Labeling Laggards and Leaders: International Organizations and the Politics of Defining Development*

Lindsay R. Dolan†

September 14, 2017

The most recent version of this paper and the Supplementary Information are available at www.lindsayrdolan.com/research.

Abstract

Many scholars view international organizations primarily as vehicles through which powerful states distribute resources. However, this view overlooks the ability of bureaucracies in international organizations to affect perceptions of countries. This paper argues that international organizations use classifications to shape how states are treated by international economic elites. Arbitrary changes in a country's classification have major effects on high-stakes decisions such as aid, investment, and credit and democracy ratings. After proposing two mechanisms—cognitive and strategic—by which classifications influence elite behavior, I show with cross-national data from 1987 to 2015 that a country's World Bank income classification correlates with the rewards it receives from actors who are susceptible to cognitive biases or accountable to less informed actors. Drawing on 50 interviews with various stakeholders, I illustrate these patterns in the "graduations" of Nepal, Botswana, and others from different classifications produced by the World Bank and the United Nations. Finally, qualitative and quantitative evidence shows that countries try to exploit these dynamics by attempting to change their classifications. The paper identifies a relatively unexamined power of international organizations in a context where its deployment significantly affects outcomes for developing countries.

*This paper was previously circulated under the title, "What's in a Name? How Bureaucratic Classifications Drive Development Outcomes." For helpful comments and conversations, I thank Chris Blattman, Sarah Bush, Justin Canfil, Allison Carnegie, Michael Clemens, Alicia Cooperman, Alan Gelb, Kolby Hanson, Mary Louis, Tamar Mitts, Scott Morris, Todd Moss, Jack Snyder, and Tara Slough, as well as participants in workshops and panels at Columbia University, the Center for Global Development, IPES 2016, and APSA 2017. I am grateful to the Center for Global Development for the fellowship that enabled this research. Jamie Park provided helpful research assistance. All remaining errors are my own. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship, grant number DGE-16-44869, and by the National Science Foundation Doctoral Dissertation Research Improvement Grant in Political Science #1647493.

†Ph.D. Candidate, Political Science, Columbia University. E-mail: lindsay.dolan@columbia.edu

In 2015, commuters in Washington, DC, walked by a new series of billboard ads demanding that the World Bank "Raise the MIC"—that is, raise the national income threshold separating "low income countries" (LICs) from "middle income countries" (MICs). The Raise the MIC campaign, which has also organized petitions and protests in addition to the ads, argues that this threshold is too low and allows needy countries to be perceived as no longer poor. This perception, they argue, has caused donors to reduce their foreign aid, banks to provide fewer concessional loans, and pharmaceutical companies to charge higher prices for essential vaccines. In short, the movement faults the World Bank for producing the classification system that others use in their operations and decision-making.[1]

The Raise the MIC campaign shines a spotlight on an important power of international organizations (IOs), the ability to classify. The World Bank's income classification is just one of many such systems mapping the landscape of global development. States may belong to the Least Developed Country (LDC) category, the Fragile and Conflict Affected State (FCAS) category, the Highly Indebted Poor Country (HIPC) category, and others. Each of these categories is created and maintained by a prominent international institution that devotes valuable staff and resources toward ensuring that the category remains fit for purpose.

Why are these classifications so powerful? As I will show, these systems often communicate no expert information but nonetheless sway the decisions of even well-informed actors such as donors, investors, and raters. Classifications can influence the behavior of influential economic and political elites through both a cognitive mechanism and a strategic one. Cognitively, classifications act as heuristic devices that simplify abundant and contradictory information and ease the process of decision-making.[2] Strategically, by making use of classifications created by a third-party actor, elites can justify their behavior to an external audience.[3] Consistent with the predictions of this theory, I show that classifications exert the strongest effects on actors who are either susceptible to cognitive biases or are accountable to less-informed audiences. Many susceptible actors are responsible for high-stakes decisions such as aid, investment, and credit and democracy ratings. This means that bureaucrats in international organizations have great ability to influence interactions between important actors in the international economy in ways that materially affect the lives of millions living in poverty.

[1] In their words, "the World Bank income classification scale sends a global message that is distorting the reality and does not accurately reflect the income level of the majority of people in these countries." Letter addressed to Jim Yong Kim, president of the World Bank, co-signed by 520 organizations. `http://raisethemic.org/wp-content/uploads/2016/01/World-Bank-Packet.pdf`

[2] See Tversky and Kahneman (1974); Kahneman (2011).

[3] See Fearon (1994).

Academic scholarship on IOs has been slow to acknowledge that these tools constitute an important source of IO power. Instead, the prevailing literature primarily views IOs as vehicles for distributing international resources, which trivializes the independent importance of IOs by focusing on their powerful state backers.[4] This "rational design" approach dominates the study of IOs, and especially the study of the World Bank and the International Monetary Fund. Numerous works demonstrate "politics at work" in the disbursement of aid and lending to recipients with greater representation on the Boards of these institutions.[5] Widespread prevalence of this view has led prominent theorists to dismiss IOs as mere reflections of an underlying distribution of power.[6]

This paper argues that IOs do more than distribute resources — they shape how countries are treated by other influential economic and political actors. The rational design approach underestimates the power of IOs to create rules, shape dialogue, and cultivate norms in ways that tangibly impact political and economic outcomes.[7] It is especially important to acknowledge these functions when studying global development institutions. Indeed, because of the low interest rate environment and new sources of finance, influential observers of the World Bank argue that its primary role has shifted from lending to facilitating international agreements and providing expertise.[8]

My project improves our understanding of bureaucratic exercises of IO power by focusing on classification. I define classification as the organization of information by creating categories and sorting objects into them. Classification can be ad hoc or systematized according to a set of rules. Michael Barnett and Martha Finnemore have called classification "one of bureaucracy's greatest sources of power," but we have little understanding of its usefulness, its limitations, or its deployment.[9] IO bureaucracies are not the only ones who classify: Scott (1998) similarly argues that states extended their bureaucratic reach by making their populations "legible" to the state through censuses, mapping, and records, all of which serve to classify individuals as "citizens." In this way, classification and governance are intimately related, both at a domestic and an international level.

[4] According to Koremenos et al. (2001), "states use international institutions to further their own goals, and they design institutions accordingly."

[5] For examples of empirical works on distributive politics in development institutions, see Kuziemko and Werker (2006); Dreher et al. (2009); Kaja and Werker (2010); Kersting and Kilby (2016).

[6] See Waltz (1979); Mearsheimer (1994); Carr (1964).

[7] See Barnett and Finnemore (1999, 2004); Autesserre (2014). My argument also comports with Chwieroth (2009), who highlights the autonomy of IMF staff to deliberate over norms surrounding capital controls and apply those norms with discretion.

[8] See Ravallion (2016); Clemens and Kremer (2016).

[9] More fully, the authors assert, "An elementary feature of bureaucracies is that they organize information and knowledge. ...The ability to classify objects, to shift their very definition and identity, is one of bureaucracy's greatest sources of power." Barnett and Finnemore (1999, 711).

A longstanding tradition in political science focuses on the informational role of international institutions. Keohane (1984), for example, argues that reducing information asymmetries between international actors is one of the main reasons that states agree to cooperate in regimes, and is one of the primary ways in which they shape global governance. By providing information, IOs empower other political actors to enforce governmental compliance with international agreements.[10] In addition, collecting and publicizing economic, health, and environmental statistics facilitates states' individual and collective abilities to develop policy in these domains.[11]

But classifications are distinct from information. Classifications organize or distill information, but in so doing, they identify certain features as more important or essential than others, and are therefore information-reducing rather than information-increasing. Classifications tell a story about data rather than present data. Political economists and economic sociologists have previously argued that narratives and ideas play a causal role in motivating economic behavior.[12] Innovations in financial models affect the strategies used by investors, narratives about the Great Depression shaped consumer and investor behavior during the 2008 financial crisis, and economic ideas about embedded liberalism facilitated institutional change in recent decades.[13] In international relations, Wendt (1992) makes a similar argument by asserting that international anarchy — a core assumption used by many realist scholars of IR to explain state behavior — is itself a social construct and motivates states' behavior only insofar as they believe it does. In other words, narratives and ideas about how the world operates can be self-fulfilling and can themselves cause the phenomena they describe. I argue that classifications are an example of this dynamic. Development classifications, in particular, tell a story about what constitutes development and coordinate the behavior of international actors; consequently, graduations from "low-income country" status are often celebrated by national news outlets as markers of success. Most important, I suggest that IO bureaucrats are often the narrators and have the power to influence international politics.

My work joins a growing literature on the importance of labels and indicators in political science. For example, previous studies have found that investors' risk perceptions are affected by a country's club mem-

[10] See Dai (2007); Büthe (2012).

[11] This is especially true in the context of global development. Clemens and Kremer (2016) note, "A multilateral institution can reduce asymmetries of information, increasing the chance of an efficient bargain. If the donors have information that a change in policy in a direction preferred by the donors might be very low cost or even beneficial from the point of view of the recipient country authorities, they would want to find a way to credibly transmit this information."

[12] See Rodrik (2013).

[13] See MacKenzie (2006); Shiller (2017); Blyth (2002).

bership and by the other peer countries with which it is grouped.[14] Classifications are also part of a broader array of indicators, rankings, and ratings, each purporting to evaluate a country's level of corruption, human rights record, level of democracy, debt sustainability, and other features of interest.[15] Frequently these indicators have been intentionally developed or deployed to exert social pressure on rated countries to comply with international norms.[16] In turn, graded countries, perceiving the benefits of desirable scores, labels, or groupings, often adopt policies targeted at changing their scores or petition the ratings organizations to amend them.[17] Yet despite scholarship's increasing awareness of the politicization of these metrics, we lack an understanding of why they exhibit these dynamics. Why would elite actors in the international economy use simplified classifications produced by an international organization instead of analyzing the data for their specific purposes? My theory draws on strategic and psychological theories to argue that actors who are susceptible to (or accountable to audiences who are susceptible to) cognitive biases will defer to classifications produced by IOs. This argument helps us to understand where classifications are most likely to give rise to the political dynamics identified by myself and others.

In what follows, I offer a new theory to account for the power of classification and derive testable hypotheses. Next, I describe the development classifications produced by the World Bank and the United Nations (UN), describing the respective research opportunities they present. Exploiting the exogeneity of thresholds separating categories and applying a difference-in-differences strategy, I show with cross-national data from 1987 to 2015 that a country's World Bank income classification correlates with its treatment by actors who are susceptible to (or accountable to audiences who are susceptible to) cognitive biases. Next, I draw on evidence from interviews with a variety of stakeholders in the World Bank and UN graduation processes to support my claims regarding the effects of classifications. Classified constituencies and countries are not only aware of classification effects but they also appear to respond strategically. I couple interview evidence with a novel source of data on revisions to national income figures to support this assertion, which logically follows from my theory, and I conclude by discussing implications for future research.

[14] See Gray (2013); Brooks et al. (2015).

[15] Indicators are defined by Davis et al. (2012) based on four features: "1) the significance of the name of the indicator and the associated assertion of its power to define and represent a phenomenon such as "the rule of law"; (2) the ordinal structure enabling comparison and ranking and exerting pressure for "improvement" as measured by the indicator; (3) the simplification of complex social phenomena; and (4) the potential to be used for evaluative purposes" (7). Classifications are a type of indicator that is particularly extreme on the third feature: simplification. While a classification could be made on the basis of multiple criteria, the result reduces an object's identity to its inclusion inside or outside a group, usually on the basis of thresholds.

[16] See Kelley and Simmons (2015); Kelley (2016); Morse (2017). Buntaine et al. (2017) argue that these strategies can perversely cause countries to shift resources toward achieving simple targets.

[17] See, for example, Cooley and Snyder (2015); Kelley and Simmons (2015); Carnegie and Samii (2017).

1 Theory

Why would classifications produced by an international organization influence the behavior of other actors in the international economy? The answer is not obvious; while scholars of international politics have highlighted the information-providing role of international organizations for decades, classifications reduce rather than add to the information known about a country. Take, for example, a focus of this paper, the World Bank income classification system under fire from the "Raise the MIC" campaign. In 2013, Bangladesh was classified as a LIC because its national income per capita was $900, below the threshold of $1,045 that separates LICs and (L)MICs. (Malawi, at $270, was also included in this category.) In 2014, however, its income was $1,080, above the threshold and qualifying it as a (L)MIC, a category which also included Armenia at $3,810. While thresholds are updated in real terms over time to account for inflation, no information other than a country's national income per capita is used to determine its income classification. The very same national income data used to classify economies is widely available through the World Development Indicators, and, as I explain in greater detail below, the thresholds separating the categories were arbitrarily selected. Consequently, an observer can arguably learn much more by looking at Bangladesh's raw income per capita than she can by looking at its classification, by which metric Bangladesh graduates from Malawi's category to Armenia's in a single year.

A more sophisticated variant of the informational argument is that even if classifications can be reproduced using publicly available data, the classification system itself communicates expert information about how a particular concept should be measured. After all, regulatory agencies regularly classify goods and behaviors to enable consumers to make informed decisions without being technical experts about their health, safety, or environmental consequences. These classifications summarize technical information that may be publicly available but difficult for a non-expert consumer to synthesize or comprehend. Indeed, technical expertise is a leading explanation for the empowerment of private regulatory authorities; governments are willing to stomach policy uncertainty in exchange for the value of costly expertise.[18] But in other domains the categories being defined are contested (e.g. democracy or development), and even if defining them also requires technical expertise, it is impossible to create these categorizations devoid of political influences.[19] Moreover, an explanation centered on technical expertise cannot explain why classifications shape the behavior of actors *who are also expert in the topic being measured*. As I will show, development classi-

[18] See Mattli and Büthe (2005); Mattli and Woods (2009).

[19] See Büthe (2012); Bush (2017) for examples, with special attention on the Freedom House Freedom in the World index.

fications do affect the behavior of actors with deep contextual and economic knowledge about the countries being measured, who would be both able to and better served to identify the most relevant features of development for their purposes. With this in mind, we have to look beyond the rationalist argument presented by the literature on regulation that classifications simply communicate expert information.

In this section, I offer a cognitive-strategic theory to account for the disproportionate power these classifications acquire. In brief, I lay out a cognitive mechanism by which classifications act as heuristic devices to simplify decision-making and a strategic mechanism by which actors use classifications produced by a third-party actor to demonstrate the impartiality of their decisions. Separating these mechanisms improves our ability to explain variation in how classifications are used and builds on existing scholarship seeking to understand indicators. Büthe (2012)'s conceptual model shows that the "users" of an indicator encompass a broader audience than just the "rule-demanders," the group that demands an indicator from a private authority. By explaining *what kinds of* indicators are likely to affect *which* actors, my theory speaks to the composition of an indicator's group of users, which in turn has implications for the kinds of indicators that will be supplied.

