole* show any signs of data manipulation. We also fail to find any statistically significant evidence that democracies are more able to sort themselves below the IDA threshold. What we do find, however, is a visually discernable over-representation of GNI per capita estimates just below the IDA threshold among aid-dependent countries in the data reported in back issues of the World Bank Atlas. McCrary tests indicate that this break in the distribution function is statistically significant. Our regression discontinuity estimates similarly suggest the appearance of a causal relationship between IDA eligibility in the current period and being aid-dependent 10 years prior.

There are two important caveats to go along with these findings. First, the statistical significance of the break in the density function in the data reported in back issues of the World Bank Atlas is sensitive to the exclusion of data from individual countries, which we explore in more detail in section II. Second, the revisions process seems to have rid the data of some of the aid-seeking induced distortions that may have originally been present. While there is a visually discernable break in the density function of current GNI per capita estimates, it is smaller and less statistically significant than the break that is evident in the data reported in back issues of the World Bank Atlas. While these findings are in no sense definitive on this matter, to the extent that the (smaller and statistically insignificant) clustering patterns in the current data are echoes of (larger and statistically significant) clustering patterns in older data releases, the results in this

paper suggests the need for some caution in the use of these data for social scientific ends. While the importance of these caveats can't be overstated, our results do suggest that a "political economy" of data production exists, which echoes recent papers by Wallace (2014) and Hollyer, Rosendorff, and Vreeland (2011) and others and suggests the need for more research into this area.

The remainder of this paper is organized as follows. Section I describes the GNI per capita measurement process and relate observations from fieldwork and anecdotal evidence to illustrate how macroeconomic data is produced in practice and how the process may sometimes be informed by national political priorities. Section II describes and presents our empirical analysis. Section III provides our conclusions.

I. The World Bank and the importance of GNI per capita cutoffs

The World Bank operates two lending divisions: the International Development Association (IDA) and the International Bank for Reconstruction and Development (IBRD). The IBRD lends at interest to middle-income and some creditworthy poorer countries.[9] The IDA serves poorer governments that either lack the creditworthiness necessary for IBRD loans or could not afford the interest rates the IBRD charges. The IDA provides grants and interest-free loans (save for a 0.75 percent service charge) that mature in 40 years and have a 10-year repayment grace period. The IDA committed $14.8 billion in 2012, 15 percent of which was disbursed in grants (IDA 2013a).[10]

Sustained eligibility for IDA loans requires a GNI per capita below a predetermined threshold.[11] The original cutoff was set at $250 in 1964 and was revised annually for inflation, reaching $940 in 1987.[12] In 1987 the IDA adopted a more restrictive operational cutoff of $580,

[9] Depending on the type of loan, interest rates currently (as of 1 July 2013) range from LIBOR plus 27 basis points to LIBOR plus 100 basis points. The IBRD does not assign a risk premium to its loans, so access to these credits is limited to countries that undergo a fairly rigorous qualitative determination of creditworthiness.

[10] World Bank 2013b. To put this number in context, according to the World Development Indicators, the IDA member countries received $57.2 billion in net official development assistance and official aid combined. World Development Indicators, retrieved 2013 from http://data.worldbank.org/.

[11] Countries must also be in good standing with the IMF to secure an IDA loan or grant.

[12] Ravillion (2013) recounts that the use of GNI-based thresholds at the World Bank dates back to 1971, when it was

which has since been annually revised for inflation (using the SDR inflation rate) but remains constant in real terms. The threshold for calendar year 2012 (fiscal year 2014) is $1,205. Countries with a GNI per capita that is persistently above the threshold but are not creditworthy at the IBRD (the so-called gap countries) can borrow from the IDA, but at less concessional rates.[13] For the most part, however, when a country's GNI per capita rises above the threshold, it begins a multi-year graduation process from the IDA to the IBRD.

Graduating from the IDA is an indication of positive developments, but it comes with substantial costs. New IDA loans cease and the repayment schedule for existing IDA loans accelerates. The accelerated repayment schedules push up debt service costs at the same time that graduating countries face the higher borrowing costs from the IBRD and on private capital markets. IDA graduates may encounter reduced access to capital because the increased burden of non-concessional loans can erode their borrowing capacity (IDA 2012c, 14). Bilateral aid flows slow significantly because many donor countries interpret graduation from the IDA as a signal that a country is no longer in dire need (Knack et al. 2013). In short, moving above the IDA's GNI per capita threshold can be an expensive form of progress.

GNI per capita (Atlas Method)

GNI represents the sum of domestic value added by all resident producers plus net receipts of primary income from abroad. While some components of GNI are likely to be relatively precise (aid inflows or foreign debt service, for example), the sum of domestic value added by all resident producers cannot be tracked so precisely and must therefore be estimated, first by the national statistical office and subsequently in revisions made by the World Bank. The GNI per capita estimates that the World Bank uses to define IDA eligibility are normalized by midyear population estimates and converted to dollars using the Atlas method, a conversion

---

decided that countries with less than $200 GNI per capita would be allowed to exercise the so-called civil works preference, meaning that they are allowed to offer preference to national suppliers when procuring certain goods and services.

[13] The financial particulars of loans to gap countries have evolved over time and have generally become less concessional and accompanied by more conditionality. Currently, gap countries can access loans with 25-year maturities and five-year grace periods and are charged interest rates up to the IBRD rate minus 200 basis points. The IDA also lends at less concessional rates to so-called blend countries that are sufficiently poor but maintain access to international credit markets. Small island nations are automatically eligible for IDA lending.

factor that averages the exchange rate between the local currency and US dollars in year t, and inflation adjusted exchange rates in years t–1 and t–2.

The GNI per capita statistics used to establish IDA eligibility by the World Bank suggest development levels that are generally similar to those implied by GDP per capita, which is unsurprising, given their conceptual similarities. Table 1 (next page) contrasts the World Development Indicators' GNI per capita data with its GDP per capita data. Column 1 rank orders the 25 countries with a GNI per capita within $500 of the IDA threshold in 2012. Column 2 shows every country with a GDP per capita that is at least as high as the poorest country from column 1 in GDP per capita terms (Haiti) and is no higher than the richest (Nicaragua). Countries with GNI per capita statistics below the IDA threshold are shaded in gray. As Table 1 indicates, the 9 poorest countries according to GNI per capita are among the 12 poorest countries according to GDP per capita (and are joined by Mali, Tanzania, and Burkina Faso, whose GNI per capita is lower than Chad's and are therefore left out of column 1).

The rankings are less consistent towards the higher end of the distribution. Nicaragua is still the richest country in the sample, but Ghana, which is the second-richest country according to GNI per capita, is in the middle of the pack in terms of GDP per capita and poorer in GDP per capita terms than several countries with GNI per capita below the IDA threshold. Senegal, Mauritania, Cameroon, and the Solomon Islands, all have GNI per capita below the IDA eligibility threshold but are richer in GDP per capita terms than many countries that do not.

Where do development indicators such as GNI and GDP statistics come from? The aggregation of macroeconomic statistics is governed by a global standard: the United Nations System of National Accounts (SNA). Committees of statistical experts during the interwar years laid out the foundations of this system, which was published in 1953 by the UN as *A System of National Accounts and Supporting Tables*.[14] The nationally produced data are collected by

[14] The standards of national accounts have since been revised three times; thus, there are four versions. In addition to SNA 1952, there are also SNA 1968, SNA 1993, and SNA 2008. However, Ward (2004) argues that "although they pay lip service to the subsequent revisions . . . many countries still adhere to the basic system and its corresponding accounting foundations as first set out" (45).

Table 1: Rank Ordering of countries near the IDA threshold in 2012 by GNI and GDP per capita.
IDA eligible countries are highlighted

Rank	Country	GNI PC	rank	Country	GDP PC
1	**Chad**	740	1	**Haiti**	459.698
2	**Benin**	750	2	**Tajikistan**	462.482
3	**Haiti**	760	3	**Mali**	476.445
4	**Kenya**	840	4	**Tanzania**	483.482
5	**Comoros**	840	5	**Burkina Faso**	494.99
6	**Bangladesh**	840	6	**Chad**	507.375
7	**Tajikistan**	860	7	**Benin**	567.914
8	**Cambodia**	880	8	**Kyrgyz Republic**	574.818
9	**Kyrgyz Republic**	990	9	**Kenya**	592.924
10	**Senegal**	1040	10	**Bangladesh**	597.494
11	**Mauritania**	1110	11	**Comoros**	606.007
12	**Solomon Islands**	1130	12	**Cambodia**	671.636
13	**Cameroon**	1170	13	Timor-Leste	690.827
14	Ivory Coast	1220	14	Lao PDR	707.19
15	Pakistan	1260	15	Ghana	724.352
16	Lao PDR	1260	16	Zambia	798.26
17	Sao Tome and Principe	1320	17	**Senegal**	799.39
18	Zambia	1350	18	Pakistan	802.451
19	Lesotho	1380	19	**Mauritania**	835.148
20	Vietnam	1400	20	Sudan	836.336
21	Nigeria	1430	21	Sao Tome and Principe	840.385
22	Sudan	1450	22	Uzbekistan	845.74
23	India	1530	23	Lesotho	928.537
24	Ghana	1550	24	Vietnam	931.031
25	Nicaragua	1650	25	Ivory Coast	957.884
			26	**Cameroon**	960.558
			27	Moldova	1038.4
			28	Nigeria	1052.34
			29	Kiribati	1072.89
			30	Papua New Guinea	1076.38
			31	India	1106.8
			32	**Solomon Islands**	1145.4
			33	Bolivia	1259.81
			34	Guyana	1276.8
			35	Nicaragua	1349.89

Source: World Bank Word Development Indicators

international organizations such as the UN and then disseminated in different datasets (Jerven 2013a). The United Nations Statistical Office collected the data and oversaw the system at the outset, but its role was supervisory and it lacked access to the raw data. According to Michael Ward (2004), this arrangement initially worked well, but particularly after the petroleum crises in 1973 and 1978, "the practical work on national accounts both in UNSO and in the member countries got farther and farther behind. Increasingly, other international agencies and the major donors began to express their frustration at the poor quality and timeliness of the national accounts data of their client countries" (96–97).[15]

Ward describes how the World Bank felt concern about timeliness particularly strongly and notes that despite the fact that "the Bank never had a mandate to compile statistics and was never involved in actual basic data collection for the national accounts," it started publishing its own GNI per capita numbers in US dollars in the 1980s (Ward 2004, 98). These estimates sometimes contradicted official data, but they became widely accepted, "because they appeared more current and consistent" (Ward 2004, 98). Furthermore, though "interactions with government at the highest level ensured that [the World Bank's] officials were granted access to data that others did not have," it "also meant that numbers that were subsequently generated were, in some political sense, endorsed by those in authority" (Ward 2004, 99). The World Bank moved towards the adoption of more fragile but 'up to date' figures that represent "agreed upon numbers that were accepted by both the Bank and its respective country members as the basis for their policy dialogue" (Ward 2004, 100).