1.1 Cognitive Mechanism

Extensive research in behavioral economics has documented that many decision-makers are susceptible to cognitive biases that sway their decisions in irrational ways. One important finding from this literature is that human brains lack the computational ability to process abundant information and behave rationally, and they turn to shortcuts or "heuristics" to assist in their decision-making.[20] Like a computer, human brains are constantly applying algorithms to distinguish relevant from irrelevant information, but this is a cognitively demanding task. Moreover, we experience deep discomfort (termed "cognitive dissonance") when we encounter conflicting information and look for shortcuts to avoid grappling with every data point. Categorizations offer one such shortcut: A decision-maker can make inferences about an unfamiliar case based on her familiarity with other cases in the same category. Studies show that categories cause individuals to think units are more similar within group and more different between group.[21] Given the amount of contradictory and poor-quality data in the development landscape, these effects may be especially pronounced, as

[20] See Simon (1955); Conlisk (1996); Kahneman (2011).

[21] Tajfel and Wilkes (1963) conduct an experiment in which individuals are asked to estimate the lengths of four short lines labeled "A" and four long lines labeled "B." The group who saw the "A" and "B" labels estimated the within-group lengths to be more similar than the group who did not see the labels. See also Taylor (1981); Wilder (1986) for social psychological studies of categorizations and intergroup bias.

classifications could be an attractive means of simplifying a messy picture of countries on the cusp of development.[22] In other words, this explanation argues that classifications are so powerful because they distill information and, in so doing, reduce the cognitive dissonance that is so perturbing to humans. Other political scientists have investigated the relevance of heuristics to international economic behavior, with evidence showing that peer country groupings influence sovereign lending practices.[23]

Cognitive psychology has also shown the existence of a "halo effect," which could amplify the heuristic effects of classifications. Social psychologists have amassed evidence that an individual's opinions about an actor's attribute A are influenced by information about that actor on unrelated attribute B. For example, evidence shows that people who are thought to be more attractive are often perceived as being more competent or intelligent.[24] If actors in the international community are also susceptible to this bias, classifications could affect international perceptions on more dimensions than just that which they seek to measure. For instance, countries that graduate in a development classification might also more easily persuade observers that they are democratic or respect human rights.

An important requirement of a good heuristic device is that the observer be familiar with the classification system. Taylor (1981) notes that the propensity to stereotype based on group affiliation depends on "the extent to which a perceiver has a well-developed concept for that attribute. ... A Korean is simply [Asian], if one is unable to distinguish Koreans from [Asians]" (86). As a result, we would expect simple, intuitive, and salient classification systems to exert the strongest effects.

Evidence from economics suggests these cognitive biases could affect the behavior of even elite, highly informed actors. Previous works have pointed to the role of euphoria and positive affect in how investors evaluate risk and returns, and these cognitive factors have been partially blamed for financial crises.[25] According to Steinbruner (1974), decision-makers actually work to avoid exposing themselves to all relevant information in order to make it possible to navigate complex problems, group decision-making dynamics, and conditions of uncertainty.[26] At the same time, it is possible that elite actors have more informed priors that are less sensitive to the heuristics I describe. Moreover, decisions that are filtered through more institutionalized processes with more veto players could be subject to greater, rather than less, rational scrutiny.[27]

[22] Jerven (2013) documents the extent of the unreliability of data in the developing world.

[23] See, for example, Gray (2013); Gray and Hicks (2014); Brooks et al. (2015).

[24] For a review, see Greenwald and Banaji (1995).

[25] See Aspara and Tikkanen (2010); Kindleberger (2005) in addition to Gray (2013); Brooks et al. (2015).

[26] These organizational conditions that exacerbate cognitive biases have been especially scrutinized in security studies. See Jervis (2006) for an example.

[27] Indeed, one of the most prevalent critiques of psychological theories of international relations is that these theories rarely offer

Psychological research shows that asking individuals to justify their treatment of cases reduces the effects of implicit bias.[28] Consequently, whether elite actors are more or less likely than the mass public to be susceptible to cognitive biases remains an empirical question, but for the purposes of my study, neither is likely to be immune to them.

1.2 Strategic Mechanism

Even actors who are not susceptible to cognitive biases could nonetheless rely on classifications for strategic reasons. When actors are responsible for making contentious distributive decisions that will produce backlash from those who are adversely affected, using a classification developed by a third party allows them to evade accusations of being partisan or political in their decisions. This logic has previously been used to explain why governments delegate regulatory powers to independent agencies. Weaver (1986) notes that one way legislators often avoid blame is to "cede discretion to the president or an independent agency for making politically costly decisions" (375).[29] Similarly, if political actors incorporate classifications formally or informally into their allocative decisions, they can credibly demonstrate their impartiality to any supporters who demand it. Since an actor's own classification system will never be perceived as credibly impartial, the actor prefers this approach to spending valuable time and resources on a proprietary system.[30]

This argument is related to but distinct from an existing explanation for the influence of "private authority" in world politics. Green (2013) defines "private authority" as "situations in which non-state actors make rules or set standards that others in world politics adopt" (42). One explanation for the influence of private authority centers on the independence and neutrality of that authority. For this reason, Green's definition specifically excludes international organizations "since they comprise state representatives who are responsible for taking or delegating decisions" and therefore are not neutral (43). While this is true in an objective sense, multilateral organizations are more widely trusted by the public than governments, especially to carry out development.[31] Indeed, the perception of the classifier's neutrality (particularly as compared to

an argument for how individual-level psychological traits explain behavior at the state level of analysis. For example, in their critique of the research on status inconsistency, Dafoe et al. (2014) argue, "IR studies that used this as their theoretical foundation did not typically contain any theory that explained why we should expect individual-level results linking inconsistency to "violent" or "dysfunctional" behavior to translate directly to world politics." (388)

[28] See Taylor (1981).

[29] See also Fiorina (1982); Mattli and Büthe (2005); Flinders and Buller (2006); Landwehr and Böhm (2011).

[30] These resources are not insignificant: The World Bank data group has received much criticism regarding its classification system. See Fantom and Serajuddin (2016).

[31] See Milner (2006).

the perception of the classification user's neutrality) is sometimes more important than the neutrality itself.[32]

This mechanism is most likely to affect actors who are accountable for distributing scarce political resources. In the context of development, donor agencies are likely to be especially sensitive to this mechanism, as they must justify their allocation decisions to the legislatures (and, by extension, mass publics) that determine their budgets. Donor agencies often try to emphasize their impartiality in their concern for development. Milner (2006), for example, argues that bilateral donors finance multilateral agencies, who are more trusted by the mass public to carry out development, in an effort to reassure their funders that they are providing needs-based aid. This aim was also behind the founding of the Millennium Challenge Corporation (MCC), which was intended to be an apolitical aid instrument for the U.S. government. Consistent with my argument, MCC has developed an eligibility and allocation rubric that relies heavily on democracy measures produced by Freedom House, an external agency.[33] This allows MCC to avoid any appearance of making partisan decisions in its development assistance. Particularly in the present political environment in which donor publics often feel aid could be better spent at home, even bilateral donors with expressly political objectives may need to illustrate the impoverishment of their recipients to maintain support for funding aid.[34] In this way, IOs can influence the behavior of states indirectly by shaping the standards the mass public and, by extension, legislators use to evaluate their governments.[35] Additionally, this logic potentially explains why researchers have observed frequently counterproductive "herding" behavior among donors, who duplicate rather than balance the assistance provided by other donors.[36]

The strategic mechanism is less likely to operate when actors are not accountable to external audiences. In other words, outside the context of a principal-agent relationship, actors have no need to consider the appearance or signal that their actions are sending. As such, I do not expect to observe the strategic mechanism at work in the perceptions expressed either by members of the mass public or by other ratings and rankings organizations (whose credibility could even be undermined by reference to another organization's classifications). Furthermore, I also do not expect to observe this mechanism at work in the case of investors. Since investors are accountable only for delivering a return on their investments and not for the fairness

[32] Sinclair (2005) and Bush (2017) argue that credit ratings agencies and Freedom House respectively have been awarded positions of private authority but are not the neutral, technical agencies they purport to be.

[33] See Bush (2017).

[34] For evidence connecting financial crises to reductions in aid, see Heinrich et al. (2016). Research on public opinion in donor countries finds that humanitarian, altruistic, and sometimes paternalistic concerns can motivate individuals to support foreign aid or immigration despite countervailing nationalist incentives (Paxton and Knack 2012; Newman et al. 2013; Bechtel et al. 2014; Baker 2015). This research suggests that emphasizing the relative need of aid beneficiaries will improve support for foreign aid.

[35] Similar arguments have been made about IOs in the context of military interventions. See Chapman (2009).

[36] See Knack et al. (2014); Frot and Santiso (2011).

implications of their decisions, any classification effects they exhibit cannot be explained by the strategic mechanism.

1.3 Testable Hypotheses

I have argued that a country's classification will affect the way international observers perceive its level of development and will therefore affect how it is treated in the global economy. My theory suggests that these "classification effects" can be explained by two mechanisms, cognitive and strategic, and that different observers are likely to be influenced by different mechanisms. To test my theory, I select a variety of international observers who are likely to exhibit variation in their sensitivity to these mechanisms. Drawing on the above mechanisms and knowledge about each type of observer, I then hypothesize *which* observers are likely to demonstrate a classification effect in their behavior. Strictly speaking, my theory is agnostic about the direction of the classification effect, but I apply existing literature and use an expert survey to add directions to my hypotheses.

All four of the observers I study — donors, investors, credit raters, and democracy raters — have profound effects on the economic and political trajectories of developing countries. I select donors and investors because they are two of the three largest sources of external finance for developing economies.[37] For decades, donors have played a crucial role in correcting market failures by supplying finance to developing countries. International targets to increase official development assistance (ODA) to .7% of national income suggest how greatly this aid is valued.[38] Increasingly, though, the development community recognizes the importance to developing countries of foreign direct investment (FDI), that is, the ownership of enterprises in developing economies by foreign firms. While academic research finds mixed results on the effectiveness of FDI as a development strategy, it is widely prioritized by international organizations and countries.[39] For instance, the 2002 Monterrey Consensus states, "Private international capital flows, particularly foreign direct investment, along with international financial stability, are vital complements to national and international development efforts. Foreign direct investment contributes toward financing sustained economic growth over the long term."[40] Private investors are thought to be so important to development that donor governments dedicate development finance, such as the U.S.'s Overseas Private Investment Corpo-

[37] See UNCTAD (2017). The third source of external finance is from remittances, but since I do not expect migrant workers abroad to be affected by classifications, I do not include this outcome in my study.

[38] `http://www.oecd.org/dac/stats/the07odagnitarget-ahistory.htm`

[39] See, for example, Kosack and Tobin (2006).

[40] `http://www.un.org/esa/ffd/monterrey/MonterreyConsensus.pdf`, paragraph 20.

ration, toward facilitating their investments in poor countries.[41] I also select two kinds of raters — credit raters and democracy raters — because their assessments are often the gateway through which developing countries can access both the forms of external finance described above as well as international bond and loan markets. Ratings of a country's creditworthiness specifically determine the rates countries will receive on international markets, but they also demonstrate political risk to other potential economic partners.[42] Democracy ratings, and especially those produced by Freedom House, are widely cited ingredients in aid and investment decisions.[43] To survive in a global economy, developing countries must concern themselves with the way they are perceived by each of these actors.

However, important differences among these four actors mean that they are not likely to be uniform in their sensitivity to either of my hypothesized mechanisms. I proceed by discussing each of these actors in turn: Given findings in the literature about the decision-making processes of each of these actors, I discuss which mechanisms are likely to transmit any classification effect and what the direction of that classification effect is likely to be.

I expect donors to exhibit a classification effect, and this effect could be transmitted through both a strategic and a cognitive mechanism. The rationale for the presence of the strategic mechanism is described above, and I expect this to account for the majority of any effect. Of all the observers involved in my study, donors are the most expert on the topic being described by a classification. This expertise could serve to reduce their cognitive bias. However, for reasons described above, donors could nonetheless exhibit such a bias if they are undertaking complex and time-restricted decisions. New experimental work seeks to test this empirically on elite samples.[44]

I expect that higher classifications will result in lower levels of aid, as a country is perceived as less needful of assistance. Lumsdaine (1993) argues that the majority of foreign aid serves the humanitarian interest of a donor public in improving conditions of poverty. While subsequent empirical research reveals that donors' strategic interests and concerns about governance in recipient countries also condition allocation behavior, ceteris paribus, donor countries will be more likely to give aid to needier recipients.[45] Because aid is a scarce resource, I do not expect donors to contribute more to a country perceived as more developed.

[41] See `https://www.foreign.senate.gov/imo/media/doc/070716_Moss_Testimony.pdf` for a description of the importance of OPIC to U.S. development goals.

[42] See Ratha et al. (2011).

[43] See Blanton and Blanton (2007); Büthe and Milner (2008); Bush (2017).

[44] See Napier et al. (2016); Dolan (2017).

[45] See Alesina and Dollar (2000); Neumayer (2005).

Hypothesis 1. Higher classifications will decrease the amount of aid a country receives. This effect will be transmitted through both cognitive and strategic mechanisms.

Table 1: Predicted mechanisms for classification effect (by observer)

	Mechanism	
	Cognitive	Strategic
Donors	✓	✓
Investors	✓	∅
Other raters	✓	∅

Given the available research, it is possible that investors would exhibit a classification effect, depending on their susceptibility to the cognitive mechanism. Unlike donors, multi-national enterprises have no need to strategically justify their use of resources so long as they return a profit on investment. Therefore, if a classification effect existed with donors, it could be attributable only to a cognitive mechanism. This is a relatively new research paradigm in international business, but early studies suggest that canonical economic models fail to explain the decisions of managers at internationalizing firms.[46] Of particular interest is the possibility that conceptions of "foreignness" and "distance" influence decisions more than can be explained objectively.[47] While not conclusive, these early works suggest that classification effects too could be transmitted cognitively.

If classifications do influence investment decisions, then I expect the direction to be positive. Investors will likely view higher levels of development as signs of stability and will be more likely to invest in countries that are perceived as more developed. An extensive literature on the determinants of FDI suggests that investors are attracted by high levels of human capital, stable political environments, and domestic institutions that protect property rights.[48] Although these factors are conceptually distinct from development, they are frequently correlated with a country's national income level. An investor susceptible to cognitive biases may be likely to attribute these characteristics to a country perceived as having a higher level of development.

Hypothesis 2. Higher classifications will increase the amount of FDI a country receives. This effect will be transmitted through only the cognitive mechanism.

Similarly, I expect raters of creditworthiness and democracy to be influenced by classifications via a cognitive mechanism. These raters should not be vulnerable to the strategic mechanism because their busi-

[46] See Aharoni (2010); Maitland and Sammartino (2015).
[47] See Williams and Grégoire (2015).
[48] See, for example, Noorbakhsh et al. (2001); Jensen (2006); Büthe and Milner (2008).

ness model succeeds on the basis of the originality of their ratings; to defer to a different classification or indicator would undermine their credibility, rather than improve it as in the case of donors. In principle, rigorous coding procedures should serve to mitigate cognitive biases. Because the Freedom House ratings are vetted more extensively than the surveys used by the *Institutional Investor*, I expect to detect a larger cognitive effect in the particular measure of creditworthiness I use than in the particular measure of democracy. However, it is not inconceivable that cognitive biases may still surface in the democracy ratings. To my knowledge, no research has investigated the extent of cognitive bias in ratings organizations themselves.

In terms of direction, I expect a classification effect to bias ratings upwards rather than downwards. The well-documented halo effect suggests that individual raters will likely reward countries perceived as more developed with higher scores on other dimensions. In no circumstance do I expect these scores to be inversely correlated.

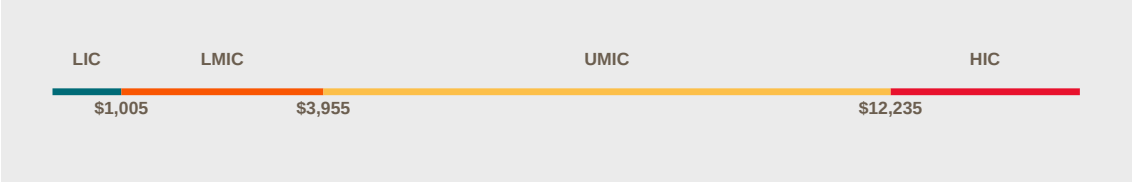
Hypothesis 3. Higher classifications will improve a country's ratings on other dimensions, such as creditworthiness and democracy scores. This effect will be transmitted through only the cognitive mechanism.

Because the directions associated with each of these hypotheses are based on assumptions that follow from existing literature, rather than deriving directly from my proposed theory, I also surveyed a group of experts about their beliefs regarding the existence of and directions of classification effects. Overall, respondents expected classifications to have effects in directions that are consistent with the extant literature. (Details and results are available in Section S1 of the Supplementary Information.)

While hypotheses 1 through 3 examine the effects of development classifications on specific actors, they also, if confirmed, suggest which actors more broadly are likely to use classifications in their analysis and decision-making. If support is found, it sheds light on the composition of the "users" in Büthe (2012)'s conceptual model of indicators.

The theory also suggests important implications for the behaviors of classified governments and domestic groups in classified countries. If classifications are as powerful as my theory suggests, then governments of classified economies may try to change their classifications as a means of achieving their political and economic goals, to the extent they are able. Also, if domestic groups in classified countries feel the effects of these classifications in their activities, then they too may try to influence their country's classification, potentially in a direction opposite to their government's preference. To fully explain the behavior of these governments and groups would require a complete theory of their preferences and action sets, which I reserve for future research. In this paper, I simply look for evidence of any strategic behavior at all by testing

Figure 1: The World Bank analytical income classification system



Note: Figure indicates the thresholds used to classify countries as "low income countries," "lower middle income countries," "upper middle income countries," and "high income countries." Thresholds are revised each year only to account for inflation and are released annually on July 1, along with the revised classifications assigned to each country. This particular set of thresholds was published on July 1, 2017. *Source:* World Bank.

the following hypothesis:

Hypothesis 4. Dissatisfied classified governments and domestic groups in classified countries will try to influence their country's classification.

2 Classification Systems in Global Development

To illustrate the power of international organizations to affect outcomes through their bureaucracies, this paper focuses on three important systems used to classify developing countries: the World Bank's analytical and operational income classification systems and the UN's Least Developed Country category. In this section I briefly describe these classification systems (with additional detail included in Section S2 of the Supplementary Information) and their role in my analysis.

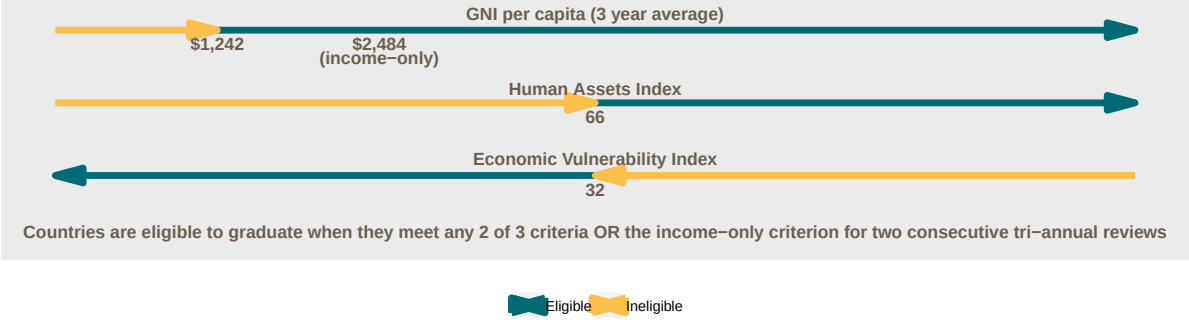
The World Bank's income classification system separates countries into "low income countries," "lower-middle income countries," "upper-middle income countries," and "high income countries" on the basis of their gross national income (GNI) per capita. A series of GNI per capita thresholds separate the categories, and no other indicators are used to determine a country's classification. Figure 1 illustrates that a country's classification is an exact function of its GNI per capita. Each year on July 1, the World Bank Development Economics Group updates the thresholds in real terms to account for inflation and releases the set of country classifications made on this basis. The current system has been in place since 1989, when it was devised for exclusively analytical purposes to facilitate World Bank research. The World Bank has never used these classifications for operational (lending) purposes, and they were not designed for this purpose. The income classifications have become essential to the vocabulary used to discuss development, and they have attracted a fair amount of controversy.

For its operational purposes, the World Bank makes use of a second, less widely-known classification system, which helps it to determine what kinds of assistance countries can qualify to receive from the World Bank. Various arms of the World Bank offer countries different kinds of assistance, ranging from the grants and concessional loans countries can receive through the International Development Association (IDA) to the non-concessional loans they can receive from the International Bank for Reconstruction and Development (IBRD).[49] Since the resource pool of IDA is limited, the World Bank defines eligibility for this category of assistance, and it uses a similar but not identical approach to the analytical income classification system. As in the analytical system, countries are classified on the basis of their GNI data (although according to different thresholds), but assessments of their creditworthiness and other factors are also taken into account when determining whether countries will have access to IDA, IBRD, or both lending arms. After these case-specific evaluations are made, countries are categorized as "IDA-only" countries, which require the most development assistance, "Blend" countries, which receive IDA financing but are now also taking on non-concessional loans from IBRD, and "IBRD" countries, which do not require any concessional assistance. However, unlike the analytical income classification system, these categories are not widely known beyond the aid and development policy community.

The other classification with widespread recognition is the UN's "Least Developed Country" (LDC) category. This classification is made on the basis of three primary indicators: GNI per capita, the Human Assets Index, and the Economic Vulnerability Index. The Human Assets Index, measuring human capital like education and health, and the Economic Vulnerability Index, measuring a country's exposure to macroeconomic or environmental shocks, both contain multiple sub-indices. The LDC category, maintained by the UN's Committee for Development Policy, uses a set of thresholds on each of the three dimensions to determine eligibility for inclusion in and graduation from the category. To be eligible to graduate from the LDC category and become a "developing country," countries must either achieve graduation thresholds on at least two of the three indicators or they can meet double the income criterion and be eligible on the basis of income alone (see Figure 2). After being found eligible in two consecutive triennial reviews and on the basis of the recommendation of UN committees in consultation with the governments in question, countries

[49] Founded to correct market failures, the World Bank offers subsidized loans to developing countries through the International Bank for Reconstruction and Development (IBRD) and the International Development Association (IDA). Intended to help rebuild Europe after World War II, the IBRD provides countries with flexible loans at favorable rates, but they are non-concessional, meaning that the IBRD is financially self-sustaining and does not rely on donors to subsidize borrowing by its members. In contrast, IDA, which is financed through replenishments from donor countries, serves the World Bank's poorest clients by providing concessional finance and grants to countries lacking creditworthiness.

Figure 2: Thresholds for graduating from the UN Least Developed Country category



Note: Figure indicates the three criteria, and the thresholds for each, used to determine a country's eligibility for graduation from the UN's LDC category. For the income criterion, the UN uses the same GNI per capita figures used by the World Bank but calculates a 3-year average. The Human Assets Index (higher is better) is a composite index of: the percentage of population undernourished, the under-five mortality rate, the gross secondary enrollment ratio, and the adult literacy rate. The Economic Vulnerability Index (higher is worse) is a composite index of: population, remoteness, merchandise export concentration, share of agriculture and forestry and fisheries in GDP, share of population in low elevated coastal zones, victims of natural disasters, instability of agricultural production, and instability of exports of goods and services. A separate set of criteria exists to assess a country's inclusion in the criteria. Information on the inclusion thresholds appears in Figure S2 in the Supplementary Information. *Source:* UN Committee on Development Policy.

graduate three years later. A different set of thresholds on the same three criteria determines countries' inclusion in the category, a process that is depicted by Figure S2 of the Supplementary Information.

Several features make these development classifications attractive objects of study. While there exist other important classifications in the development landscape, these that are produced by the World Bank and the UN are arguably the most influential and substantively important. Moreover, useful variation between them introduces promising research opportunities. A particularly helpful source of variation is the process by which countries change categories within a classification system. Because the World Bank's analytical income classification system is predicated on a relatively volatile variable — national income — there is a great deal of movement across threshold boundaries in this system.[50] This creates a unique empirical opportunity to statistically quantify the effects of these classifications in a large data set. In contrast, only seven countries have ever graduated from the LDC category, and three of them were in 2015. This limits the ability to conduct quantitative analysis because there are insufficient opportunities to observe countries' experiences under different classifications. However, the deliberation involved in these graduations offers considerable qualitative data, as international organizations and country experts carefully assess the expected effects of each of these graduations. Consequently, studying a variety of classification systems permits a

[50]Several, though fewer, countries have "graduated" from IDA-only to "Blend" and subsequently IBRD status, and some have "reverse graduated."

multi-method approach that applies both quantitative and qualitative evidence to document the significance of classifications. Finally, a focus on classifications is particularly relevant in the current economic context. Global changes in the demography of poverty have caused many high-profile "graduations" of countries who continue to experience high levels of poverty or inequality.[51] This has cast a spotlight on these classification systems, causing significant controversy surrounding whether they continue to be fit for purpose. This increases the policy relevance of my research (it is important to understand what political consequences these transitions will entail) as well as the evidence available to test my hypotheses (many actors are engaged in a vocal debate about whether and how these systems should be reformed).

3 Quantifying Classification Effects

The aim of this study is to illustrate systematic relationships between a country's classification in a given year and the way it is perceived and treated in the global economy. To do so quantitatively requires a large data set with sufficient variation in the classifications countries receive cross-nationally and over time. The World Bank analytical income classification system exhibits this necessary variation and is ideal for this quantitative analysis. However, a naive regression of outcomes on World Bank income classifications would not adequately show this relationship, since international observers could be influenced directly by the factors that determine a country's classification and not at all by the classification itself. My estimation strategy addresses this obstacle by controlling for the determinant of the classification (GNI per capita) and estimating the effect of being above a threshold determining eligibility for a classification. Subsequent analysis addresses endogeneity in this relationship. Before outlining these models, I begin by introducing my data.

3.1 Data

To test my hypotheses, I combine data on a variety of economic outcomes with the classifications countries received from the World Bank and the indicators used to determine those classifications. Since the World Bank began its current classification system in 1989, my sample includes all country-years that ever received classifications during the period 1987 to 2015, excluding countries that have been continually classified as

[51] See Moss and Leo (2011); Morris and Gleave (2015).

HICs since before 1989.[52]

The main explanatory variables are historical data on gross national income (GNI) per capita and the historical thresholds that were used to classify countries in each year. I obtained the original GNI data that were used to classify countries at their time of classification from the World Bank Development Economics Data Group.[53] It is important to note that this figure can differ significantly from the estimate of GNI that can be obtained by downloading the most current World Development Indicators online. This is because income estimates are often revised over time on the basis of updated economic assumptions, new population data, and other causes. I use the historical income data and the historical thresholds, which is the best way of approximating how a country was perceived at the time of its classification, regardless of how it would be classified ex post on the basis of contemporary data. The historical thresholds are available online from the World Bank.[54]

Dependent variables include aid, foreign direct investment, creditworthiness, and democracy ratings. In keeping with the literature, I measure *Aid* as net disbursements of official development assistance (ODA) as reported by the OECD, and I impute zeroes for missing values.[55] *FDI* refers to logged net annual inward FDI flows from UNCTAD's Handbook of Statistics.[56] The measure of *Creditworthiness* comes from the biannual country risk ratings published in the investment journal *Institutional Investor*. These ratings range from 0 to 100 and are the most widely available continuous indicator of creditworthiness.[57] I obtained the IIR ratings by scanning and transcribing the tables published in March and September issues of the U.S. edition, 1987-2012. Since income classifications are released on July 1, I operationalize this variable by

[52]Because the present categories were defined in FY89, the first observation for most countries therefore contains the classification awarded in FY89 (July 1, 1988-June 30, 1989), which was determined in spring 1988 on the basis of data from calendar year 1987. Unless otherwise indicated, years are reported in terms of data years.

[53]The World Bank provided the GNI data used for the analytical classifications from 1999-2015 (i.e., used to classify economies for FY01 onwards) and the GNI data used for the operational classifications from 1973-2007. In this paper, I use the analytical figure for all years after 1999 and the operational figure for years before. Where they overlap, the numbers are very close to one another.

[54]`http://databank.worldbank.org/data/download/site-content/OGHIST.xls`

[55]See Knack et al. (2014) for precedent.

[56]While some scholars prefer to use FDI flows as a percentage of GDP (see Büthe and Milner (2008)), I prefer the use of levels to avoid detecting a mechanical relationship as the result of changes in GNI. However, the results are the same in both specifications (see Table S3 in the Supplementary Information).

[57]According to the magazine, the IIR ratings "are based on information provided by senior economists and sovereign-risk analysts at leading global banks and money management and securities firms. They have graded each country on a scale of zero to 100, with 100 representing those countries that have the least chance of default. ... The individual credit responses are weighted using a formula that gives more importance to responses from institutions with greater worldwide exposure and more sophisticated country analysis systems." The creditworthiness ratings are preferable to credit ratings because they are available for many countries before they were formally rated. While other work has been able to extract "shadow ratings" for these countries, having a real measure of their perceived risk is helpful for detecting what will likely be subtle effects. For works comparing various measures of creditworthiness, see Vij (2005); Oetzel et al. (2001); Ratha et al. (2011).

taking the mean of the September rating and the March rating from the following year, resulting in a single rating for a fiscal year, and for my baseline model, I impute zeroes for missing values to capture the poor reputation associated with being unrated. For ratings of a country's level of *Democracy*, I use the political rights score from Freedom House, treating it as a perceptual rather than latent variable. I flip the score so that higher levels refer to higher scores on political rights, from 1 to 7. Summary statistics appear in Table S2 in the Supplementary Information.