As Jerven (2013b, ch. 4) has described for African economies, the process of agreeing upon the final numbers can be political. National statistical office representatives present forthcoming GDP numbers and World Bank representatives decide whether or not to accept the estimates the statistical office has produced. Jerven (2013b, 99) notes that consultants who have helped prepare estimates or have been involved in meetings where new estimates are discussed confirm that concerns about sustained eligibility at the IDA come up frequently. If there is no

[15] This was also when the World Bank started converting to US dollars using the Atlas method in order to deal with volatile nominal exchange rates. The World Bank also made its own midyear population estimates. See Ward (2004, 97–99) for the full account.

agreement or if there are no regular close relations between IMF or World Bank missions and the national government, the country and the World Bank may publish different numbers (see Jerven 2013b, Table 1.2, 24–25). World Bank data often get revised substantially and the process is documented in the World Bank Group Archives.[16] The revised figures are not published until the country economist and the World Bank's International Economics Department have approved them.

II. Evidence

The empirical portion of this paper asks whether the process through which GNI per capita statistics are produced allows countries to manipulate their data in order to maintain or establish IDA eligibility. Doing so, we should reiterate, would not be easy. The World Bank revises these data for accuracy, and the population data used as the denominator in GNI per capita estimates are released after the macroeconomic data used in the numerator has been established. Nonetheless, if countries were able to manipulate their data, there should be evidence in the form of an abnormally high density of country-year observations just below the IDA eligibility threshold and a relatively low density of country-year observations just above it. Our expectations in this regard are informed by the regression discontinuity literature (see, for example Lee 2008; Imbens and Lemieux 2008). Regression discontinuity is a method for establishing the causal effect of treatment that is administered to units on the basis of their being above or below a threshold value of some continuous measure (the assignment variable). Regression discontinuity compares outcomes between a treated group of observations just above/below and a control group of observations just below/above the threshold under the assumption that those units are otherwise comparable to each other except for their different exposures to the treatment. The comparability of these groups requires that the units under study did not sort themselves into or out of the treatment group. McCrary (2008) proposes a formal test of this assumption by looking for discontinuities in the density function of the assignment

[16] Folder 1353586, WB IBRD/IDA DEC-03, Records of the Office of the Vice President, Development Economics and Chief Economist and later Senior Vice President, Development Economics and Chief Economist (DECVP), World Bank Group Archives, Washington, DC.

variable at a threshold: "too many" observations just below/above the treatment threshold is an indication that such sorting behavior occurred.

In our application we consider aid-seeking data manipulation as being analogous to the sorting behavior the McCrary test is meant to identify. In the absence of data manipulation the density function of GNI per capita observations should pass smoothly over the IDA eligibility threshold; if countries are able to manipulate their data in order to remain IDA eligible there should be a discontinuity in the density function at the IDA eligibility threshold with too many observations just below it. If some countries but not others were able to manipulate their data we would expect these discontinuities to exist in in subsamples.

We consider two possibilities concerning the conditions under which manipulated GNI estimates might be most likely to manifest themselves in the data. The first considers the motivations that countries have to keep themselves IDA eligible. The benefit of being IDA eligible is the promise of more concessional aid, both from the IDA as well as from bilateral and other donors that use the IDA eligibility benchmark in their own lending decisions. Doing so would mean that a country would appear poorer than necessary, however, which may come at a political cost to the government. Indeed, as Aronow, Carnegie and Samii (2014) show, graduating from World Bank lending programs can be a point of pride for governments and presented as evidence of their competence. To the extent that countries consider the benefits of appearing richer against the benefits of remaining IDA eligible, we hypothesize that greater levels of aid dependence would motivate countries towards biasing macroeconomic statistics downwards in an attempt to remain IDA eligible.

The second potential causal mechanism that we consider concerns countries' capacity to have their opportunistically low estimates represented in World Bank released data (assuming that countries are choosing opportunistically low estimates in the first place). The divide in this case is not which countries want to manipulate their data to remain IDA eligible, but which countries can successfully do so. We think it is plausible that the data emanating from countries that are otherwise perceived as being more credible and transparent should face less scrutiny. While perceptions of credibility are not directly observable, we think democracy should make for a good proxy. Democracies are empirically associated with lower levels of corruption than

autocracies (Lederman, Loayza, and Soares 2005) and democracies – especially longstanding ones – are often *perceived* to be less corrupt in country risk ratings, domestic and international business surveys, and citizen polls (Treisman 2007). More specific to this application, Hollyer, Rosendorff, and Vreeland (2011) suggest that the greater levels of transparency in democratic regimes extend to their release of more macroeconomic data. The World Bank itself, while officially neutral with respect to governing institutions in recipient countries, appears to place particular value on democratic institutional development (Heckleman et al. 2011, Aronow, Carnegie and Samii 2014). However, to the extent that a general perception of democracies as trustworthy extends to the reliability of their data, this trust is likely misplaced. De Castro, Perez, and Vives (2011) note that releases of macroeconomic data before elections tend be subject to larger subsequent revisions. Alt, Lassen, and Wehner (2014) show that politically informed manipulation of macroeconomic data is widespread within European democracies (see also, Von Hagen and Wolff (2006) and Milessi-Ferretti (2004)).[17]

Method

We use three methods to establish whether GNI per capita data cluster in ways that suggest manipulation: 1) visual inspection of histograms noting the density of country-year observations of GNI per capita near the IDA threshold and 2) McCrary tests of discontinuities in density function around a threshold (McCrary 2008), and 3) regression discontinuity models looking for an "effect" of IDA eligibility in the current period on past realizations of the conditioning variable.

Sample

We restrict our analysis to the time since 1987, when the IDA established the operational cutoffs that are currently used. We exclude countries from our sample that are not members of the World Bank and the small island nations that are eligible for IDA lending regardless of their

[17] We acknowledge alternative causal mechanisms that can plausibly relate democracy to data manipulation. For example, if we are right about democracies being perceived as more credible than autocracies, it may the case that they actually are and therefore less likely to engage in aid-seeking data manipulation. Similarly, democratic governments might reasonably be expected to place relatively more weight on their ability to take credit for a growing economy, which also suggests that they should be less likely to disproportionately appear just below the IDA threshold.

GNI per capita.[18] We limit the graphical representations of our analysis to country-year observations with GNIs per capita within $1,000 on either side of the IDA eligibility threshold. While we cannot know how near a country's "true" GNI per capita must be to the IDA threshold for its statistical office to plausibly claim a GNI that would place it below the IDA eligibility threshold, we think $2,000 is a large enough range to capture the full set of country-year observations for which IDA eligibility-seeking behavior may be relevant. This restriction is entirely inconsequential to our statistical tests, which focus on the density of GNI per capita observations in bands around the IDA threshold that are much narrower than $1,000.

Data

Our variable of interest is a country's GNI per capita (converted into dollars using the Atlas method) minus the IDA eligibility threshold.[19] The resulting variable is $0 for countries whose GNI per capita is exactly at the eligibility threshold, negative for countries with GNIs per capita below the IDA eligibility threshold and positive for countries with GNIs per capita above the IDA eligibility thresholds.

We use two different sources of GNI per capita data. The first is data that was downloaded from the WDI website in August 2013. These are the data that are most commonly used by researchers and represent the World Bank's current (as of the time of download) estimate of a country's GNI per capita. The accuracy of these data comes in part from a periodic and often substantial revision process. While the revision process almost certainly generates more accurate data, it may in doing so obscure the clustering patterns that would arise from data manipulation. For that reason we also examine GNI per capita figures gathered from back issues of the print editions of the World Bank's Atlas. These data are less revised and provide a better proxy for the data at the time they are first made available to researchers. These data are often revised from the time that IDA eligibility is established, however, and so they differ to some degree from the GNI per capita estimates that are used to establish IDA eligibility.

[18] For a list of these countries, see IDA 2012a.

[19] Threshold data is available at siteresources.worldbank.org/DATASTATISTICS/Resources/OGHIST.xls.

Our first set of tests mirrors diagnostic tests carried out with respect to the IDA threshold in in Knack et al. (2013) and Knack et al. (2014) and with respect to the IBRD eligibility threshold in Aronow, Carnegie and Samii (2014). Here we ask whether there is any indication of sorting behavior in the data overall, without considering any differences across sub-samples. The histogram in Figure 1 shows the distribution of GNI per capita for all country-year observations in our sample. The X-axis indicates the re-centered GNI per capita data. The vertical line marks the IDA threshold at 0. We separated the data into 16 equal-sized bins of $50 each, so that the column to the left of the IDA threshold represents the number of observations with re-centered GNI per capita less than 0 and greater than or equal to –$50, the next column to the left represents the number of country-year observations with GNI per capita less than –$50 and greater than or equal to –$100, and so forth. The top panel relates to the downloaded data and the bottom panel relates to the historical data. Neither histogram suggests an obvious abnormality in the distribution function. Both histograms show that there are more data just below the IDA eligibility threshold than just above it, but in both cases the overall trend in the density function is downward sloping and the patterns around the IDA threshold are not obviously out of line with that trend.