3.2 Estimation Strategy

The main challenge to identifying the effect of classifications is that international observers could be influenced by the determinants of a country's classification rather than by the classification itself. Fortunately, since the World Bank uses exact, public, and arbitrary cutoffs in GNI per capita to determine classifications, these factors can be controlled for directly. My main specification regresses outcomes of interest on both a dummy variable indicating that a country's GNI per capita surpasses a relevant cutoff and also on the continuous GNI per capita variable, which conceivably directly shapes observers' perceptions of the country independent of its classification. Controlling for this variable allows me to estimate the effect of the classification itself, β. The functional form for this specification is:

$$Y_{i,t} = \alpha + \beta Above\ cutoff_{i,t-1} + \delta log(GNIpc)_{i,t-1} + \gamma \mathbf{X}_{i,t-1} + \mu_i + \tau_t + \varepsilon \tag{1}$$

where Y is an outcome, $\mathbf{X}$ represents a vector of covariates, and t denotes the year or period of analysis. Country and period/year fixed effects are used, and I cluster standard errors by country. A similar specification is used by Knack et al. (2014), who study the effect of crossing the operational cutoff on aid allocations.[58] I evaluate the effect of different cutoffs within the same classification system in separate models, as different categories within the same classification system could understandably produce effects of various magnitudes or directions. In order to improve comparability across outcomes, I standardize all dependent variables.

The unit of analysis differs between *Aid* and other regressions. In keeping with Knack et al. (2014), *Aid* is aggregated into three-year periods corresponding with the IDA replenishment cycles of the World

[58]Despite using the same empirical model, the results I present in the next section will differ from those reported in Knack et al. (2014). This is because my sample includes two additional periods (data through 2015) that were not available at the time the authors conducted their analysis. A complete replication of the paper and more detailed explanation of our different findings appears in Section S6 of the Supplementary Information.

Bank, a grouping which reflects the natural decision-making timeline of most donors and also smooths an otherwise volatile variable.[59] In *Aid* regressions, I restrict the sample to just those countries that have ever benefited from IDA during the time frame. This is because countries that have never benefited from IDA are unlikely to exhibit any change in their aid receipts, so their inclusion only serves to reduce my power. Since all countries may receive FDI, credit ratings, and democracy ratings, I use samples of all countries in these regressions.

Many factors shape economic perceptions of countries. I therefore include several controls that are thought to influence the views and decisions of a variety of global economic elites. Larger countries are more likely to receive aid and attract international attention, so I control for the country's logged population. A country's level of democracy is known to influence the behavior of donors, investors, and credit raters, so I include lagged values of the Freedom House Political Rights score as a control in these regressions (but not when it is the dependent variable).[60] Any actor participating in a financial transaction with a country must also take into account the country's financial assets. I therefore include logged gross capital formation, which has been found by Vij (2005) to be one of the biggest predictors of the *IIR* rating.

I include these controls primarily to improve the precision of my coefficient estimates rather than to address bias in my coefficient estimates. This is because there are very few variables that would be systematically correlated with my main explanatory variable: whether a country is above the LIC or the LMIC threshold. I acknowledge, however, in Section 5.2, that sometimes a country may try to influence its position above or below important thresholds. Some evidence may suggest that this "juking the stats" behavior appears more commonly in authoritarian regimes.[61] Controlling for democracy will therefore mitigate this concern.

3.3 Results

The main results appear in Table 2. Consistent with the hypotheses, certain classifications affect the aid a country is able to attract from foreign donors, investors' perceptions of its creditworthiness, and ratings of its democracy, although its FDI inflows do not exhibit any effect. I proceed by discussing each of these

[59] All variables are aggregated using means, while income variables and the cutoff dummy are taken from the final year in each period.

[60] See, for example, Alesina and Dollar (2000); Li and Resnick (2003). Knack et al. (2014) also control for logged population and Freedom House. In addition, they include the World Bank's Country Policy and Institutional Assessment (CPIA) score, but these data are not publicly available, and the authors report that their results are robust to dropping this control.

[61] See Magee and Doces (2015) and Wallace (2016).

results in turn.

Overall, the income classification system affects a country's ability to attract aid, but these effects only appear when countries become "upper-middle income." Column 1 of Table 2 illustrates that "upper-middle income" countries experience a decline in their aid by a quarter of a standard deviation as a result of their label, even when raw GNI is taken into account. However, there is no detectable difference between "low income" and "lower-middle income" countries in the aggregate development assistance countries receive. This suggests that contrary to the concerns of countries approaching "middle-income country" status, donors do not appear to penalize countries for their growth until those countries have advanced into "upper middle-income" status, but when they do, the effect is sizable. Table 3 finds that this result is primarily driven by bilateral donors; multilateral assistance is not significantly affected by any country category. Interestingly, the effects are even stronger for non-traditional donors than they are for traditional donors, even though non-traditional donors are widely thought to ignore need in their foreign aid policy.[62]

Private investors, on the other hand, do not exhibit sensitivity to income classifications, at least in the aggregate. Column 2 of Table 2 shows that graduating to "lower-middle income" and later "upper-middle income" country status does not affect a country's net FDI inflows.

Sovereign risk analysts do appear to be swayed by the income classifications in their risk assessments. Column 3 of Table 2 shows that becoming "upper-middle income" is associated with a bump in a country's creditworthiness score, and this result is significant at the .1 level. This specification likely understates the effect at this threshold, as it imputes zeroes for all country-years with no rating. Table 4 presents the results without these imputations, in other words, using only the data for countries that received a creditworthiness score. In this specification, the coefficient size on graduating to "upper-middle income" status doubles and becomes significant at the .05 level. However, graduating to "lower-middle income" status also becomes significant yet produces a negative rather than a positive effect. This is because countries are more likely to be rated when they become "lower-middle income," as shown in column 3 of Table 4. This finding too suggests the importance of the income classifications, as being rated itself communicates an improvement in a country's standing. Despite the fact that sovereign lending involves careful analysis of economic fundamentals, these results suggest that even senior analysts responsible for these decisions are biased by how a country is classified.

[62] "Traditional" donors are those that are members of the Organization for Economic Cooperation and Development (OECD)'s Development Assistance Committee (DAC).

Table 2: The effects of classifications on behavior of donors, investors, and raters

	(1) Aid	(2) FDI inflows	(3) Creditworthiness	(4) Democracy
		A. Above LIC ceiling		
Above LIC ceiling (t-1)	−0.022	0.011	0.053	0.170**
	(0.063)	(0.069)	(0.059)	(0.071)
Constant	3.151	−0.121	−2.719	−1.669
	(5.124)	(4.078)	(3.841)	(4.421)
Covariates	✓	✓	✓	✓
Country F.E.	✓	✓	✓	✓
Period F.E.	✓	✓	✓	✓
Period	3-Year	Year	Year	Year
Observations	632	3,062	2,148	3,204
R^2	0.886	0.833	0.916	0.828
		B. Above LMIC ceiling		
Above LMIC ceiling (t-1)	−0.246**	−0.032	0.111*	0.027
	(0.107)	(0.048)	(0.064)	(0.069)
Constant	4.374	0.032	−3.661	−2.805
	(4.874)	(4.123)	(3.944)	(4.445)
Covariates	✓	✓	✓	✓
Country F.E.	✓	✓	✓	✓
Period F.E.	✓	✓	✓	✓
Period	3-Year	Year	Year	Year
Observations	632	3,062	2,148	3,204
R^2	0.887	0.833	0.917	0.827

Standard errors in parentheses
*p<0.1; **p<0.05; ***p<0.01

Note: The table reports coefficients from OLS regressions of the outcome on a dummy variable coded 1 if a country is above the cutoff, controlling for GNI per capita. Standard errors are clustered at the country level. Covariates include lagged values of log population, log gross capital formation, and Freedom House political rights score, and they include country and either year or period fixed effects. The Freedom House political rights score is inverted so that positive values are more democratic, and this covariate is omitted when it is the dependent variable. In the aid regressions, the sample is restricted to countries that have ever benefited from IDA after 1987. All dependent variables have been standardized for ease of comparison.

Table 3: The effects of classifications on aid by type of donor

	(1) DAC	(2) Non-DAC	(3) Multilateral
	A. Above LIC ceiling		
Above LIC ceiling (t-1)	−0.036	0.141	−0.035
	(0.069)	(0.106)	(0.078)
Constant	0.548	12.545*	2.040
	(5.364)	(7.073)	(4.852)
Covariates	✓	✓	✓
Country F.E.	✓	✓	✓
Period F.E.	✓	✓	✓
Period	3-Year	3-Year	3-Year
Observations	631	585	629
R^2	0.882	0.630	0.879
	B. Above LMIC ceiling		
Above LMIC ceiling (t-1)	−0.339**	−0.426***	−0.165
	(0.133)	(0.154)	(0.103)
Constant	2.340	14.984**	2.940
	(4.808)	(6.831)	(4.784)
Covariates	✓	✓	✓
Country F.E.	✓	✓	✓
Period F.E.	✓	✓	✓
Period	3-Year	3-Year	3-Year
Observations	631	585	629
R^2	0.884	0.633	0.880

Standard errors in parentheses
*p<0.1; **p<0.05; ***p<0.01

Note: The table reports coefficients from OLS regressions of the outcome on a dummy variable coded 1 if a country is above the cutoff, controlling for GNI per capita. Standard errors are clustered at the country level. Covariates include lagged values of log population, log gross capital formation, and Freedom House political rights score. The sample is restricted to countries that have ever benefited from IDA after 1987. All dependent variables have been standardized for ease of comparison.

Democracy scorers similarly exhibit a classification bias. Countries that reach "lower-middle income" status have systematically higher Freedom House ratings than their "low income" counterparts, and the result is statistically significant at the .05 level. This result is especially surprising, since Freedom House follows a rigorous coding procedure intended to capture the variable of interest, democracy. The presence of a classification effect even in these evaluations of a dimension unrelated to economic development is consistent with the widely documented "halo effect," by which units sharing one characteristic are assumed to share other characteristics. However, these classifications appear to more powerfully influence Freedom House ratings at the lower end of the income spectrum; no difference is detected between "lower-middle" and "upper-middle" countries.

Table 4: The effects of classifications on credit ratings

	(1) With imputations	(2) Without imputations	(3) Any rating (0-1)
		A. Above LIC ceiling	
Above LIC ceiling (t-1)	0.053	-0.156^{***}	0.185^{*}
	(0.059)	(0.055)	(0.104)
Constant	-2.719	14.489^{***}	-10.678
	(3.841)	(5.496)	(6.992)
Covariates	✓	✓	✓
Country F.E.	✓	✓	✓
Year F.E.	✓	✓	✓
Period	Year	Year	Year
Observations	2,938	2,148	3,202
R^2	0.851	0.916	0.768
		B. Above LMIC ceiling	
Above LMIC ceiling (t-1)	0.111^{*}	0.197^{**}	0.054
	(0.064)	(0.077)	(0.094)
Constant	-3.661	13.820^{**}	-12.039^{*}
	(3.944)	(5.546)	(7.066)
Covariates	✓	✓	✓
Country F.E.	✓	✓	✓
Year F.E.	✓	✓	✓
Period	Year	Year	Year
Observations	2,938	2,148	3,202
R^2	0.851	0.917	0.767

Standard errors in parentheses
*p<0.1; **p<0.05; ***p<0.01

Note: The table reports coefficients from OLS regressions of the outcome on a dummy variable coded 1 if a country is above the cutoff, controlling for GNI per capita. Standard errors are clustered at the country level. Covariates include lagged values of log population, log gross capital formation, and Freedom House political rights score. All dependent variables have been standardized for ease of comparison.

These results are robust to several alternative specifications. One concern is that including Freedom House as a control variable in the aid, FDI, and creditworthiness regressions introduces endogeneity, since I have demonstrated that the classifications influence this measure. In the Supplementary Information, I show that my results are robust to dropping this control as well as all controls. The aid results also do not change by using yearly rather than periodic observations. Finally, Section S2 in the Supplementary Information presents the results from an alternative model, which uses change scores and holds changes in GNI per capita constant to identify the effect of crossing a cutoff. This model focuses on explaining year-to-year variation in how observers react to a single country's change in category, omitting the country-to-country variation in how observers react to countries just above and just below a threshold. The results from this

exercise are similar to those yielded by the baseline model.

3.4 Discussion

With some exceptions, the classification effects I discovered are consistent with those predicted by my theory. I found strong support for hypothesis 1: Donors are particularly susceptible to the income classifications. Donors are significantly more likely to prefer allocating aid to LMICs than to UMICs. Curiously, there is no evidence of donor departure following a country's graduation from LIC to LMIC status. It is possible that publicity about the challenges LMICs face has succeeded in mitigating this classification effect, while UMICS receive no such attention. This is consistent with the strategic story I outline, as donors must justify their allocation of a scarce resource across a pool of competitive applicants based on common understandings of country need.

I also found evidence in favor of hypothesis 3: Countries with higher classifications are significantly more likely than countries with lower classifications to receive favorable democracy and creditworthiness scores. This is true at both the LIC/LMIC and LMIC/UMIC transitions. When a country crosses the LIC ceiling, it receives a higher democracy rating and it is more likely to receive a credit rating (even if it's a poor one). When a country crosses the LMIC ceiling, again, both its credit and democracy rise. If, as I argue, a cognitive mechanism is responsible for these results, it is not surprising that both ceilings produce effects, as rational bases for treating them differently should not apply.

I did not find support for hypothesis 2, as classifications do not produce any observable change in FDI. This null finding could result either because private investors are simply not be susceptible to the cognitive bias produced by classifications, or because my assumption regarding the direction of the classification effect is flawed. If higher classifications attract some investors but repel others, then the aggregate effect could be ambiguous. My study is not able to distinguish between these. Nonetheless, a finding that investors do not exhibit a cognitive bias simply speaks to the limitations of the classification effect rather than negating its power over other users.

Taken together, my findings suggest that the cognitive mechanism likely accounts for at least some portion of the classification effect. This is because the strategic mechanism is unable to account for the effects I detected in raters' behavior. It is not possible given this design to explain what portion of the aid findings are attributable to a cognitive versus a strategic mechanism. However, I argued that donors are the least susceptible to cognitive biases of any of my observers, suggesting that some portion of their sensitivity

to the classifications is likely explained strategically. Distinguishing these mechanisms is a topic for future research.

4 Classification Experiences

In the previous section, I exploited the opportunities for quantitative analysis that result from the strict set of thresholds that make up the World Bank's income classification system. This section, however, exploits complementary opportunities for qualitative analysis that result from systems in which countries have greater discretion over their categorization, such as the UN's LDC category and the World Bank's IDA category. In these systems, graduation is a deliberative process involving numerous stakeholders tasked with evaluating the country's preparedness for graduation. I interviewed 51 of these stakeholders regarding the graduation processes of 16 countries to investigate the anticipated or historical effects of each country's graduation. The results not only support the hypotheses qualitatively but also illustrate that these dynamics can be found across multiple classification systems.

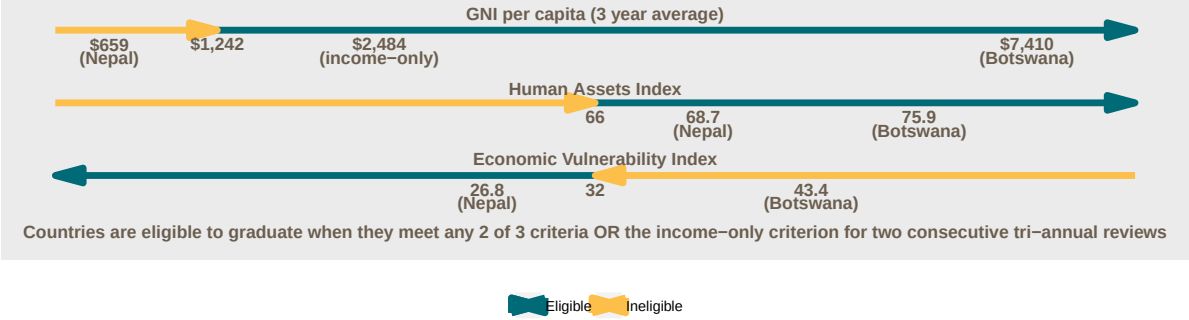
I start by telling the stories of Nepal's (upcoming) and Botswana's (historical) graduations from the UN's LDC category from the perspectives of various domestic groups. My theory implies that actors who benefit from their country's international reputation will embrace the rewards that come with graduation, while groups who depend on foreign aid will suffer its costs. This is exactly what I find. Next, I show that these two examples are consistent with the experiences of a broader sample of World Bank graduates, as well as with the hypotheses.

4.1 Examples from Nepal and Botswana

Consider the cases of Nepal and Botswana, where I conducted 31 interviews with current and former government officials, local experts, and leaders in the business community and civil society. Examining these graduations from the perspective of various domestic groups illustrates that a country's international classification produces many different effects, creating both winners and losers. On the one hand, certain political and business elites have incentives to pursue the improved international reputation that comes with graduation. On the other hand, the costs of losing international assistance and advocacy are diffuse and are frequently born by the most vulnerable populations.

Nepal will likely soon graduate from the LDC category. As can be seen in Figure 3, in 2015, the UN

Figure 3: Nepal's and Botswana's performances relative to LDC graduation criteria (2015)



Note: Figure plots Nepal's and Botswana's 2015 performances on the indicators used to determine a country's LDC categorization. Nepal is currently classified as a LDC, but, as can be seen, in 2015 it became "pre-eligible" to graduate because it met graduation thresholds on two of three criteria. If this continues in the next 2018 triennial review, Nepal will be eligible to graduate. This process is not automatic, however, as the CDP or UNCTAD could choose not to advance a graduation recommendation. While Nepal receives the opportunity to voice its own views during deliberation, it must accept the final recommendation. Botswana graduated from the LDC category in 1994 on the basis of the criteria at that time. This figure shows, however, that its EVI remains quite high relative to the graduation threshold. However, this is not sufficient for Botswana to be re-included in the LDC category, a process which is depicted in Figure S3 in the Supplementary Information.

concluded that Nepal was eligible to graduate from the LDC category, and according to UN protocol, Nepal may (but is not required to) graduate in 2022, so long as it is found eligible again in its upcoming 2018 review. Nepal is a LIC and fails to meet the income criterion for LDC graduation, but has met the required two of three criteria due to its achievements on the Human Assets Index and the Economic Vulnerability Index.

Nepal first committed to explicitly pursuing LDC graduation in the 13th National Plan (2013-2016), which was drafted shortly after Nepal chaired an international meeting in Istanbul in which LDCs collectively aimed for a 50 percent graduation rate within the next decade. The Vice Chairman of the NPC at the time summarized his experience at this meeting:

> I had a chance to talk to many leaders of LDCs, and we would ask each other: How long do you want to be called an LDC? Do you ever want to graduate or shall we just continue being LDCs forever? Because LDCs, you know, they are thought to have lazy people and want to depend on donors, and donors are just tired of providing for them. If anyone is going to develop, it has to come from us. We had very strong feelings that we wanted to send a message to the developed world that we could take care of ourselves. Well, Nepal was signatory to this message at the International Conference of LDCs. After this, I became Vice Chairman of the NPC. And I prepared the 13th plan and included that objective to do whatever was necessary to get out of

> LDC status by 2022. When I turned to the government to get that document approved, I said: Let us tell the world that we are capable of graduating and will no longer be part of the LDC community.[63]

But many individuals associated with think tanks and aid organizations are concerned about the departure of donors following graduation. One director of a research institute with experience in the National Planning Commission and Ministry of Foreign Affairs in Nepal argued that Nepal's LDC status was vital to the assistance it was able to secure from European donors. In particular, he cited the recent closing of the Danish embassy in Kathmandu as evidence of the perception that Nepal no longer needed assistance: "They argued that their priority had shifted to poor African countries that needed their support. If we cannot brand ourselves as a country in need of development, we will lose these things."[64] Another think tank director pointed to the importance of the special treatment and duty-free access Nepal receives in the World Trade Organization as result of its LDC status, arguing that if Nepal graduates, "the international community should know that it doesn't reflect a real structural transformation, and they should continue to offer these benefits to us."[65] The leader of a widely-respected Nepali NGO grounded his expectation of donor flight in reports he had heard from his colleagues in Vietnam, who encountered aid losses after Vietnam graduated from LIC status: "Why should aid agencies give money to them when there are so many other countries that need aid?"[66] As the UNDP Country Director in Nepal commented, "Nepal has so much more to gain by being the focus of development partners rather than trying to signal to the rest of the world, 'Oh we can go it alone.'"[67] The expectation of donor flight has led one development consultant who works with NGOs in Nepal to encourage his clients to diversify their sources of funding away from traditional donors.[68]

Indeed, momentum for Nepal's LDC graduation has slowed, and the new cohort of government officials I spoke with are ambivalent about the prudence of pursuing LDC graduation in 2022. However, they carefully avoid voicing concerns about aid in their rationales, instead critiquing graduation protocol on technical grounds. A current member of the NPC responsible for the most recent national plan argued that the in-

[63] Author's interview with Dr. Rabindra Kumar Shkaya, Former Vice Chairman of the NPC, July 20, 2017, Kathmandu, Nepal.

[64] Author's interview with Dr. Nishchal Pandey, Executive Director, Center for South Asian Studies, June 22, 2017, Kathmandu, Nepal.

[65] Author's interview with Dr. Posh Raj Pandey, Executive Director, South Asia Watch on Trade, Economics, and the Environment, July 20, 2017, Kathmandu, Nepal.

[66] Author's interview with Mr. Gyan Adhikari, Executive Director, Rural Reconstruction Nepal, July 11, 2017, Kathmandu, Nepal.

[67] Author's interview with Mr. Renaud Meyer, UNDP Country Director, July 12, 2017, Kathmandu, Nepal.

[68] Author's interview with development consultant, June 23, 2017, Kathmandu, Nepal.

dicators used did not accurately capture development in a landlocked country like Nepal, concluding, "We see little point in endorsing this technical process when its criteria are unhinged from reality."[69] Instead, government officials intimate that Nepal should delay its graduation until it can graduate on all three criteria, otherwise graduation would not be "meaningful."[70] At the same time, government officials are quick to dismiss the concerns of "NGO activists who think the sky will fall when we graduate because of the loss of aid. NGOs of course survive on portraying Nepal as a helpless country. But look at the earthquake rescue money from the government relative to what we received from abroad — it's miniscule."[71] The real reason for stalling graduation remains unclear, since the concept of a "meaningful" graduation exists nowhere else.

Despite this portrayal of the NGO community, very few NGOs I spoke with were familiar with LDC graduation, as were surprisingly few business leaders. Even though one of the main benefits of LDC categorization is special treatment in the WTO, the chairman of the Export Council of Nepal and the Export Committee of the Nepal Chamber of Commerce was not familiar with the LDC category or its benefits.[72] In contrast, a small but vocal group of elites push at the highest levels for Nepal's graduation, even and especially if aid losses occur. One director of an economic think tank argued, "Nepal has always projected itself as a poor country as a means of rent-seeking. They know the right story to tell you, one that feeds into the business of development. Nepalis must change their mindset from poverty brokers to prosperity catalysts."[73] As a result, one expert concluded that if the UN recommends graduation, "our politicians will happily accept that decision and tell people it was their achievement. Civil society and the business community aren't aware enough to mobilize so politicians get to claim an accomplishment in the news."[74] While specific political interests exist to support graduation, its consequences will be born by an uncoordinated and unaware constituency.

These dynamics can be seen in hindsight in the case of Botswana, an UMIC and one of only five historical LDC graduates.[75] Despite having left the category over two decades ago in 1994 and being widely considered a poster child for development success, Botswana tops global rankings in inequality and HIV prevalence. In addition, lack of economic diversification has caused Botswana's Economic Vulnerability

[69] Author's interview with member of National Planning Commission, July 9, 2017, Kathmandu, Nepal.
[70] NPC Vice-Chairman Govinda Raj Pokharel in *The Kathmandu Post*, April 2015.
[71] Author's interview with member of National Planning Commission, July 9, 2017, Kathmandu, Nepal.
[72] Author's interview, July 14, 2017, Kathmandu, Nepal.
[73] Author's interview with Mr. Sujeev Shakya, Chairman, Nepal Economic Forum, July 18, 2017, Kathmandu, Nepal.
[74] Author's interview with Dr. Posh Raj Pandey, Executive Director, South Asia Watch on Trade, Economics, and the Environment, July 20, 2017, Kathmandu, Nepal.
[75] Others include Cabo Verde, the Maldives, Samoa, and Equatorial Guinea. Vanuatu will graduate in 2020 and Angola in 2021.

Index to slip below the LDC inclusion threshold, although stagnation on this single indicator is insufficient to push a country back into LDC status.[76] For this reason, several interviewees expressed skepticism regarding the value of the income classification system or the LDC category. As one civil society leader stated, "The media account of Botswana is not what's on the ground, and international NGOs are prejudiced by the label. They look at Botswana and see a wealthy country, but they don't see the needs of the neediest."[77] Not only is the label perceived to be inaccurate, but voices in civil society and in the business community agree that Botswana's outstanding international reputation has actually played a role in contributing to the the country's political and economic stagnation by concealing its problems.

The concept of donor flight is well-known in Botswana. According to one economist with decades of experience in the country, Botswana managed to delay the departure of donors for some time after Botswana became a MIC by arguing that the government used resources effectively, but the LDC graduation again highlighted Botswana's success and made it more challenging to justify continued assistance. As in Nepal, many countries closed their embassies, which had primarily existed for the purposes of bilateral aid.[78] Development agencies from Sweden, the United Kingdom, and Norway all pulled out, and the U.S. Agency for International Development (USAID) ceased supporting the salaries of technical officers in the government.[79] One civil society leader mentioned that his organization was recently rejected for funding by the Gates Foundation, which stated that it was trying to dedicate more funding for Africa. He commented, "Botswana isn't Africa enough for the Gates Foundation. You can't talk to any international donor and say you're from Botswana. It's perceived as too wealthy."[80]

The dearth of international support for Botswana creates problems not because Botswana needs external resources but because it has crippled the development of a robust civil society that can hold the government to account. Most (although not all) acknowledge that Botswana is relatively more successful than other developing countries and that there are sufficient resources in the country to address its problems. But one civil society leader pointed out, "Developing buildings and infrastructure is not congruent with developing people. Organizations like the World Bank and the African Development Bank say we have enough re-

[76] See Figure S3 in the Supplementary Information.

[77] Author's interview with Mr. Oscar Motsumi, Executive Secretary, Botswana Network of AIDS Service Organizations (BONASO), August 8, 2017, Gaborone, Botswana.

[78] Author's interview with Dr. Keith Jefferis, Managing Director, Econsult Botswana, August 10, 2017, Gaborone, Botswana.

[79] Author's interview with Ms. Alice Mogwe, Director, Ditshwanelo Botswana Centre for Human Rights, August 10, 2017, Gaborone, Botswana. Author's interview with Dr. Jay Salkin, Director of Research, Bank of Botswana, August 9, 2017, Gaborone, Botswana.

[80] Author's interview with Mr. Oscar Motsumi, Executive Secretary, Botswana Network of AIDS Service Organizations (BONASO), August 8, 2017, Gaborone, Botswana.

sources to develop our people, but that aspect of development is not happening. This is why NGOs must complement the work of the government. But how do NGOs generate funding? They ask for money from donors. Unfortunately, because of Botswana's classification, most bilateral agreements that we used to have no longer exist."[81] But civil society organizations primarily rely on the government for funding, and as one leader pointed out, "they don't bite the hands that feeds them. You see NGOs leaning more and more toward the government, which limits the advocacy agenda."[82] What limited international funding exists for Botswana exists only for organizations focused on HIV/AIDS and LGBTI issues, what one interviewee called "the international development community's flavor of the month."[83] While this dynamic exists in all aid recipient countries, it is exaggerated in a country like Botswana, where a claim for international assistance must be based on a specific thematic objective as opposed to general need. It is no surprise that most civil society organizations work on these issues or attempt to anchor their own agenda on these issues. But as a result, marginalized groups falling outside these bounds are neglected. For example, an initiative to serve people with disabilities under the office of the President, while well-intentioned, is judged by observers to be significantly outdated, yet there are no NGOs actively holding government to account on this issue.[84]

Of course, Botswana's outstanding international reputation does come with perks, allowing Botswana to "punch above its weight in international affairs."[85] But there are also concerns that the reputation is also a liability because "it produces the assumption that we don't need to change anything. Change is already hard. It makes it more difficult when the rest of the world says you're wonderful."[86] Most of Botswana's historic growth rates occurred in the 1970s and 1980s as the result of prudent management following the discovery of diamond reserves and the successful transition from the agrarian to the mineral economy. However, economists stress the importance of making a second transition from a mineral to an industrial economy. The last few decades have produced slow growth, few jobs, little entrepreneurship, and failure to adopt important policy reforms. More controversially, others criticize the government for regressing in

[81] Author's interview with Ms. Cindy Kelemi, Executive Director, Botswana Network on Ethics, Law, and HIV/AIDS, August 15, 2017, Gaborone, Botswana.

[82] Author's interview with Mr. Oscar Motsumi, Executive Secretary, Botswana Network of AIDS Service Organizations (BONASO), August 8, 2017, Gaborone, Botswana.

[83] Author's interview with Ms. Alice Mogwe, Director, Ditshwanelo Botswana Centre for Human Rights, August 10, 2017, Gaborone, Botswana.

[84] Author's interview with Ms. Cindy Kelemi, Executive Director, Botswana Network on Ethics, Law, and HIV/AIDS, August 15, 2017, Gaborone, Botswana.

[85] Information in this paragraph based on author's interview with Dr. Keith Jefferis, Managing Director, Econsult Botswana, August 10, 2017, Gaborone, Botswana.

[86] A similar sentiment was expressed by Diamond Hub coordinator Mmetla Masire in a national newspaper: "We must stop gloating about international ratings which are saying we have the best run economy in Africa and face the reality because we are on the ground and know what is happening." *The Patriot*, November 19, 2014.

transparency and civil liberties.[87] To be sure, continued failure to match Botswana's reputation with reality will likely eventually attract global attention, but reputations are sticky, and Botswana's elevated status arguably permits early backsliding to go unnoticed. As one respondent summarized, "Nobody wants to rock the boat and point out that for the last 10 years, Botswana has been on a downward spiral."[88] The case of Botswana raises the possibility that the so-called "middle income trap" may be caused partially by the "middle income country" category itself.