The results of a McCrary test confirm that impression. As noted above, the McCrary test establishes the existence (or non-existence) of discontinuities in a density function at a specified threshold.[20] It does so by calculating the density of observed data within discrete bins and using local, non-parametric regressions to estimate the density of observations at the threshold. The McCrary test does this separately for data on both sides of the threshold and calculates a test statistic based on the log difference between the two estimates. If no sorting occurs the discontinuity between the two density estimates at the threshold should be statistically indistinguishable from zero (satisfying the requirements for regression discontinuity research designs). Sorting, however, should lead to a statistically significant difference between the two density estimates at the threshold. Because the discontinuity estimate is calculated as the density

[20] We implement this tests using the DCdensity package in Stata. http://emlab.berkeley.edu/~jmccrary/DCdensity/DCdensity.ado.

above the threshold minus the density below the threshold, aid-seeking data manipulation would yield a negative test statistic.

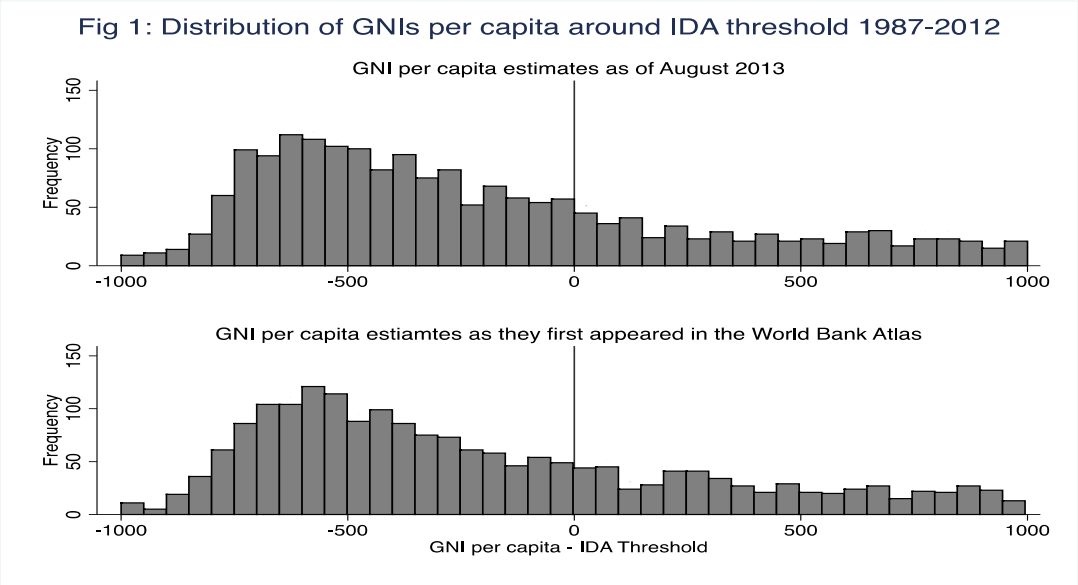
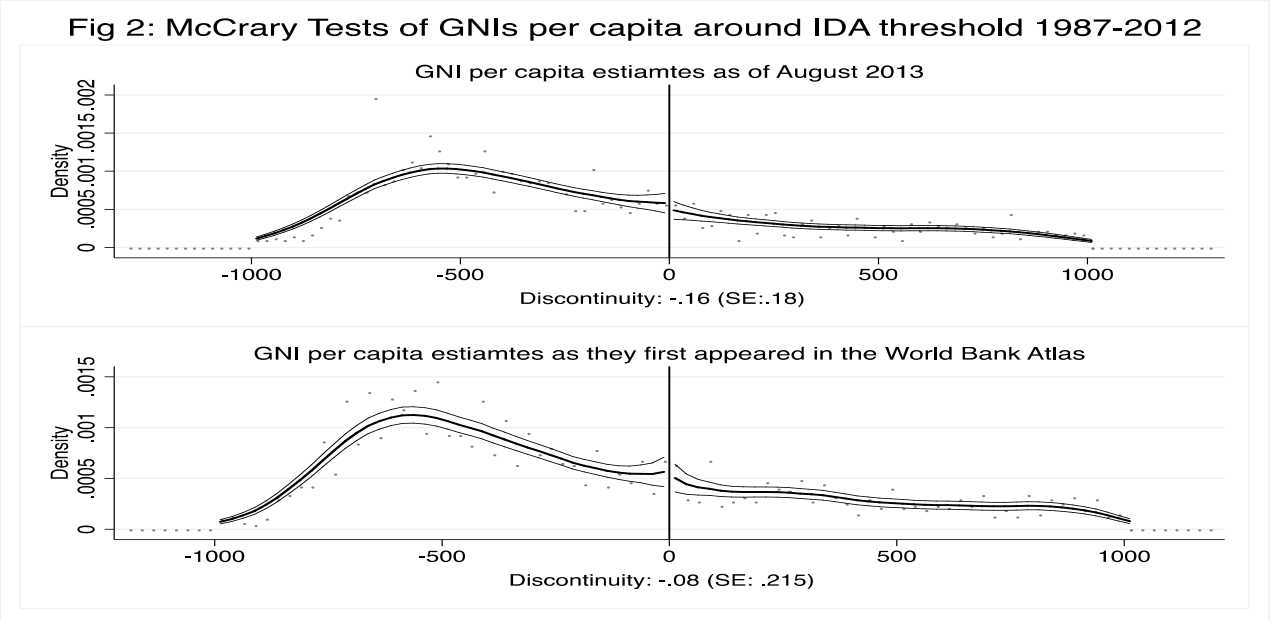


Figure 2 shows the graphical output from McCrary tests estimated with a bandwidth of 300.[21] The thicker lines show the estimated density function and the thinner lines show a 95% confidence interval. The vertical line at 0 indicates the IDA eligibility threshold. These estimates indicate a log difference between the densities at the threshold of –.08 with a standard error of .22 in the historical data and a log difference of –.16 and a standard error of .18 in the recently downloaded data. Neither estimate is statistical significant. These results are consistent with the diagnostic tests performed in Knack et al. (2013), Knack et al. (2014) and Aronow, Carnegie and Samii (2014).[22] While not the main purpose of this paper, this non-finding is notable (and comforting) in itself.

[21] Kernel based estimators such as this are highly sensitive to bandwidth selection (see McCrary 2008, Imbens and Lemieux 2003). Our main tests all employ a bandwidth of 300. We note in the text how these estimates compare to estimates derived using alternative bandwidths of 100, 200 and 400. In general, there is no meaningful difference in the results regardless of the bandwidth employed. Except where otherwise noted, we use the McCrary tests default bin width selector, which is equal to 2*sd(runvar)*length(runvar)^(-.5), where the "runvar" refers to the GNI per capita – the IDA threshold. See McCrary (2008).

[22] These non-findings persist in alternative specification using bandwidths of 100, 200 and 400.



Aid-dependent countries sort into IDA eligibility

The main purpose of this paper is to explore the possibility that certain kinds of countries are more likely and/or more able to sort themselves into IDA eligibility by claiming low GNI estimates. Our first hypothesis is that aid-dependent countries should be more willing to appear poorer if doing so helps them maintain access to aid. Our first strategy to evaluate this hypothesis is to bisect the data into aid-dependent and an aid-independent subsamples, and replicate the analyses carried out above on the full sample on the subsamples. We measure aid dependence using the World Bank's measure of net ODA per capita and separate the sample into an aid-dependent group whose values of aid per capita is at or above the median and an aid independent sample whose values of aid/GNI is below the median.[23]

The most direct measure of aid dependence would utilize the 1-year lag of aid/per capita, but this measure is potentially compromised by reverse causality. Because crossing the IDA threshold decreases aid flow we would expect to observe a degree of clustering of aid-dependent countries just below the IDA threshold whether or not the data is being manipulated. The one year lag mitigates that somewhat, but GNI per capita moves slowly enough relative to the IDA threshold that being just below the IDA threshold in year t–1 is highly correlated with being

[23] The World Bank's measure is stated in current dollars, which we converted into constant 2005 dollars.

below the IDA threshold in year t, making it unlikely that a one-year lag of aid dependence is solely capturing an exogenous component of aid dependence. To illustrate, from 1988 onward, 92% of country-year observations that were within $300 under the IDA eligibility threshold were also under the IDA eligibility threshold in the year prior compared to only 20% of country-year observations that were within $300 over the IDA eligibility threshold had a GNI per capita under the IDA eligibility threshold in the prior year. This difference is statistically significant at the .001 level.[24]

To avoid this problem we use the 10-year lag of aid dependence as a proxy for aid dependence at the time that a country submits its data to avoid the problem. The intuition behind this measure is twofold. First, many aspects of aid dependence are long-lived and countries that were aid-dependent 10 years prior are likely to be aid-dependent in the current period.[25] Second, the 10-year lag structure is sufficiently long to be confident that it is capturing components of aid dependence other than IDA eligibility. To be precise, 78% of country-year observations in our sample that are within $300 above the IDA threshold in year t were below the IDA threshold in year t–10 compared to 82% of country-year observations that are within $300 below the IDA threshold in year t–10. This difference is not statistically significant (p value: 0.22). A Kolmogorov-Smirnov test confirms the lack of statistically significant distinctions in the distribution of the 10-year lagged variable across country-year observations above and below the IDA eligibility threshold and within $300 of it.[26] We are therefore confident that the 10-year lagged measure of aid per capita acts as suitable instrument for aid dependence in the current period. We considered using infant mortality measured in 1980, which has been used as an instrument for aid dependence elsewhere in the literature (ex. Burnside and Dollar 2000, Knack 2001) but found that it is not well correlated with aid per capita in our sample. Regressions demonstrating as much can be found in the appendix.