Taken together, the examples Nepal and Botswana illustrate that political leaders enjoy and pursue the improved international reputation associated with graduation, but domestic populations often pay for it. Donors pull out (hypothesis 1) because they correctly perceive that governments are capable of addressing their own problems, but when governments are unwilling to do so, vulnerable populations are left with little recourse. It is not only that they lose access to international resources, making them more dependent on the government. In addition, inflated international perceptions of the country's performance (hypothesis 3) make it even more challenging for marginalized groups to call international attention to the government's failure to address their needs. This suggests that classifications can exacerbate inequality and hinder not only development but also democracy. By revealing the benefits and costs of graduation as experienced by involved parties, these cases provide further evidence that classification effects are real and have distributive consequences.

4.2 Themes in Graduation Experiences

In this section, I explicitly connect the experiences of Nepal and Botswana to the hypotheses and mechanisms, while also illustrating the external validity of my argument by presenting evidence from an additional, broader sample of 14 graduating countries in the World Bank. This evidence comes from 20 interviews I conducted in New York and Washington, DC with diverse stakeholders involved in the IDA graduation process and the income classification system. I include the perspectives of advisors to the World Bank from past and future graduating countries, representatives from donor countries to the IDA replenishment meetings, desk officers, and critical observers of the World Bank.

Consistent with hypothesis 1, which predicted that higher classifications would result in less aid, both

[87] Botswana, coded a dictatorship by the DD approach, has been ruled by the Botswana Democratic Party since independence, although its previous presidents have been globally recognized for their commitments to democracy.

[88] Author's interview with Mr. Oscar Motsumi, Executive Secretary, Botswana Network of AIDS Service Organizations (BONASO), August 8, 2017, Gaborone, Botswana.

Nepal and Botswana very obviously illustrated donor flight following graduation. Indeed, in the World Bank sample, many countries associate graduation with losses in aid. Several interviewees reported that graduating from IDA caused them to lose access to other sources of aid. One of the current candidates for graduation, Vietnam, has already reported losing access to bilateral funds as a result of its candidacy.[89] The same has occurred for previous graduates Armenia and Georgia. A former World Bank official from Georgia stated that one donor backed out of a trust fund and cited Georgia's loss of IDA status as its reason for doing so. Furthermore, following Armenia's and Georgia's 2014 graduation, Georgia nearly lost its regional hub because it was viewed as less important to World Bank operations.[90] A World Bank Board official representing Armenia and Georgia noted that during 2016 negotiations over graduation policy reform, a proposal for access to an additional source of financing called IDA+ for recently graduated countries was withdrawn because "IDA is for the poorest countries."[91] Even the countries that have not yet become categorized as Blend anticipate this dynamic. One advisor spoke of Zambian officials' concern about losing access to aid, particularly after crossing the LMIC/UMIC threshold.[92] In summary, one advisor stated, "Proposed IDA graduation reforms do not address the issue of perception of countries as IDA and non-IDA by the donor community."[93] The perception that donors coordinate on World Bank classifications was widespread; one IDA deputy stated, "we all know it goes on."[94]

While some of the post-graduation aid losses could be explained through cognitive biases at work in the minds of donors, other pieces of evidence are better explained by a strategic mechanism. Many donors formally tie eligibility for their assistance to the World Bank analytical income classification system. Even though the World Bank has never attached any material benefits to these categories, the Global Fund to Fight AIDS, TB, and Malaria,[95] the Millennium Challenge Corporation,[96] Gavi,[97] and the OECD DAC all include income classifications in their eligibility policies.[98] Médecins Sans Frontières notes that MICs are often excluded from voluntary license agreements with pharmaceutical companies and are therefore penalized in

[89] Author's interview with IDA deputy, July 22, 2016, Washington, DC.

[90] Author's interview with former advisor to Executive Director, August 2, 2016, Washington, DC.

[91] Author's interview with advisor to Executive Director, August 3, 2016, Washington, DC. The recent graduates were Angola, Armenia, Bosnia-Herzegovina, and Georgia.

[92] Author's interview with advisor to Executive Director, July 21, 2016, Washington, DC.

[93] Author's interview with former advisor to Executive Director, August 2, 2016, Washington, DC.

[94] Author's interview with IDA deputy, July 22, 2016, Washington, DC.

[95] http://www.theglobalfund.org/documents/core/eligibility/Core_EligibilityAndCounterpartFinancing_Policy_en/

[96] https://www.mcc.gov/resources/doc/report-candidate-country-fy-2016

[97] http://www.gavi.org/support/apply/countries-eligible-for-support/

[98] http://www.oecd.org/dac/stats/daclist.htm

tiered pricing schemes, limiting the provision of vaccines to countries with some of the highest rates of HIV/AIDS.[99] To rely on classifications as heuristics is one thing; to formally commit to using them as a matter of policy demands a higher level of scrutiny and defense, making it less likely that cognitive biases account for these behaviors. Furthermore, while representatives of donor governments are disincentivized to be candid about strategic incentives for their allocations, some nonetheless are. As one U.S. Treasury official commented, "When we can show that most of our aid is benefiting LDCs or fragile and conflict-affected states, this helps our numbers."[100]

The Nepal and Botswana cases present mixed evidence on hypothesis 2, concerning FDI. One World Bank country economist for Nepal spoke of the "positive psychological effect" Nepal's graduation from the LDC category would have on private investment, which he viewed as beneficial to a country with insufficient public investment.[101] An expert with former government experience agreed, anticipating that graduation would improve investor confidence in Nepal. He cited multiple instances, including the puzzling departure of a hydroelectric multinational corporation and the unwillingness of European insurers to cover domestic airlines in Nepal, in which he believed Nepal's status as an LDC had undermined international confidence in its capability, since in both cases there was a rational economic basis for investment in Nepal.[102] Both respondents felt that while graduation would not change the underlying fundamentals of the Nepalese economy, it could psychologically reassure private foreign investors. Botswana's outstanding reputation, according to hypothesis 2, should attract significant amounts of FDI. But Botswana has tried to attract FDI for decades without much success, except in the sectors of mining, banking, and retail. Economists attribute this failure to difficulties in obtaining work permits and outdated immigration policies that make it challenging for foreign firms to operate.[103] This evidence suggests that classifications are not enough to shift investor perceptions, although given its lack of economic diversification, Botswana is perhaps a very hard test of this.

Hypothesis 3 predicted that higher classifications would improve a country's reputation on unrelated indicators, democracy and creditworthiness. Indeed, many respondents' comments implicitly supported the notion of a "halo effect" that exists with respect to classifications.[104] Even respondents with precise technical

[99] https://www.msfaccess.org/sites/default/files/MSF_UTW_17th_Edition_4_b.pdf

[100] Author's interview with U.S. Treasury official, August 5, 2016, Washington, DC.

[101] Author's interview with World Bank economist, June 28, 2017, Kathmandu, Nepal.

[102] Author's interview with Dr. Nishchal Pandey, Executive Director, Center for South Asian Studies, June 22, 2017, Kathmandu, Nepal.

[103] Author's interview with Dr. Jay Salkin, Director of Research, Bank of Botswana, August 9, 2017, Gaborone, Botswana.

[104] One graduation that does clearly correspond to an increase in creditworthiness is the graduation from IDA to IBRD, although

knowledge of the criteria involved in LDC graduation cited irrelevant accomplishments that graduation would signal, such as overcoming civil war, increasing mobile phone coverage, and better governance. One respondent generalized to say that Nepal's graduation would signal that "we are a normal country," the most general possible halo.[105] In Botswana, the halo effect was exactly what frustrated civil society groups representing marginalized groups. One respondent explicitly connected Botswana's status as a UMIC and an LDC graduate to its international reputation as a democracy, arguing that international actors (incorrectly) perceive Botswana to be a place of peace, political stability, and respect for civil liberties.[106]

5 Responses to the Classification Effect

According to hypothesis 4, if classifications are so powerful, then affected countries and groups will try to influence which classification they receive to the extent that they are able. This section shows that countries and groups indeed try to manipulate their classifications. According to my theory, whether a country seeks a higher or a lower classification should depend on how much that country values aid relative to improved investment or borrowing opportunities. However, testing this theory would involve imposing assumptions about how the diverse interests of government officials and various domestic groups are aggregated into a unitary foreign policy. Since this preference aggregation process is not obvious, I leave it to future research to explain which objectives states pursue, using this space primarily to explain how they can manipulate classifications and that they do.

5.1 Lobbying over Classifications

One extreme method by which the classified can respond strategically is to attempt to affect the classification system itself. The "Raise the MIC" movement described in the introduction is organized by an advocacy group claiming to represent civil society organizations in classified countries who are most disadvantaged by higher classifications. Organized by the AIDS Healthcare Foundation (AHF), the group argues that the "middle-income label" is a primary obstacle for NGOs trying to obtain foreign aid from the Global Fund

this relationship can be explained rationally, since the World Bank's assessment of a country's creditworthiness is one (subjective) component in this process. An IDA officer formerly in charge of assessing the creditworthiness of IDA graduates for IBRD lending noted that countries often pursue IDA graduation as a "stamp of approval" because it helps reduce their borrowing costs on the open market. Author's interview with IBRD credit risk officer, July 26, 2016, phone interview.

[105] Author's interview with Dr. Nishchal Pandey, Executive Director, Center for South Asian Studies, June 22, 2017, Kathmandu, Nepal.

[106] Author's interview with Mr. Oscar Motsumi, Executive Secretary, Botswana Network of AIDS Service Organizations (BONASO), August 8, 2017, Gaborone, Botswana.

and discounted vaccine prices from pharmaceutical companies. The movement has been endorsed by hundreds of health, education, clean water, and human rights organizations in middle-income countries. These organizations are often primarily funded by foreign donors and may even be at odds with their governments, making it unlikely that their interests will be represented following a country's "graduation."[107] Even the particular choice of strategy adopted by the Raise the MIC movement itself points to the power of classifications: The Raise the MIC movement primarily seeks change on the part of the classifying organization and not the users of the classification. While the group also lobbies the Global Fund (a user) directly, most of its resources are spent lobbying the World Bank (a producer) to publicly and unequivocally clarify that its classifications are not intended to influence funding or pricing. According to the movement, this "would indicate to the Global Fund that they need to change the system... [and] would give more ammunition to civil society to pressure drug companies to lower their prices in middle income countries."[108] This logic suggests that the movement recognizes the widespread legitimacy of the World Bank to define a classification and that the pressure applied by civil society will only succeed when backed by the World Bank's endorsement. That the World Bank receives the brunt of the blame for other actors' use of a classification suggests that the blame-avoiding strategies employed by classification users are successful.

But while seeking change in the method of classification is an effective approach for groups representing international constituencies (e.g. health NGOs), simply seeking to change which classification they receive is a more effective strategy at the national level. Indeed, in the case of IDA graduation and LDC graduation, country consultations are required before eligible countries are recommended for graduation, presenting eligible countries with the opportunity to support or oppose their own graduations. In the case of IDA graduation, steep losses in aid and borrowing volumes unanimously deter countries from beginning the graduation process prematurely and frequently cause them to protest their graduation to influential donors through back-door conversations.[109] In the case of LDC graduation, while some countries do resist graduation, others accept or even accelerate it. Bangladesh, for example, has negotiated with the UN's CDP so that

[107] Author's interview with Denys Nazarov, Associate Director of Global Policy, AIDS Healthcare Foundation, April 7, 2017, phone interview.

[108] Author's interview with Denys Nazarov, Associate Director of Global Policy, AIDS Healthcare Foundation, April 7, 2017, phone interview.

[109] Vietnam and Sri Lanka vocally opposed their graduation from IDA during replenishment meetings in 2016. Author's interview with IDA deputy, July 22, 2016, Washington, DC. In the case of IDA graduation, whatever improved credit countries can privately access as a result of graduation is far outweighed by the losses in borrowing volumes they experience within the World Bank. Currently the IBRD faces significant capital constraints due to the low interest rate environment and is unable to accommodate the volumes of lending demanded by recent IDA graduates. Moreover, the "blend" status allows countries to access both IDA and IBRD, so prolonging their time at this status is nearly always advantageous.

it can graduate ahead of schedule in time for its 50 year anniversary in 2021.[110]

5.2 Gaming the System

When graduation is deterministic, no amount of persuasion behind closed doors can influence a classification. In the case of the World Bank's income classification system, countries must actually change their national income figures. But there is compelling reason to believe that they do so. A growing literature investigates the extent to which national bureaus of statistics are capable of manipulating national economic information, finding mixed results. Recent evidence finds that that national governments "juke the stats" to appear more successful.[111] Using the GDP figures reported in the World Development Indicators, Magee and Doces (2015) and Wallace (2016) provide evidence that GDP growth rates reported by authoritarian regimes are inflated when benchmarked against growth rates of measures that are harder to manipulate — nighttime lights and electricity, respectively.[112] These works argue that all regimes have an incentive to report rosy economic data to improve the odds of re-election or reappointment and reduce the possibility of collective action from below but that democracies have institutional checks that make this impossible. Authoritarian regimes may not be the only ones interested in manipulating figures; Sandefur and Glassman (2013) argue that pay-for-performance incentives from international donors encourage aid recipients to inflate their vaccination rates. Conversely, it is possible that international organizations are complicit in statistical manipulation: Dreher et al. (2008); Lang and Presbitero (2017) find evidence of bias in IMF growth and inflation statistics and debt sustainability analyses, which favor countries that are aligned with prominent shareholders.

Some studies even show that statistical manipulation takes place at relevant thresholds. Previous works especially germane to my study have used the McCrary statistical density test to demonstrate systematic underreporting of GNI per capita in an attempt to remain below the World Bank's operational threshold.[113] These tests, however, are misleading for two reasons. First, they cannot illustrate heterogeneous strategies. I have argued that the classifications are associated with both costs and benefits, suggesting that certain countries may prefer to be under-classified, while others may prefer to be over-classified. If countries try to

[110] Author's interview with UN official, May 9, 2017, New York, NY.

[111] See Wallace (2016).

[112] Hollyer et al. (2011) provide the basis for these tests by showing that authoritarian regimes are less likely to provide data; the subsequent works have then examined the bias of the provided data.

[113] Kerner et al. (2015) observe significantly more observations just below the operational threshold in the historical data from the World Bank Atlas, but not in the revised figures currently found in the WDI.

move in different directions, these effects could cancel each other out and present the illusion of continuity, or small discontinuities may underestimate the extent of manipulation occurring in both directions. Second, these tests will underestimate strategic behavior if some attempts fail. For example, if a country close to a threshold tries to jump just over it but fails, it can easily appear to be a country that looks like a country that tried to slide just under the threshold. Since countries are operating in a tight range, it is likely that failures are not uncommon.

To overcome these limitations, I test whether countries are more likely to revise their national income figures as they approach an important threshold. Revision to national income data can occur for many benign reasons, but if countries revise their income data to influence their classifications, then we would expect these revisions to be strategically timed as countries try to avoid or accelerate crossing certain cutoffs. This approach avoids any assumptions about whether figures are revised upwards or downwards, and it measures countries' attempts rather than their successes. In what follows, I describe the unique "time machine" data that supports this exercise and present the results.

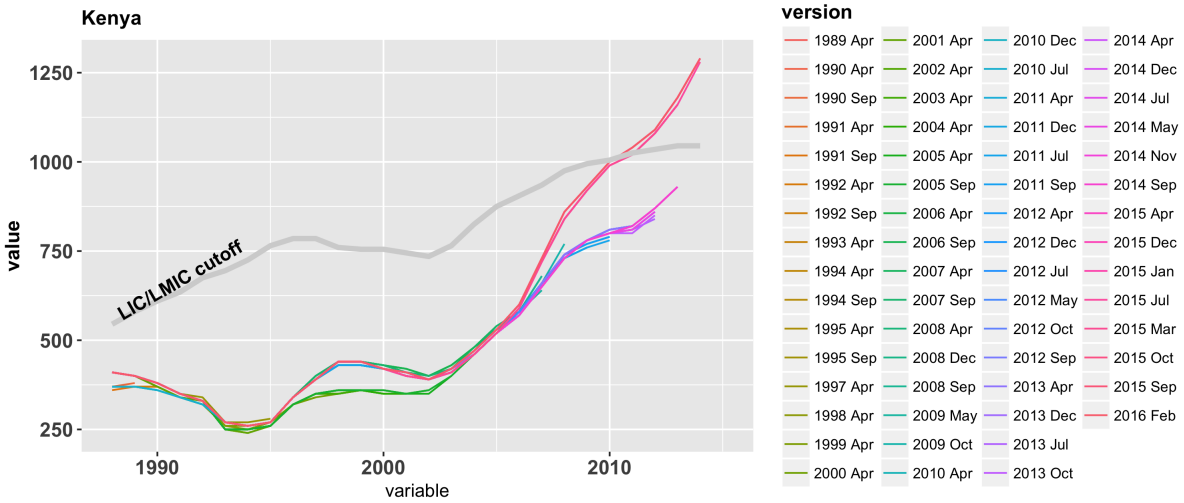
As previously mentioned, the GNI figures that were used at the time of a country's classification sometimes differ from the current version of those historical estimates. These differences are not inconsequential: In 288 country-years, a country received a different classification than the one it would have received on the basis of our current estimate of its GNI in that year.[114] Fortunately, the World Bank makes all previous versions of the World Development Indicators available for download, allowing me to detect at what point in time substantial revisions to GNI are made.

Figure 4 illustrates the unique variation reported in this "time machine" data set. Each panel plots a different country's income over time according to every version of World Bank data that has ever been published, with each version in a different color. Wherever there is vertical distance between colored lines, ex post estimates of an economy's national income have been revised in a version of the WDI indicated by the color. In these examples, the estimates diverge dramatically in the 2010s, when the countries rebased and revised their estimates of previous years ex post. The fact that different colors of lines straddle the cutoff separating LICs and LMICs indicates that these countries would have been classified differently had researchers known what we know now.

Intuitively, my objective is to calculate the magnitude of the revisions reported in each of the versions of the WDI (the area between one colored line and the colored line from the previous version of the WDI) and to

[114] Descriptive statistics on these "misclassifications" are available in Table S18 in the Supplementary Information.

Figure 4: National income data is revised over time: the example of Kenya



Note: Figure depicts Kenya's GNI per capita over time according to every available version of the World Development Indicators, each depicted by a line with a different color. The x-axis represents the year described by the data, while the color of the line indicates the version of the data. The separation between lines occurring in the early 2010s results from a statistical rebasing exercise in which Kenya re-estimated the size of its economy. The grey line indicates the the ceiling separating LICs and LMICs, which changes over time only to account for inflation. *Source*: World Bank Database Archives, OGHIST.

relate that magnitude to the country's distance from a threshold. To operationalize this, I start by calculating, for each version of the WDI, how much a country's estimates of the previous three years of GNI differed from its estimates of those same years as reported by the previous version of the WDI. Mathematically, I define a revision for each version as the following:

$$ThreeYrRevision_{i,j} = (GNI_{i,j} - GNI_{i-1,j}) + (GNI_{i,j-1} - GNI_{i-1,j-1}) + (GNI_{i,j-2} - GNI_{i-1,j-2}) \quad (2)$$

where i denotes a version of the WDI and j denotes the year described by the data. Since my "treatment," a country's distance to the threshold, is measured only annually and I am interested only in the total amount of revision carried out in a single year, I take the sum of all revisions reported in all versions in a given year. To match the data and classification schedule of the World Bank, I aggregate by fiscal year, beginning in July. Intuitively, this is a measure of the total dollars a country added or subtracted to its estimates of its GNI in the three most recent years. Since these are the years that will most likely be affected in a country's attempts to engineer its GNI for the present year, they are a good basis for a proxy for a country's revision activity in a given year. Finally, because this paper highlights reasons why countries may choose to either seek or avoid certain classifications, I remain agnostic about the direction of the revision by taking the absolute value, resulting in the following measure of revision activity:

$$ThreeYrRevisionAbs_j = |\Sigma ThreeYrRevision_{i,j}| \quad (3)$$

With this measure defined, I proceed to investigate whether a country's revision activity increases as it approaches an income classification cutoff. To do so, I make use of the analytic tools developed for regression discontinuity designs, which allow me to test whether data just before and just after a cutoff exhibit significantly different patterns. Regression discontinuities are typically employed in the service of causal identification. In these designs, an author makes use of a situation in which some "treatment" or intervention occurs only after a certain threshold is met. If this threshold is truly arbitrary, and if actors do not behave strategically, then in expectation, units that have just met and have just failed to meet the threshold should be similar. Using this design, a researcher will test whether outcomes for units just above and just below the threshold are significantly different, and if they are, this may be causally attributed to the intervention. My use of this design differs, since I am using it to illustrate that units' behavior *is* strategic.

Specifically, I estimate the size of the jump in revision activity at the discontinuity of the threshold through the use of local linear regression.[115] This approach models linear relationships on either side of the threshold and estimates the difference between them. Since regression discontinuities model the difference only in the neighborhood of the cutpoint, distant observations are omitted. Selection of the bandwidth within which observations are included, therefore, is an important modeling decision. There is some debate in the literature whether bandwidths should be selected through a data-driven process or through researcher discretion; I simply estimate the results at several reasonable bandwidths to demonstrate the sensitivity of the results.[116] In order to cluster standard errors by country, I use block bootstrapping, since there are fewer than 50 clusters within these bandwidths. Using the form

$$Y_{it} = \alpha + \beta AboveCutoff_{it} + \delta DistanceToCutoff_{it} + \gamma AboveCutoff * DistanceToCutoff_{it} + \varepsilon_{it} \quad (4)$$

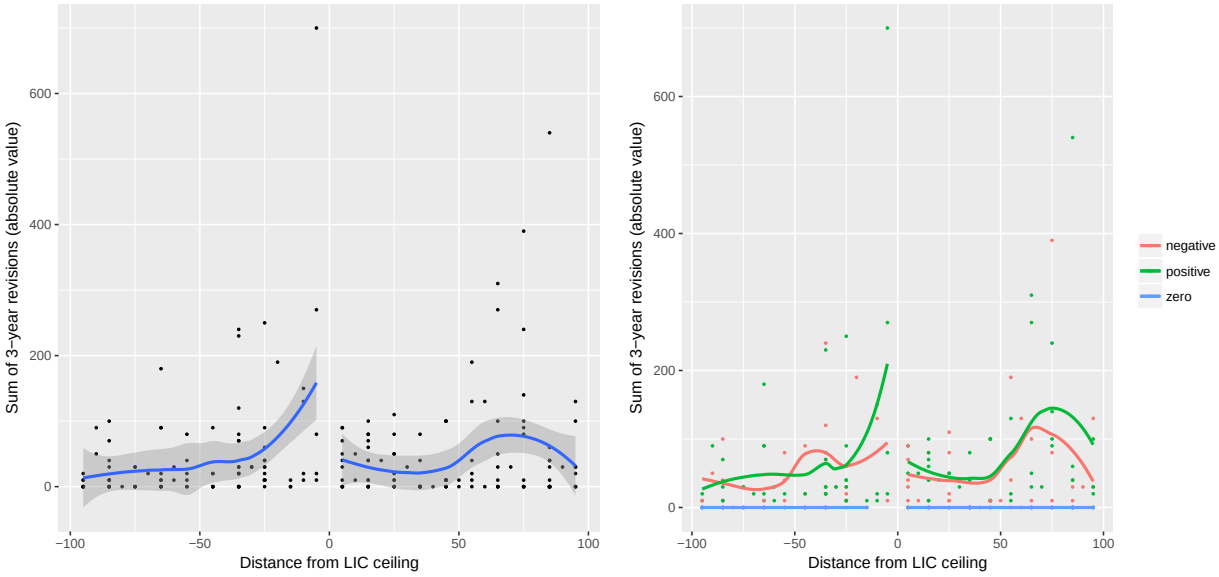
I estimate β, which is equivalent to the size of the jump at the discontinuity.

Revisions to GNI increase as countries approach the LIC ceiling and significantly fall as soon as they have passed the cutoff. This trend is visually apparent in Figure 5. Panel A plots the absolute value of

[115] For methodological reasons to use local linear regression instead of the use of higher-order polynomials, see Skovron and Titiunik (2015).

[116] See Calonico et al. (2014).

Figure 5: Revisions to GNIpc increase as a country approaches the LIC ceiling



Note: Figure plots the magnitude of revisions to recent national income estimates against the distance between a country's most recent current estimate and the LIC ceiling. Revisions to national income data increase as a country approaches this ceiling. The figure plots the raw data, with each point representing a country-year observation. Loess smoothing lines fitted using bandwidths of 100. *Source:* World Development Indicators Archives.

revisions to GNI, positive and negative, while in Panel B, positive and negative revisions to GNI are plotted separately. It is clear that some countries add to their GNI as they near the cutoff, while others subtract from it, suggesting heterogeneous strategies in response to the cutoff. Point estimates appear in Table 5, Panel A. The data-driven Imbens-Kalyanaraman bandwidth is estimated to be 24. This proximity to the threshold is so close as to possibly obviate the need for strategic behavior, so I prefer bandwidths of 50, 100, and 150, which are far enough out for data manipulation to be helpful in making the difference one way or the other. The third specification, using the bandwidth of 100, suggests that countries revise their GNI per capita by about $70 more just before the cutpoint, and the effect is statistically significant at conventional levels. While the other specifications are not statistically significant at conventional levels, with p-values between .1 and .2, the overall results are similarly negative. This is cautious evidence that countries are more likely to revise their GNI just before the LIC cutoff although the results are sensitive to the selection of bandwidth. In contrast, there is clearly no elevated GNI revision just before reaching the ceiling of the LMIC category. Results from Table 5, Panel B are slightly positive but extremely imprecise, suggesting that there is no systematic relationship between data revisions and this cutoff.

This analysis, both qualitative and quantitative, illustrates that international constituencies and national governments recognize the effects of classifications and respond strategically. Notably, strategies do not

Table 5: Discontinuities in revisions to national income data

	Absolute revisions to GNIpc in previous 3 years			
	(1)	(2)	(3)	(4)
	A. LIC/LMIC Discontinuity			
Above cutoff	−205.01	−107.68	−69.89**	−28.47
	(175.55)	(75.46)	(31.49)	(23.39)
Bandwidth	24	50	100	150
Observations	40	114	218	344
	B. LMIC/UMIC Discontinuity			
Above cutoff	16.06	71.52	−0.358	88.34
	(114.97)	(119.64)	(87.62)	(67.60)
Bandwidth	39	50	100	150
Observations	31	34	82	113

Standard errors in parentheses
*p<0.1; **p<0.05; ***p<0.01

Note: Estimates come from a local linear regression of the outcome on the above-cutoff indicator, the running variable (GNIpc distance to the cutoff), and the interaction. Model 1 uses the data-driven IK bandwidth. All standard errors are calculated using block bootstrapping, clustered by country.

appear to be uniform, as some strategies are designed to improve a country's classification, while others designed to hinder it. While the particular outcomes I have discussed in this paper—aid, FDI, creditworthiness, and democracy ratings—are largely important for material reasons, social motivations also surfaced in discussions of graduation. On the topic of Nepal's LDC graduation, one consultant conjectured, "When the country graduates, politicians will get the credit. They are working for their own interests, not the people. Leaders want to travel to foreign countries and receive the international credibility associated with being a developing country, not an LDC."[117] Further exploring heterogeneity in the strategies classified actors pursue and the extent to which they are motivated by the desire for improved status is a topic for future research.

6 Conclusion

Far from simply describing the world, classifications structure it in tangible ways. In this paper, I presented quantitative and qualitative evidence to illustrate that when a country moves up in its development classification, it receives less aid and improved creditworthiness and democracy ratings. While the quantitative evidence did not support a relationship between classification and investor perceptions, this relationship mer-

[117] Author's interview with international NGO consultant, July 23, 2017, Kathmandu, Nepal.

its further exploration given qualitative accounts of graduation processes. The patterns I observe in the data and interviews offer strong support for my cognitive-strategic theory of classifications, indicating multiple mechanisms through which classifications shape global politics. I demonstrated that even highly-informed elite actors appear to be susceptible to the heuristic effects of classifications and that foreign donors consider their need to justify their allocation of scarce resources to a donor public. More broadly, my theory implies that actors in principal-agent relationships and those who are sensitive to cognitive biases will be the most frequent "users" of a classification system.

I also presented preliminary evidence that classified constituencies and countries appear to win and lose from certain classification effects and strategically try to influence their classifications accordingly. These strategies are not uniform, however, and may be shaped not only by material incentives but also by social ones. Further exploration of this heterogeneity has the potential to shed light on economic strategies available to countries who wish to acquire improved status in the global economy. In the meantime, results presented here further substantiate my core claim that classifications are a powerful tool international organizations may deploy, either intentionally or unintentionally, to affect outcomes and behaviors in the international economy.

This raises another question for future research: If classifications are so powerful, how do international organizations wield this power, and to what ends? Although the World Bank income classifications appear to have acquired global influence unintentionally, they may have been preserved strategically. Organizers of the Raise the MIC movement claim that the World Bank has resisted their demands because powerful donors do not wish to expose the extent of global poverty and the inadequacy of their assistance.[118] Indeed, it is possible that classifications are one other way through which powerful countries use international organizations to achieve their goals. Alternatively, organizational theories could account for the choices made in developing and maintaining a classification, as bureaucracies try to extend their relevance and reach. Consistent with this view, the UN CDP released a report encouraging international organizations to make more extensive use of its LDC category.[119] These appeals, if successful, will exaggerate the effects I have described in this paper. Regardless, future work should investigate the strategies used by the classifiers to better understand the evolving role of international organizations in the global economy.

[118] Author's interview with Denys Nazarov, Associate Director of Global Policy, AIDS Healthcare Foundation, April 7, 2017, phone interview.

[119] https://www.un.org/development/desa/dpad/publication/recognition-and-application-of-the-least-developed-country-category-by-un-development-system-organizations/

But to the extent that international organizations hope to use classifications to promote development, this study offers policy implications. When international organizations consider graduating or reverse graduating a country from a particular grouping or classification, they should take into consideration the expected response of external actors when estimating the impact of such a move. Moreover, international organizers should plan for the politicization of any classification they introduce, as my project illustrates that global development institutions, intentionally or unintentionally, affect how developing countries are perceived and treated in the international economy.