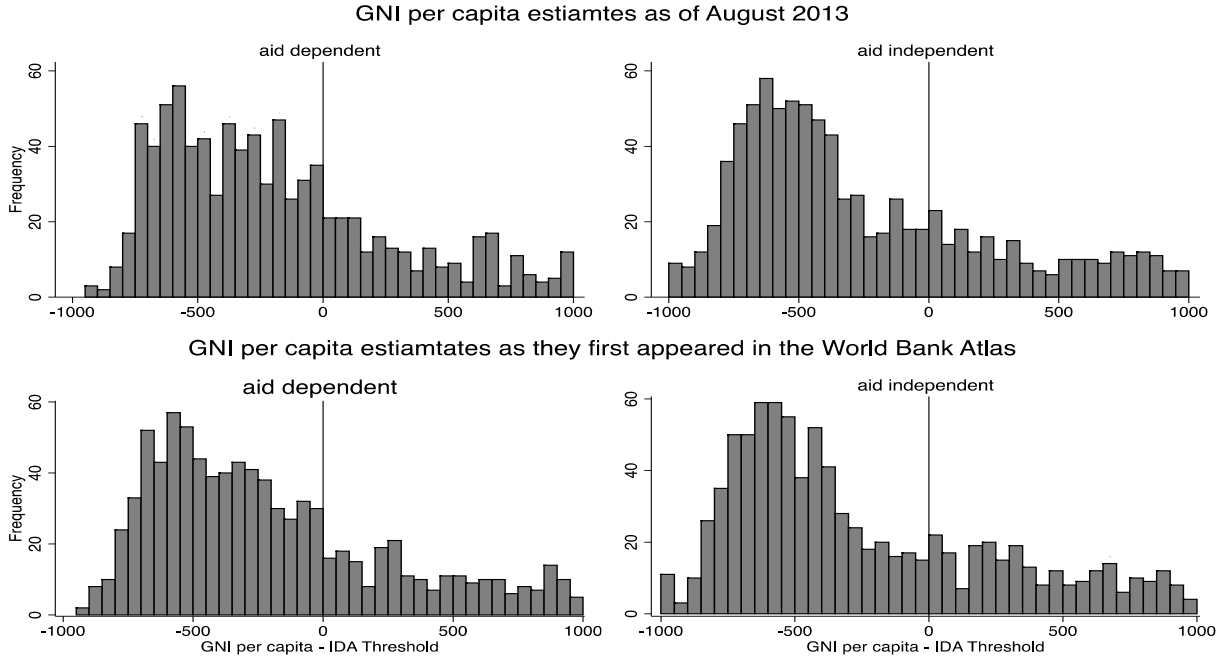
---

[24] We are particularly concerned about differences in lagged eligibility within a $300 of the IDA threshold given our use of a $300 bandwidth in our primary McCrary tests. A Kolmogorov-Smirnov test confirms the unsuitability of this variable for our purposes.

[25] The correlation between aid per capita and its 10-year lag is .578 in our sample and highly statistically significant.

[26] The five-year lag of aid per capita failed this test.

Fig 3: Distribution of GNIs per capita around IDA threshold 1987-2012



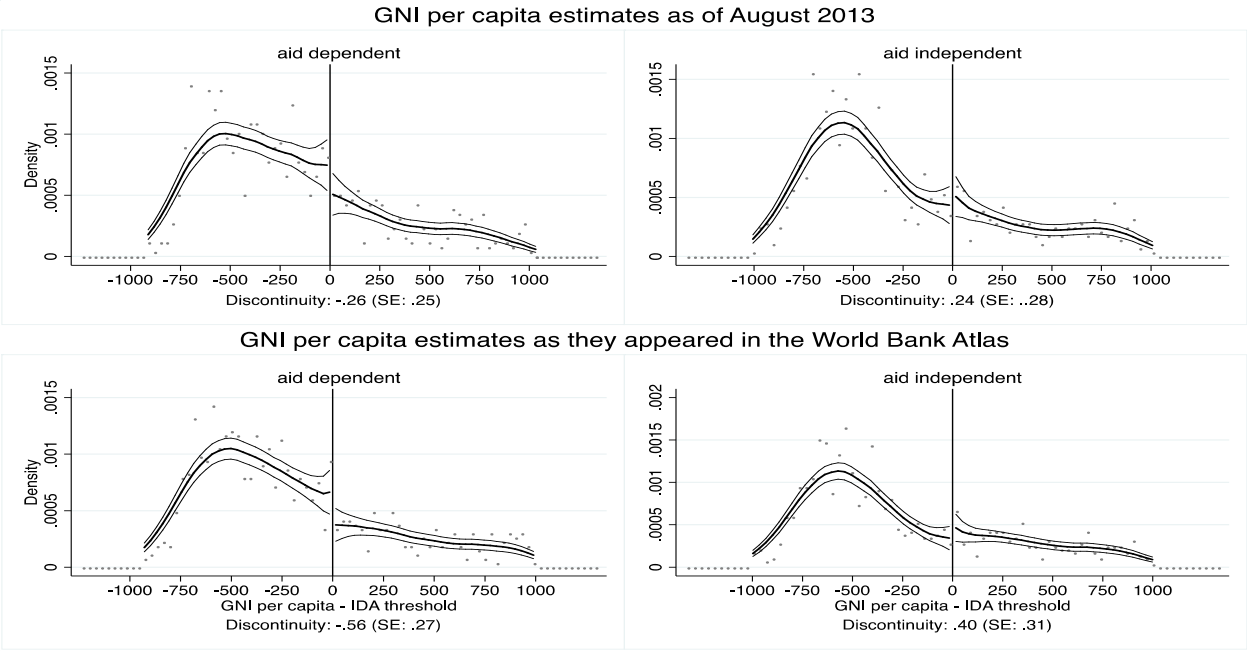
see text for definition of aid dependence

Figure 3 shows the distribution of GNIs per capita for the two datasets, with the data for each divided according to whether or not that country-year's 10-year lag of aid dependence is above or below the sample median. The two left-side histograms refer to country-year observations that are relatively aid-dependent. Both plots exhibit a noticeable gap in the density function at the IDA eligibility threshold. In the data taken from the World Bank Atlas, the 31 observations between –$50 and $0 are almost twice as numerous as the 16 observations between $0 and $50. The 63 observations between –$100 and $0 are similarly nearly twice as numerous as the 35 observations between $0 and $100. There is a similar, but slightly more muted pattern in the more recent GNI per capita estimates. In these data there are 32 observations between –$50 and $0 compared to 21 observations between $0 and $50 and 62 observations between –$100 and $0 compared to 40 observations between $0 and $100. Another difference between the two distributions is that the density function for the data taken from the World Bank Atlas is relatively smooth in areas other than the IDA threshold, which makes the discontinuity at that threshold appear more like a break in the density function to the naked eye. The density function

for the more recent estimates is relatively choppy, and the apparent discontinuity at the IDA threshold is not as obviously out of line with the overall distribution of the data.

Figure 4 shows the graphical output from the McCrary tests on the data separated in the same manner as in Figure 3. The results of the McCrary tests largely reflect the intuition in the above histograms. In the more recent GNI per capita estimates there is a discontinuity in the aid-dependent sample in the hypothesized, negative direction (–.26, standard error: .25), but it is statistically insignificant. The discontinuity estimate in the data taken from the World Bank Atlas, however, is larger (–.56, standard error: .27) and statistically significant at the .05 level (t statistic: – 2.08).[27] While aid-independent countries show a slight discontinuity in the opposite direction in both data sets, these discontinuities do not approach statistical significance in either.

Fig 4: McCrary Tests of GNI per capita around IDA threshold 1987-2012



see text for definition of aid dependence

The finding that there are "too many" aid-dependent country-year observations with GNIs per capita just below the IDA threshold in the data reported in the World Bank Atlas is consistent with data manipulation by aid-dependent countries. This is a provocative result and we

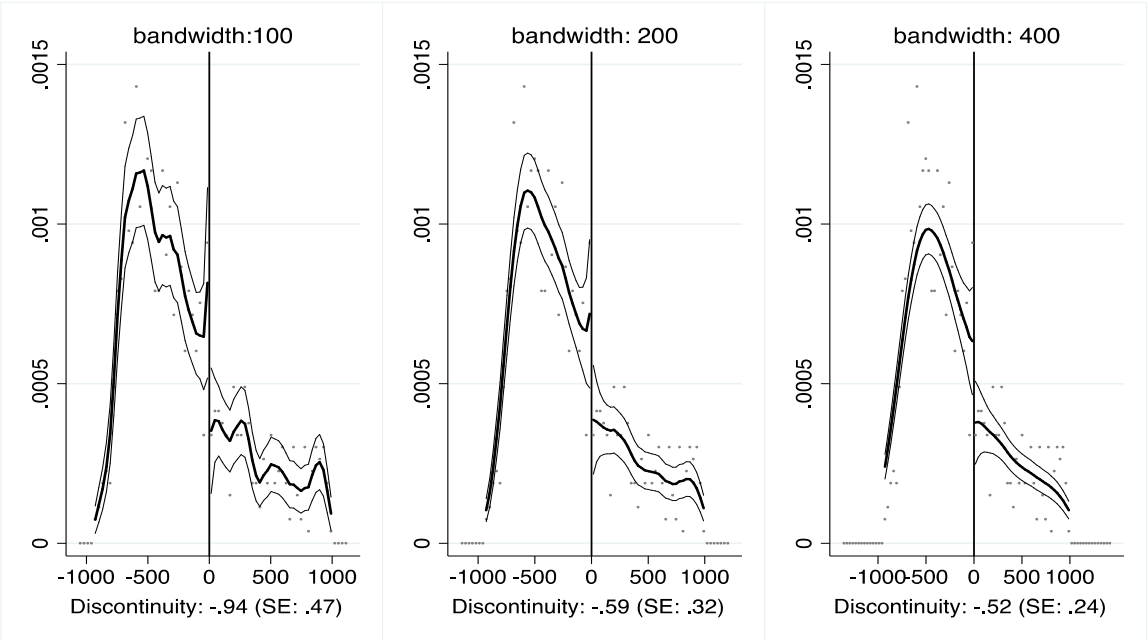
[27] Note that the slight overlap of the confidence intervals does not preclude there being a statistically significant difference between the two estimates (Schenker and Gentleman 2001).

re-estimated the analysis pictured in the bottom left quadrant of Figure 4 in several ways to explore its robustness. First, we re-estimated the analysis using alternative bandwidths. Figure 5 shows graphically the results of McCrary tests estimated with bandwidths of 100, 200 and 400. The basic result of a statistically significant discontinuity in the density function in the expected direction obtains in each estimate. However, the size of the estimated discontinuity varies substantially, and its statistical significance when estimated with a bandwidth of 200 falls below the conventional .05 level, but retains statistical significance at the .1 level (t statistic of 1.81).

Second, we re-estimated our McCrary test using alternative bin-widths. The McCrary test defaults to a bin-width of 30.4 in the above models. Figure 6 shows the results of (roughly) halving and doubling the bin-width to 15 and 60. Both estimates suggest a discontinuity in the expected direction and of roughly comparable size to those reported in other tests. Discontinuity estimates using a bin-width of 15 retain statistical significance at conventional levels, though estimates obtained using a bin-width of 60 only achieve statistical significance at the .1 level (t statistic of 1.85).

Fig 5: McCrary Tests of GNI per capita around IDA threshold 1987-2012

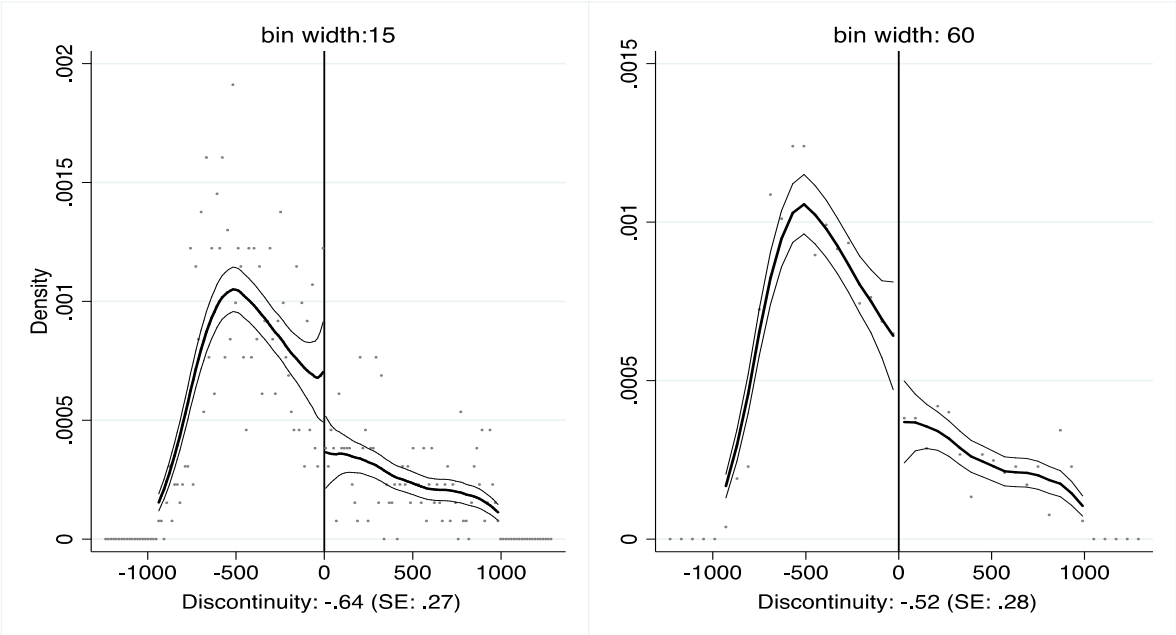
GNI per capita estiamtes as they appeared in the World Bank Atlas; aid dependent countries



see text for definition of aid dependence

Fig 6: McCrary Tests of GNI per capita around IDA threshold 1987-2012

GNI per capita estiamtes as they appeared in the World Bank Atlas; aid dependent countries



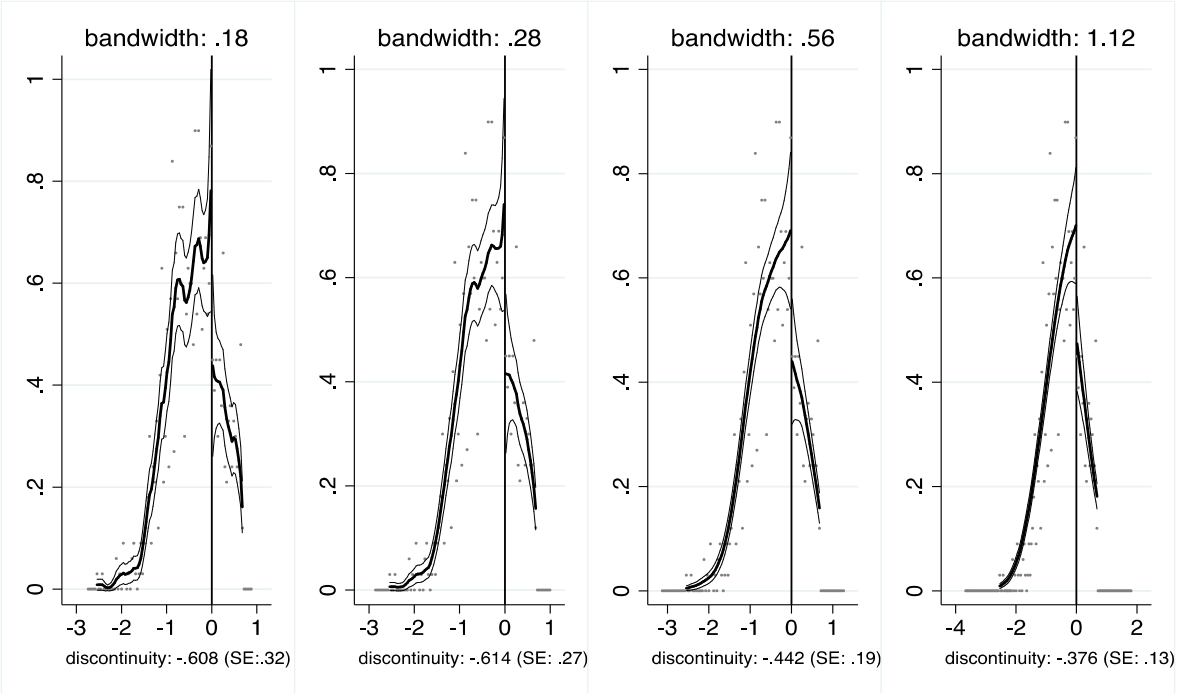
see text for definition of aid dependence

Third, we re-estimated all of our McCrary tests using the log of GNI per capita rather than the unlogged version.[28] Using logged GNI per capita reduces the left skew in the data and mitigates the possibility that this skew could be driving the results. The logged variable exhibits patterns that are extremely similar to what we observe in the unlogged variable. Figure 7 shows the results of McCrary tests estimated on aid dependent countries using the logged GNI per capita variable across a range of bandwidths. Each estimate suggests a negative and statistically significant discontinuity, which is consistent with data manipulation.

Fourth, we estimated a series of regression discontinuity models to look for evidence that being IDA eligible "causes" countries to be aid-dependent 10 years earlier (see Lee 2008). Such as causal effect would, of course, be illusory, but it would indicate that aid-dependent countries have sorted themselves below the IDA threshold. A benefit of this test is that it treats aid

[28] Our re-centered GNI per capita variable has negative values. To take the log of we added the minimum value + 1 (in this case, 1016), logged the resulting variable and then subtracted the log of 1016 to re-center the variable so that the IDA threshold is at 0.

Fig 7: McCrary Tests of GNI per capita around IDA threshold 1987-2012
GNI per capita data as they appeared in the World Bank Atlas; aid dependent countries



see text for definition of aid dependence

dependence as a continuous variable, rather than forcing us to artificially bisect the data at the median to designate aid-dependent and aid-independent sub-samples. Another benefit to this test is that it allows us to cluster standard errors in order to better accommodate the panel nature of the data. We estimate the following equation:

$$\text{Aid per capita}_{it-10} = \alpha + \beta_{01}\text{GNIPCrc}_{it} + \beta_{02}\text{GNIPCrc}^2_{it}\ldots + \rho\text{IDA}_{it} + \beta_{11}\text{IDA}_{it}*\text{GNIPCrc}_{it} + \beta_{12}\text{IDA}_{it}*\text{GNIPCrc}^2_{it}\ldots + \varepsilon_{it}$$

where the i subscript indicates country and the t subscript indicates year. IDA is a dichotomous variable coded 1 if a country has a GNI per capita value at or below the IDA eligibility threshold and 0 otherwise and GNIPCrc is GNI per capita re-centered such that 0 indicates a GNI per capita exactly at the IDA threshold. We estimate these models with OLS using a triangular kernel weighting scheme (the results are nearly identical using a rectangular kernel) and standard errors clustered by country. ρ captures the difference in the 10-year lag if aid per capita between countries IDA eligible countries and IDA ineligible countries, estimated at the IDA eligibility

threshold (when GNIPCrc is equal to 0). A statistically significant ρ would indicate a discontinuity at the IDA threshold in the relationship between GNI per capita and the 10-year lag of aid per capita. In normal applications this discontinuity would indicate the causal effect of IDA eligibility on the 10-year lag of aid per capita. In this case any "effect" would be spurious; IDA eligibility in year t cannot cause aid dependence to go up 10 years earlier. Rather, a positive and statistically significant ρ suggests at aid dependent countries were able to sort themselves below the IDA threshold.