References

Yair Aharoni. Behavioral Elements in Foreign Direct Investments. *Advances in International Management*, 23:73–111, 2010.

Alberto Alesina and David Dollar. Who Gives Foreign Aid to Whom and Why? *Journal of Economic Growth*, 5:33–63, 2000.

Jaakko Aspara and Henrikki Tikkanen. The Role of Company Affect in Stock Investments: Towards Blind, Undemanding, Non-Comparative and Committed Love. *Journal of Behavioral Finance*, 11(2), 2010.

Severine Autesserre. *Peaceland: Conflict Resolution and the Everyday Politics of International Intervention*. Cambridge University Press, 2014.

Andy Baker. Race, Paternalism, and Foreign Aid: Evidence from U.S. Public Opinion. *American Political Science Review*, 109(1), 2015.

Michael Barnett and Martha Finnemore. The Politics, Power, and Pathologies of International Organizations. *International Organization*, 53(4):699–732, 1999.

Michael Barnett and Martha Finnemore. *Rules for the World*. Cornell University Press, 2004.

Michael M. Bechtel, Jens Hainmueller, and Yotam Margalit. Preferences for International Redistribution: The Divide over the Eurozone Bailouts. *American Journal of Political Science*, 58(4):835–856, 2014.

Shannon Lindsey Blanton and Robert Blanton. What Attracts Foreign Investors? An Examination of Human Rights and Foreign Direct Investment. *The Journal of Politics*, 69(1):143–155, 2007.

Mark Blyth. *Great Transformations: Economic Ideas and Institutional Change in the Twentieth Century*. Cambridge University Press, 2002.

Sarah M. Brooks, Raphael Cunha, and Layna Mosley. Categories, Creditworthiness, and Contagion: How Investors' Shortcuts Affect Sovereign Debt Markets. *International Studies Quarterly*, 59(3):587–601, 2015.

Mark. T. Buntaine, Bradley C. Parks, and Benjamin P. Buch. Aiming at the Wrong Targets: The Domestic Consequences of International Efforts to Build Institutions. *International Studies Quarterly*, 2017.

Sarah Bush. The Politics of Rating Freedom: The Role of Ideological Affinity in the Puzzling Authority of the Freedom in the World Ratings. *Perspectives on Politics*, 2017.

Tim Büthe. Beyond Supply and Demand: A Political-Economic Conceptual Model. In *Governance by Indicators: Global Power through Quantification and Rankings*. Oxford University Press, 2012.

Tim Büthe and Helen Milner. The Politics of Foreign Direct Investment into Developing Countries: Increasing FDI through International Trade Agreements? *American Journal of Political Science*, 52(04):741 – 762, 2008.

Sebastian Calonico, Matias D. Cattaneo, and Rocio Titiunik. Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. *Econometrica*, 82(6):2295–2326, 2014.

Allison Carnegie and Cyrus Samii. International Institutions and Political Liberalization: Evidence from the World Bank Loans Program. *British Journal of Political Science*, (Forthcoming), 2017.

Edward Hallett Carr. *The Twenty Years' Crisis, 1919-1939*. Harper Perennial, 1964.

Terrence L. Chapman. Audience Beliefs and International Organization Legitimacy. *International Organization*, 63(X):733, 2009.

Jeffrey Chwieroth. *Capital Ideas: The IMF and the Rise of Financial Liberalization*. Princeton University Press, 2009.

Michael A. Clemens and Michael Kremer. The New Role for the World Bank. *Journal of Economic Perspectives*, 30(1):53–76, 2016.

John Conlisk. Why Bounded Rationality? *Journal of Economic Literature*, 34(2):669–700, 1996.

Alexander Cooley and Jack Snyder, editors. *Ranking the World: Grading States as a Tool of Global Governance*. Cambridge University Press, 2015.

Allan Dafoe, Jonathan Renshon, and Paul Huth. Reputation and Status as Motives for War. *Annual Review of Political Science*, 17(1):371–393, 2014.

Xinyuan Dai. *International Institutions and National Policies*. Cambridge University Press, Cambridge, 2007.

Kevin E. Davis, Angelina Fisher, Benedict Kingsbury, and Sally Engel Merry, editors. *Governance by Indicators: Global Power through Quantifications and Rankings*. Oxford University Press, 2012.

Lindsay Dolan. The Distortionary Effects of Classifications on Aid Allocation Decisions: Evidence from an Elite Lab Experiment. 2017.

Axel Dreher, Silvia Marchesi, and James Raymond Vreeland. The Political Economy of IMF Forecasts. *Public Choice*, 137(1-2):145–171, 2008.

Axel Dreher, Jan Egbert Sturm, and James Raymond Vreeland. Development Aid and International Politics: Does Membership on the UN Security Council Influence World Bank Decisions? *Journal of Development Economics*, 88(1):1–18, 2009.

Neil Fantom and Umar Serajuddin. The World Bank's Classification of Countries by Income. *World Bank Policy Research Working Paper*, 7528, 2016.

James D. Fearon. Domestic Political Audiences and the Escalation of International Disputes. *The American Political Science Review*, 88(3):577–592, 1994.

Morris P. Fiorina. Legislative Choice of Regulatory Forms - Legal Process or Administrative Process. *Public Choice*, 39(1):33–66, 1982.

M. Flinders and J. Buller. Depoliticisation: Principles, Tactics and Tools. *British Politics*, 1:293–318, 2006.

Emmanuel Frot and Javier Santiso. Herding in Aid Allocation. *Kyklos*, 64(1):54–74, 2011.

Julia Gray. *The Company States Keep: International Organizations and Investor Perceptions*. Cambridge University Press, 2013.

Julia Gray and Raymond Hicks. Reputations, Perceptions, and International Economic Agreements. *International Interactions*, 40(3):325–349, 2014.

Jessica Green. *Rethinking Private Authority: Agents and Entrepreneurs in Environmental Governance*. Princeton University Press, 2013.

Anthony G. Greenwald and Mahzarin R. Banaji. Implicit Social Cognition: Attitude, Self, and Stereotypes. *Psychological Review*, 102(1):4–27, 1995.

Tobias Heinrich, Yoshiharu Kobayashi, and Kristin A. Bryant. Public Opinion and Foreign Aid Cuts in Economic Crises. *World Development*, 77:66–79, 2016.

James R. Hollyer, B. Peter Rosendorff, and James Raymond Vreeland. Democracy and Transparency. *The Journal of Politics*, 73(04):1191–1205, 2011.

Nathan M. Jensen. *Nation-States and the Multinational Corporation: A Political Economy of Foreign Direct Investment*. Princeton University Press, 2006.

Morten Jerven. *Poor Numbers: How We Are Misled by African Development Statistics*. Cornell University Press, Ithaca, NY, 2013.

Robert Jervis. Reports, Politics, and Intelligence Failures: The Case of Iraq. *Journal of Strategic Studies*, 29(1):3–52, 2006.

Daniel Kahneman. *Thinking Fast and Slow*. Farrar, Straus, and Giroux, 2011.

Ashwin Kaja and Eric Werker. Corporate Governance at the World Bank and the Dilemma of Global Governance. *World Bank Economic Review*, 24(2):171–198, 2010.

Judith Kelley. *Scorecard Diplomacy: Rating Nations on the Fight against Human Trafficking*. Book manuscript, 2016.

Judith G. Kelley and Beth A. Simmons. Politics by Number: Indicators as Social Pressure in International Relations. *American Journal of Political Science*, 59(1):55–70, 2015.

Robert O. Keohane. *After Hegemony: Cooperation and Discord in the World Political Economy*. Princeton University Press, 1984.

Andrew Kerner, Morten Jerven, and Alison Beatty. Does It Pay to Be Poor? Testing for Systematically Underreported GNI Estimates. *The Review of International Organizations*, pages 1–38, 2015.

Erasmus K. Kersting and Christopher Kilby. With a Little Help from My Friends: Global Electioneering and World Bank Lending. *Journal of Development Economics*, 121:153–165, 2016.

Charles Kindleberger. *Maniacs, Panics, and Crashes: A History of Financial Crises*. John Wiley and Sons, Inc., 5th edition, 2005.

Stephen Knack, Colin Xu, and Ben Zou. Interactions among Donors' Aid Allocations: Evidence from an Exogenous World Bank Income Threshold. *World Bank Policy Research Working Paper*, 7039, 2014.

Barbara Koremenos, Charles Lipson, and Duncan Snidal. The Rational Design of International Institutions. *International Organization*, 55(4):761–799, 2001.

Stephen Kosack and Jennifer Tobin. Funding Self-Sustaining Development: The Role of Aid, FDI, and Government in Economic Success. *International Organization*, 60(Winter), 2006.

Ilyana Kuziemko and Eric Werker. How Much Is a Seat on the Security Council Worth? Foreign Aid and Bribery at the United Nations. *Journal of Political Economy*, 114(5):905–930, 2006.

Claudia Landwehr and Katharina Böhm. Delegation and Institutional Design in Health-Care Rationing. *Governance*, 24(4):665–688, 2011.

Valentin F. Lang and Andrea F. Presbitero. Room for Discretion? Biased Decision Making in International Financial Institutions. 2017.

Quan Li and Adam Resnick. Reversal of Fortunes: Democratic Institutions and Foreign Direct Investment Inflows to Developing Countries. *International Organization*, 57(1):175–211, 2003. ISSN 0020-8183. doi: 10.1017/S002081.

David Halloran Lumsdaine. *Moral Vision in International Politics: The Foreign Aid Regime, 1949-1989*. Princeton University Press, 1993.

Donald MacKenzie. *An Engine, Not a Camera: How Financial Models Shape Markets*. MIT Press, Cambridge, 2006.

Christopher S. P. Magee and John A. Doces. Reconsidering Regime Type and Growth: Lies, Dictatorships, and Statistics. *International Studies Quarterly*, 59(2):223–237, 2015.

Elizabeth Maitland and André Sammartino. Material Cognition and Internationalization. *Journal of International Business Studies*, 46(7):733–760, 2015.

Walter Mattli and Tim Büthe. Accountability in Accounting? The Politics of Private Rule-Making in the Public Interest. *Governance*, 18(3):399–429, 2005.

Walter Mattli and Ngaire Woods. In Whose Benefit? Explaining Regulatory Change in Global Politics. In Walter Mattli and Ngaire Woods, editors, *The Politics of Global Regulation*, pages 1–43. Princeton University Press, Princeton, 2009.

John J. Mearsheimer. The False Promise of International Institutions. *International Security*, 19(3):5–49, 1994.

Helen Milner. Why Multilateralism? Foreign Aid and Domestic Principal-Agent Problems. In Darren Hawkins, David Lake, Daniel Nielsen, and Michael J. Tierney, editors, *Delegation and Agency in International Organizations*. Cambridge University Press, 2006.

Scott Morris and Madeleine Gleave. The World Bank at 75. *Center for Global Development Policy Paper*, (58), 2015.

Julia C. Morse. International Institutions and Indirect Market Enforcement: Analyzing the Regime to Combat Terrorist Financing. 2017.

Todd Moss and Benjamin Leo. IDA at 65: Heading Toward Retirement or a Fragile Lease on Life? *Center for Global Development Working Paper*, 246, 2011.

Ashley Napier, Angie Acquatella, Michael G. Findley, Daniel L. Nielson, Brian O'Donnell, Michael J. Tierney, and Catherine E. Weaver. Mapping for Impact: Testing the Effects of Geospatial Information on Development Aid Decision-Making. 2016.

Eric Neumayer. Is the Allocation of Food Aid Free from Donor Interest Bias? *Journal of Development Studies*, 41(3):394–411, 2005.

Benjamin J. Newman, Todd K. Hartman, Patrick L. Lown, and Stanley Feldman. Easing the Heavy Hand: Humanitarian Concern, Empathy, and Opinion on Immigration. *British Journal of Political Science*, 45 (03):583–607, 2013.

Farhad Noorbakhsh, Alberto Paloni, and Ali Youssef. Human Capital and FDI Inflows to Developing Countries: New Empirical Evidence. *World Development*, 29(9):1593–1610, 2001.

Jennifer M. Oetzel, Richard A. Bettis, and Marc Zenner. Country Risk Measures: How Risky Are They? *Journal of World Business*, 36(2):128–145, 2001.

Pamela Paxton and Stephen Knack. Individual and Country-Level Factors Affecting Support for Foreign Aid. *International Political Science Review*, 33(2):171–192, 2012.

Dilip Ratha, Prabal K. De, and Sanket Mohapatra. Shadow Sovereign Ratings for Unrated Developing Countries. *World Development*, 39(3):295–307, 2011.

Martin Ravallion. The World Bank: Why It Is Still Needed and Why It Still Disappoints. *Journal of Economic Perspectives*, 30(1):77–94, 2016.

Dani Rodrik. When Ideas Trump Interests: Preferences, World Views, and Policy Innovations. 2013.

Justin Sandefur and Amanda Glassman. The Political Economy of Bad Data: Evidence from African Survey and Administrative Statistics. *The Journal of Development Studies*, 51(2):116–132, 2013.

James C. Scott. *Seeing Like a State*. Yale University Press, New Haven, 1998.

Robert J. Shiller. Narrative Economics. 2017.

Herbert A. Simon. A Behavioral Model of Rational Choice. *Quarterly Journal of Economics*, 69(1):99–118, 1955.

Timothy Sinclair. *The New Masters of Capital*. Cornell University Press, Ithaca, NY, 2005.

Christopher Skovron and Rocio Titiunik. A Practical Guide to Regression Discontinuity Designs in Political Science. 2015.

John D. Steinbruner. *The Cybernetic Theory of Decision: New Dimensions of Political Analysis*. Princeton University Press, 1974.

Henri Tajfel and A. L. Wilkes. Classification and Quantitative Judgement. *British Journal of Psychology*, 54(2):101–114, 1963.

Shelley E. Taylor. A Categorization Approach to Stereotyping. In David L. Hamilton, editor, *Cognitive Processes in Stereotyping and Intergroup Behavior*. Psychology Press, 1981.

Amos Tversky and Daniel Kahneman. Judgment under Uncertainty: Heuristics and Biases. *Science*, 185 (4157):1124–1131, 1974.

UNCTAD. *World Investment Report 2017*. 2017.

Madhu Vij. The Determinants of Country Risk Analysis: An Empirical Approach. *Journal of Management Research*, 5(1):20–31, 2005.

Jeremy L. Wallace. Juking the Stats? Authoritarian Information Problems in China. *British Journal of Political Science*, 46(01):11–29, 2016.

Kenneth N. Waltz. *Theory of International Politics*. Waveland Press, 1979.

R. Kent Weaver. The Politics of Blame Avoidance. *Journal of Public Policy*, 6(4):371–398, 1986.

Alexander Wendt. Anarchy is What States Make of It: The Social Construction of Power Politics. *International Organization*, 46(2):391–425, 1992.

David A. Wilder. Social Categorization: Implications for Creation and Reduction of Intergroup Bias. *Advances in Experimental Social Psychology*, 19(C):291–355, 1986.

David W. Williams and Denis A. Grégoire. Seeking Commonalities or Avoiding Differences? Re-Conceptualizing Distance and Its Effects on Internationalization Decisions. *Journal of International Business Studies*, 46(3), 2015.