Table 2: Regression Discontinuity Estimates of IDA eligibilityit on Aid Per Capita$_{it-10}$

Model #	1	2	3	4	5	6	7	8
Bandwidth	100	200	300	400	100	200	300	400
	coef./SE	coef./SE	coef./SE	coef./SE	coef./SE	coef./SE	coef./SE	coef./SE
IDA	48.280**	37.530**	35.690**	36.593*	51.13	55.566**	67.164**	46.278*
	(21.74)	(17.54)	(16.39)	(18.38)	(60.26)	(25.97)	(26.92)	(25.50)
GNIPCrc	-0.105	0.039	0.092	0.073	-0.688	-0.996	0.727	-0.062
	(0.26)	(0.16)	(0.12)	(0.12)	(4.06)	(1.07)	(0.64)	(0.63)
IDA*GNIPCrc	0.561	0.123	-0.012	0.035	1.065	2.684	0.094	0.597
	(0.60)	(0.31)	(0.18)	(0.14)	(3.45)	(1.71)	(1.16)	(0.95)
GNIPCrc^2					0.023	0.017	-0.007	0.001
					(0.10)	(0.02)	(0.01)	(0.00)
IDA*GNIPCrc^2					-0.039	-0.001	0.012	0.002
					(0.15)	(0.02)	(0.01)	(0.00)
GNIPCrc^3					0.000	0.000	0.000	0.000
					(0.00)	(0.00)	(0.00)	(0.00)
IDA*GNIPCrc^3					0.00	0.000*	0.00	0.00
					(0.00)	(0.00)	(0.00)	(0.00)
Constant	75.668***	73.827****	69.424****	71.493****	77.453**	82.972***	60.139**	74.802***
	(23.33)	(18.82)	(18.95)	(18.85)	(35.54)	(28.93)	(25.12)	(24.96)
r2	0.051	0.019	0.01	0.014	0.057	0.035	0.019	0.017
N	165	304	497	706	165	304	497	706

All models estimated using OLS with standard errors clustered by country. * 0.10 ** 0.05 *** 0.01 **** 0.001

Table 2 reports the results of this model using all combinations of bandwidths of 100, 200, 300 and 400 and first and third order polynomial strings of GNIPCrc.[29] The key information in Table 2 is the coefficient associated with the IDA variable, which corresponds to the ρ term. Models 1–4, in which GNIPCrc enters the equation linearly, all suggest that that IDA eligibility

[29] The results of models using second order polynomials are virtually identical to those using first order polynomials and included in the replication files. The replication files also contain regression discontinuity estimates using the rdrobust command in stata (see Calonico, Cattaneo and Titiunik 2014). These estimates are very similar to the regression estimates that we report here.

has a positive and statistically significant "effect" on aid per capita 10 years prior. Models 5–8 replicate models 1–4 except that GNIPCrc enters the equation as a third order polynomial. These estimates are similar, though at the narrowest bandwidth ($100 above and below the IDA threshold), the effect is statistically insignificant. While the failure to find a statistically significant effect in model 5 is notable, on the whole we think these estimate provide fairly strong support for the possibility of aid-seeking data manipulation by aid-dependent countries.

As a final robustness check we re-estimated the McCrary test multiple times, dropping one country from the sample in each iteration. Given the small sample size, it is possible that the discontinuity is being driven entirely by a single country's data. The results of this exercise are summarized in Table 3. Table 3 summarizes the results of this exercise by listing the excluded country, the discontinuity estimate, the standard error and the associated t statistic for each iteration in which a country's exclusion from the sample altered the discontinuity estimate by more than .01. Table 3 is sorted by the discontinuity estimate in descending order. The estimates reported in Figure 4 are noted as "Baseline Estimate".

One thing to notice is that the estimates reported in Table 3 are consistently negative, suggesting that the directionality of the discontinuity is not an artifact of data from any particular country. However, the exclusion of individual countries can have substantial impacts on the size of the discontinuity estimate and its statistical significance. In two cases – Honduras and the Ivory Coast – excluding those countries yields discontinuity estimates that are statistically significant and substantially larger than those found in the full sample. In eight cases – Zambia, Nicaragua, Bolivia, the Solomon Islands, Egypt, Djibouti and Guyana – the exclusion of that country from the sample generates discontinuity estimates that do not achieve statistical significance at the .05 level, but retain statistical significance at the .1 level. Excluding Senegal reduces statistical significance below the .1 level and reduces the estimated size the discontinuity by almost 30%. As such, the evidence reported above should be interpreted with some caution. While the data taken from back issues of the World Bank Atlas indicate a discontinuous density function that is consistent with data manipulation among aid-dependent countries, the magnitude

and statistical significance (but not the directionality) of this discontinuity is fairly sensitive to the inclusion or exclusion of a few key countries.[30]

Table 3: Discontinuity Estimates among Aid-dependent Countries Excluding One Country at a Time

Excluded Country	Discontinuity Estimate	SE	t statistic
Honduras	-0.675	0.286	-2.363
Ivory Coast	-0.647	0.282	-2.291
Sri Lanka	-0.597	0.285	-2.092
Lao PDR	-0.591	0.277	-2.132
Dominican Republic	-0.583	0.271	-2.152
Papua New Guinea	-0.583	0.283	-2.059
Equatorial Guinea	-0.574	0.277	-2.070
Armenia	-0.573	0.270	-2.122
BASELINE ESTIMATE	-0.561	0.269	-2.083
Jordan	-0.550	0.270	-2.039
Guyana	-0.549	0.291	-1.884
El Salvador	-0.547	0.269	-2.036
Jamaica	-0.547	0.269	-2.036
Congo, Rep.	-0.547	0.275	-1.990
Lesotho	-0.540	0.271	-1.992
Bolivia	-0.539	0.301	-1.790
Angola	-0.538	0.271	-1.990
Mongolia	-0.538	0.271	-1.984
Zambia	-0.525	0.272	-1.930
Djibouti	-0.503	0.303	-1.663
Nicaragua	-0.502	0.278	-1.810
Egypt, Arab Rep.	-0.502	0.275	-1.823
Solomon Islands	-0.481	0.282	-1.707
Senegal	-0.395	0.279	-1.415

[30] We repeated this exercise of excluding one country at a time for our regression discontinuity estimates as well. The results of doing so suggests that the regression discontinuity results are, in general, more robust to country-by-country exclusion than the McCrary tests. For most combinations of bandwidths and polynomial orders, the exclusions of some countries reduces the magnitude and statistical significance of the discontinuity estimate somewhat, but not as dramatically as in the McCrary tests noted in the text. Interestingly, Senegal is never the most "problematic" country to exclude; more often than not, excluding Djibouti has the greatest downward influence of the discontinuity estimate. Excluding the Philippines typically results in substantially larger and more statistically significant discontinuity estimates. Stata code for replicating these models is included in the replication files.

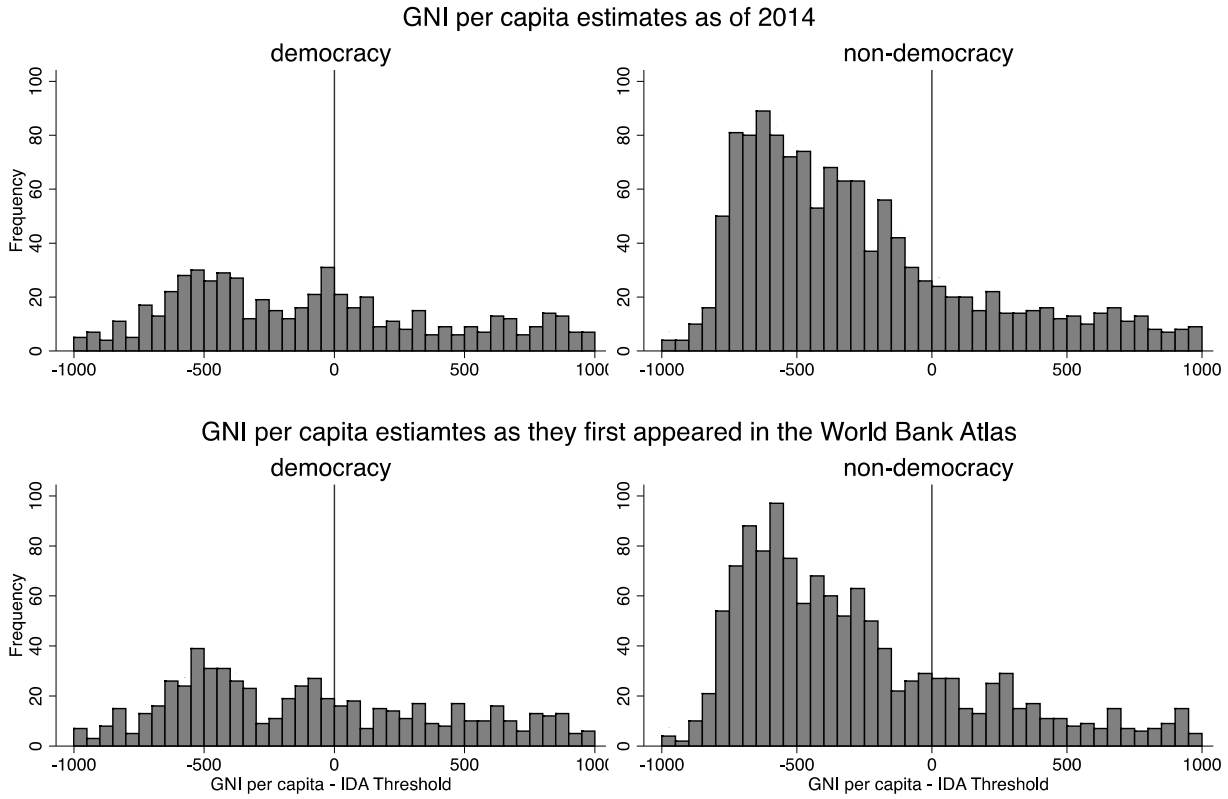
Democracies sort into IDA eligibility

We also considered the possibility that democracies are disproportionately represented just below the IDA eligibility threshold. We think this is a plausible expectation if the perception that democracies are trustworthy allows their aid-seeking estimates to be more frequently represented in the data the World Bank releases. We identify countries as democracies and non-democracies according to their score in the Polity IV dataset (Marshall et al. 2013). We code a country-year as democratic if it has a polity score greater than or equal to 6 (on a scale from –10 to 10), which is the convention in the political science literature.

Figure 8 plots the data separating country-years by whether or not they are democratic. The left side plots show the data from democracies. One thing to notice is that there are not that many democratic observations. In the recently downloaded data there are 568 democratic observations within $1,000 of the IDA threshold and 198 within $300. In the data taken form the World Bank Atlas there are 609 democratic observations within $1,000 of the IDA threshold and 190 within $300. There is little convincing evidence in these plots to suggest that democracies are more able to sort into IDA eligibility. In data taken from the World Bank Atlas, the 19 observations between –$50 and $0 below the IDA threshold are hardly different than the 16 observations between $0 and $50 above it. There is a spike in the density function that occurs below the IDA threshold, but it is only really noticeable when one considers much larger bins of data. For example, the 70 observations within $150 below the IDA threshold are 55% percent more numerous than the 56 observations within $150 just above it. The recently downloaded estimates of GNI per capita show a noticeable spike in the density function just below the IDA threshold, but this spike appears to be an artifact of data revisions that are unlinked to aid distribution and, in any event, is not obviously out of sync with the overall density function. If anything, the histograms on the left side of Figure 8 suggest that democracies tend to cluster near, but not directly below the IDA threshold. This is an interesting pattern, but it does not obviously follows from our expectations.[31] The data from non-democracies show no evidence of clustering below or near the IDA threshold.

[31] Moreover, the McCrary test is not designed to establish whether this sort of clustering pattern is different than one could expect to observe from chance alone.

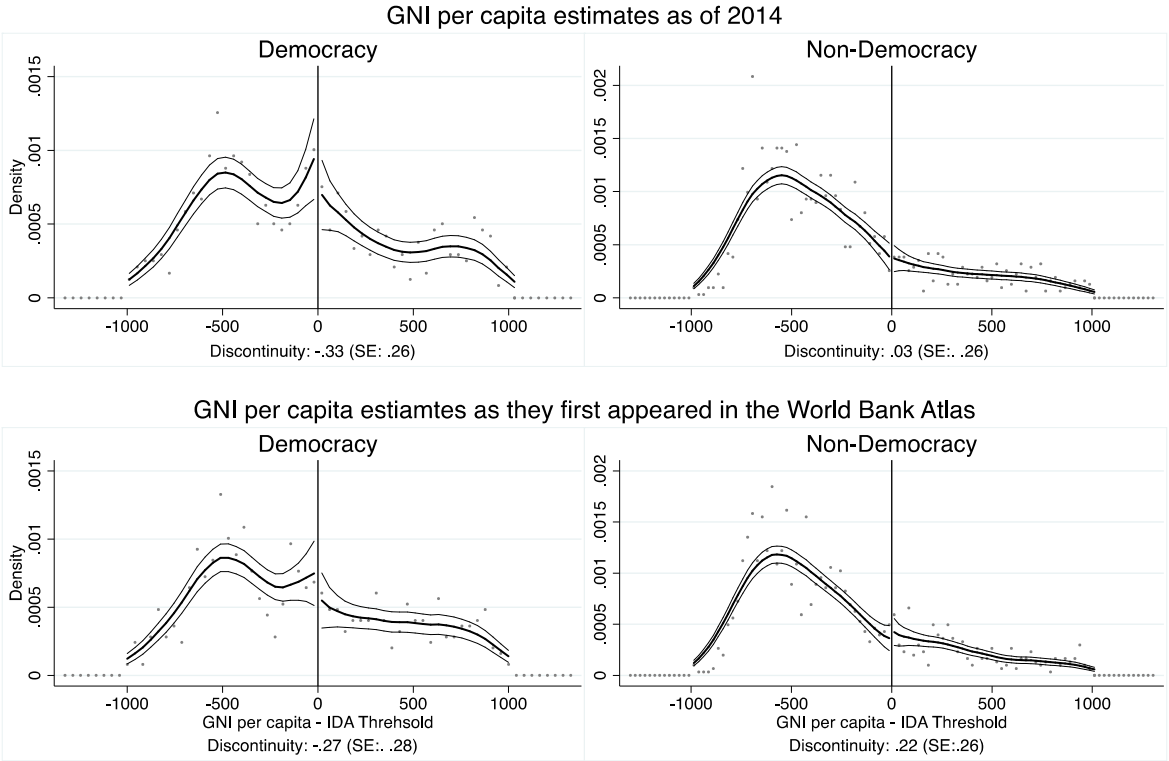
Fig 8: Distribution of GNIs per capita around IDA threshold 1987-2012



see text for defintion of democracy

The results of McCrary tests, shown graphically in Figure 9, corroborate the lack of any obvious clustering below the IDA eligibility threshold. Among democracies, the estimated discontinuity in the density function at the IDA eligibility threshold using data from the World Bank Atlas is in the expected, negative direction (–.27, standard error: .28) but does not approach statistical significance. Our analysis of the more recently downloaded data produces similar results, as do re-estimates using alternative bandwidths.

Fig 9: McCrary Tests of GNI per capita around IDA threshold 1987-2012



McCrary tests with bandwidth of 300; see text for definition of democracy

As a final test we replicated the regression discontinuity models summarized in Table 2 using lagged values of polity scores (the lag structure is entirely innocuous; we get substantively similar results when we use polity scores in their contemporaneous form). The purpose of these tests is to gauge if IDA eligibility in the present appears to "cause" democracy in the past, which would be an indication of sorting behavior. It doesn't. Table 4 summarizes the same regressions as Table 2 (bandwidths of 100, 200, 300 and 400; GNIPCrc entering linearly and as a third order polynomial; estimated with OLS using standard errors clustered by country). As before, the key information in Table 4 is the coefficient associated with the IDA variable, which corresponds to the ρ term. This variable is statistically insignificant in every specification, indicating no clear evidence that democracies are disproportionately able to sort themselves below the IDA threshold.

Table 4: Regression Discontinuity Estimates of IDA Eligibility$_{it}$ on Democracy $_{It-1}$

Model #	1	2	3	4	5	6	7	8
Bandwidth	100	200	300	400	100	200	300	400
	coef./SE	coef./SE	coef./SE	coef./SE	coef./SE	coef./SE	coef./SE	coef./SE
IDA	-1.247	1.114	1.345	1.05	-0.623	-2.357	-1.488	-0.177
	(2.09)	(1.63)	(1.36)	(1.18)	(3.69)	(2.84)	(2.54)	(2.10)
GNIPCrc	-0.026	0	-0.007	-0.004	0.029	-0.106	-0.028	0.002
	(0.03)	(0.01)	(0.01)	(0.01)	(0.19)	(0.10)	(0.06)	(0.04)
IDA*GNIPCrc	-0.001	0.007	0.025**	0.015**	-0.09	0.047	-0.05	-0.042
	(0.03)	(0.01)	(0.01)	(0.01)	(0.23)	(0.13)	(0.07)	(0.04)
GNIPCrc^2					-0.001	0.001	0.000	0.000
					(0.00)	(0.00)	(0.00)	(0.00)
IDA*GNIPCrc^2					-0.001	-0.002	-0.001	0.000
					(0.01)	(0.00)	(0.00)	(0.00)
GNIPCrc^3					0.000	0.000	0.000	0.000
					(0.00)	(0.00)	(0.00)	(0.00)
IDA*GNIPCrc^3					0.00	0.00	0.00	-0.000*
					(0.00)	(0.00)	(0.00)	(0.00)
Constant	2.391	1.401	1.826	1.444	1.853	3.281	2.164	1.608
	(1.64)	(1.46)	(1.42)	(1.34)	(2.35)	(2.10)	(1.90)	(1.63)
r2	0.025	0.004	0.039	0.032	0.03	0.02	0.053	0.05
N	187	340	553	776	187	340	553	776

All models estimated with OLS with standard errors clustered by country. * 0.10 ** 0.05 *** 0.01 **** 0.001

To summarize, we explored two possibilities concerning a "political economy" of data manipulation. The first was that aid-dependent countries would be more likely to try to sort themselves below the IDA eligibility threshold by submitting low GNI estimates to the World Bank. There is (qualified) support for this hypothesis in the data taken from the World Bank Atlas, but less so in the data downloaded from WDI, which is to be expected given the nature of the revision process. We also explored the possibility that democracies would be more likely to have their low estimate pass through the World Bank revision process and we found little evidence to suggest that is the case.

III. Conclusions

Foreign aid distribution relies heavily on evaluations of which countries are poor and which countries are not. While poverty and development (however conceived) exist along a continuum, the official designations relevant to aid distribution are often dichotomous. The IDA threshold is one such example. Countries with GNIs per capita below the IDA threshold are eligible for more concessional loan programs than countries that are not. In practice, GNI per capita is estimated by the World Bank from data that emanates from the countries themselves, which creates an opportunity for national statistical offices to submit estimates that, while not necessarily wrong, are chosen for their potential to maintain or establish a country's IDA eligibility.

This paper asks whether some countries are able to remain IDA eligible by submitting low estimates of their GNI to the World Bank. This possibility is interesting in its right and has important implications for social science research on the topics of aid and development. If some countries appear poorer in the data because it is advantageous for them to appear that way, it can become difficult to accurately interpret empirical correlations.

The evidence in this paper is reasonably suggestive that aid-dependent countries tend to cluster just below the IDA threshold, which could be evidence of sorting behavior brought about through opportunistic estimates of a country's GNI. For the reasons noted above, we do not think the results of this analysis are sufficient to conclude anything definitively, but we do think that the results of our empirical exercise are sufficiently suggestive to indicate that questions about the politics of data production in general and the reliability of development data in social science applications should be explored further.

Bibliography

Abbott, K. W., and D. Snidal 1998. Why States Act Through Formal International Organizations. *Journal of Conflict Resolution* 42 (1): 3–32.

Alt, J., D. Lassen, and J. Wehner. 2012. "Moral Hazard in an Economic Union: Politics, Economics, and Fiscal Gimmicks in Europe." Paper prepared for the annual meeting of the American Political Science Association, Washington, DC, 28–31 August.

Aronow, Peter M. and Carnegie, Allison and Samii, Cyrus. 2014. International Institutions and Political Liberalization: Evidence from the World Bank Loans Program (March 5, 2014). Available at SSRN: http://ssrn.com/abstract=2405297 or http://dx.doi.org/10.2139/ssrn.2405297.

Barnett, M., and M. Finnemore. 2004. *Rules for the World: International Organizations in Global Politics*. Ithaca, NY: Cornell University Press.

Calonico, S., Cattaneo, M. D., & Titiunik, R. *forthcoming*. "Robust nonparametric confidence intervals for regression-discontinuity designs." Econometrica.

Castro, F. de, J. Perez, and M. Rodríguez Vives. 2011. Fiscal Data Revisions in Europe. Working Paper Series no. 1342, May, European Central Bank. Available from http://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1342.pdf. Accessed 12 December 2013.

Caughey, D., & Sekhon, J. S. 2011. "Elections and the regression discontinuity design: Lessons from close us house races, 1942–2008." *Political Analysis* 19(4), 385–408.

Development Initiatives. 2012. ODA. Available at http://www.devinit.org/wp-content/uploads/di-data-guides-oda-09-10-12-E.pdf.

The Economist. 2/9/2013. "Motion of Censure: The Fund Blows the Whistle".

Finnemore, M. 1997. "Redefining Development at the World Bank. In *International Development and the Social Sciences: Essays on the History and Politics of Knowledge*, edited by Frederick Cooper and Randall M. Packard, 203–27. Berkeley: University of California Press.

———. 2013. Constructing Statistics for Global Governance. Working Paper. Unpublished paper in authors' possession.

Heckelman, Jac, Stephen Knack, and Halsey Rogers. 2011. Crossing the threshold: An analysis of IBRD graduation policy. World Bank Policy Research Working Paper no. 5511. Available at http://elibrary.worldbank.org/doi/pdf/10.1596/1813-9450-5531.

Hollyer, J. R., B. P. Rosendorff, and J. R. Vreeland. 2011. "Democracy and Transparency." *Journal of Politics* 73 (4): 1191–1205.

IDA. 2012a. *IDA Borrowing Countries*. Available at http://www.worldbank.org/ida/borrowing-countries.html.

———. 2012b. GNI per Capita Operational Guidelines & Analytical Classifications. Available at siteresources.worldbank.org/DATASTATISTICS/Resources/OGHIST.xls.

———. 2012c. Review of IDA's Graduation Policy. IDA Resource Mobilization Department. Concessional Finance and Global Partnerships. Available at http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2012/10/24/000333038_20121024232535/Rendered/PDF/733630BR0IDA0R0Official0Use0Only090.pdf.

———. 2013a. ABCs of IDA: Thematic and Country Results. Available at http://www.worldbank.org/ida/ida_abc.html.

———. 2013b. IDA's Long Term Financial Capacity and Financial Instruments. IDA Resource Mobilization Department (CFPIR). Available at http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2013/03/08/000356161_20130308142851/Rendered/PDF/758330BR0IDA0S0Official0use0Only090.pdf.

Imbens, Guido, and Thomas Lemieux. 2008. "Regression discontinuity designs." *Journal of Econometrics* 142: 615–35.

Jerven, Morten. 2013ba."For richer, for poorer: GDP revisions and Africa's statistical tragedy." *African Affairs* 112 (446): 138–47.

———. 2013b. *Poor Numbers: How We Are Misled by African Development Statistics and What to Do about It*. Ithaca, NY: Cornell University Press.

Kenny, Charles. 2011. What does it mean to be low income? Global development: Views from the center. Available at http://www.cgdev.org//blog/what-does-it-mean-be-low-income.

Keohane, R. O. 2005. *After hegemony: Cooperation and discord in the world political economy*. Princeton, NJ: Princeton University Press.

Keohane, R. O., and J. S. Nye. 1974. "Transgovernmental Relations and International Organizations." *World Politics* 27 (1): 39–62.

Keohane, Robert O., and Joseph S. Nye. 1974. "Transgovernmental relations and international organizations." *World Politics* 27 (1): 39–62.

Knack, Stephen, Lixin Colin Xu, and Ben Zou. 2013. Income thresholds and aid responses. Paper presented at the 17th annual meeting of the International Society for New Institutional Economics, Florence, 20–22 June.

Koremenos, B., C. Lipson, and D. Snidal. 2001. "The Rational Design of International Institutions." *International Organization* 55 (4): 761–99.

Lederman, D., Loayza, N. V., & Soares, R. R. 2005. "Accountability and corruption: Political institutions matter." *Economics & Politics* 17(1), 1-35.

Lee, David S. 2008. "Randomized experiments from non-random selection in U.S. House elections." *Journal of Econometrics* 142:675–97.

Marshall, Monty G., Keith Jaggers, and Ted Robert Gurr. 2013. Polity IV project. Center for International Development and Conflict Management at the University of Maryland College Park.

McCrary, J. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics* 142 (2): 698–714.

Milesi-Ferretti, G. M. 2004. "Good, Bad or Ugly? On the Effects of Fiscal Rules with Creative Accounting." *Journal of Public Economics* (1): 377–94.

OECD. 2013. Aid to Poor Countries Slips Further as Governments Tighten Budgets. *OECD Newsroom*, 4 March. Available at http://www.oecd.org/newsroom/aidtopoorcountriesslipsfurtherasgovernmentstightenbudgets.htm.

Ravillion, Martin. 2013. Why $12,616? Global Development: Views from the Center blog, 8 July. Available at http://www.cgdev.org/blog/why-12616.

Schenker, N., & Gentleman, J. F. 2001. "On judging the significance of differences by examining the overlap between confidence intervals." *The American Statistician* 55(3), 182-186.

Sumner, A. 2011. The New Bottom Billion: What If Most of the World's Poor Live in Middle-Income Countries? Center of Global Development blog, 18 March. Available at http://www.cgdev.org/publication/new-bottom-billion-what-if-most-worlds-poor-live-middle-income-countries.

Treisman, D. (2007). "What have we learned about the causes of corruption from ten years of cross-national empirical research?" *Annual Review of Political Science* 10: 211-244.

Von Hagen, J., and G. B. Wolff. 2006. "What Do Deficits Tell Us About Debt? Empirical Evidence on Creative Accounting with Fiscal Rules in the EU." *Journal of Banking & Finance* 30 (12): 3259–79.

Wallace, Jeremy. L. 2014. "Juking the Stats? Authoritarian Information Problems in China." *British Journal of Political Science* 1-19.

Ward, M. 2004. *Quantifying the World: UN Ideas and Statistics*. Bloomington: Indiana University Press.

World Bank. 2012. About the World Bank. Available at http://go.worldbank.org/1M3PFQQMD0.

———. 2013a. The World Bank Atlas Method. Available at http://data.worldbank.org/about/country-classifications/world-bank-atlas-method.

———. 2013b. The World Bank's Fund for the Poorest. Available at http://www.worldbank.org/ida/what-is-ida/fund-for-the-poorest.pdf.

Appendix

Appendix Table 1 shows the results of regressions relating to the relative predictive power of two different proxies for aid dependence – the ten-year lag of aid per capita and infant mortality in 1980, which has been used elsewhere in the literature (ex. Burnside and Dollar 2000, Knack 2001). These are OLS regressions with standard errors clustered by country. Our dependent variable is aid per capita in the current period. To better approximate our tests we limit the sample to World Bank members within $300 of the IDA threshold, excluding countries subject to the small islands exemption. GNI per capita is implicitly controlled for as the assignment variable in the McCrary tests (and explicitly so in our regression discontinuity models) and so we control for it explicitly here.

Model one regresses aid per capita on the one-year lag of GNI per capita. The coefficient on GNI per capita is negative and highly statistically significant. The R^2 is .024. Model two adds the measure of infant mortality in 1980. The coefficient is negative, which is contrary to expectations, and statistically insignificant. The R^2 increases slightly, to .047. Model three replaces infant mortality in 1980 with the 10-year lagged measure of aid per capita. The coefficient is positive and statistically significant, and the R^2 increases substantially from .024 in model 1 to .255, suggesting that the 10-year lag of the aid to GNI ratio is a substantively and well as statistically significant predictor of the level of aid dependence among the countries in our sample. When both proxies are included in the same model (model 4), the coefficient estimate on the 10-year lag retains its statistical significance and magnitude, while the estimated coefficient on infant mortality remains statistically significant and continues to carry the wrong sign. Models 5-8 replicate models 1-4, but extend the sample to country-years within $1000 of the IDA eligibility threshold. The results are substantively similar and indicate the 10-year lag as the more appropriate instrument.

Appendix Table 1: Regressions of aid dependency on proxies of aid dependency

DV: Aid per capita								
Sample:	With $300 of IDA eligibility threshold				With $1000 of IDA eligibility threshold			
	1	2	3	4	5	6	7	8
	coef./SE	coef./SE	coef./SE	coef./SE	coef./SE	coef./SE	coef./SE	coef./SE
GNIpc (one year lag)	-0.051***	-0.056**	-0.036**	-0.033**	0.006	-0.01	-0.002	-0.013
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
infant_mortality80		-0.385		-0.226		-0.271		-0.163
		(0.48)		(0.33)		(0.26)		(0.21)
aid per capital (ten year lag)			0.367****	0.393****			0.329****	0.339****
			(0.06)	(0.05)			(0.06)	(0.06)
Constant	117.212****	154.123***	74.399****	88.137**	57.133****	92.716**	39.997****	61.023**
	(18.71)	(55.00)	(15.22)	(33.91)	(6.08)	(36.07)	(5.36)	(28.17)
r2	0.024	0.047	0.255	0.311	0.002	0.018	0.232	0.276
N	492	452	455	430	1704	1625	1615	1563

All data taken from World Development Indicators
All Models estiamted using OLS with custered robust standard errors
* .1, **.05, ***, .01, ***.001