Political Science Research and Methods (2025), 1–21
doi:10.1017/psrm.2025.19

ORIGINAL ARTICLE

Reassessing World Bank conditionality: beyond count measures

Jacob Winter[1], Ben Cormier[2], Teresa Kramarz[3] and Mark S. Manger[4]

[1]Political Science, University of Toronto, Toronto, ON, Canada; [2]Government and Public Policy, University of Strathclyde, Glasgow, UK; [3]School of the Environment, University of Toronto, Toronto, ON, Canada and [4]Munk School of Global Affairs and Public Policy, University of Toronto, Toronto, ON, Canada
Corresponding author: Mark S. Manger; Email: mark.manger@utoronto.ca

(Received 22 April 2024; revised 3 July 2024; accepted 2 September 2024)

Abstract
Many studies argue that the World Bank grants favorable loan conditions to allies of its powerful principals. These studies typically use the count of conditions as a proxy for how demanding loans are on borrowers, even though some conditions are more difficult to comply with than others. We propose a new operationalization: a measure of conditionality stringency in Bank loans constructed using Latent Semantic Scaling. Using this new measure, we find little evidence of a generalizable influence of powerful principals. Instead, the stringency of loan conditions is associated with bureaucratic assessments of risk. To facilitate future research, we provide a new dataset of World Bank loan condition texts and our measure of text stringency for all loans in the dataset.

Keywords: conditionality; World Bank; text and content analysis

1. Introduction

Not all World Bank loan conditions are created equal. Why do some borrowers receive relatively lenient conditions while others receive more stringent conditions? This question relates to a foundational debate in International Relations: do international organizations (IOs) primarily do the bidding of their most powerful member states or do they have a life of their own as organizations (Barnett and Finnemore, 1999)? Applied to the Bank, are loan conditions a function of powerful principals or bureaucratic tendencies?

A large literature finds powerful member states influence World Bank conditionality. Prominent studies, drawing on principal-agent (PA) theory, rely on the *count of loan conditions* to proxy how onerous loans are for borrowers (inter alia, Kilby, 2009; Clark and Dolan, 2021) and show that states aligned with powerful Bank principals in forums such as the UN receive fewer conditions. This study finds that important aspects of Bank conditionality other than count are not determined by powerful principals.

First, reviews of conditionality by Bank researchers submit that what is most impactful on a borrower is not the number of loan conditions, but their *stringency*—the degree to which loans require verifiable and costly reforms rather than vaguely defined changes of little consequence (Kapur and Webb, 2000; Koeberle, 2005b). Studies of policy loans find that success depends on "substantive […] policy- and action-oriented" conditions or prior actions requiring "actionable" reforms

that can "lead to tangible results" (Moll et al., 2015, 9, 14). We define stringency based on this research, then operationalize and measure stringency by applying quantitative text analysis to a new dataset of Bank loan conditions. Second, using this new measure, we find that conditionality language is primarily determined by the Bank's bureaucratic priorities, not the influence of powerful principals.

This yields two contributions. First, we document that across time and borrowers, the stringency of Bank conditionality is primarily associated with attributes of bureaucratic organizations. Despite the influence of principals, Bank staff obtain enough agency slack that these tendencies inform conditionality content more than principal interventions. This is new quantitative evidence supporting theories of Bank lending that draw on sociological insights (Barnett and Finnemore, 2004; Gutner, 2005; Weaver, 2008; Honig, 2019; Kramarz, 2020) and theories of informal governance by staff across international financial institutions (on the International Monetary Fund [IMF], see Chwieroth, 2013; Nelson, 2017).

We do *not* claim that principals have no influence on loan conditions. Rather, we refine studies of Bank lending by highlighting that (1) the language of conditionality text is at least as important as, if not more important than, the number of conditions and (2) variation in conditionality language is not associated with the same factors as variation in conditionality count. This does not rule out that principals will occasionally try to influence the content of conditions that borrowing countries receive, but indicates that such influence may be too rare, muted, or mediated by the Bank's own bureaucracy to be discernible in standard statistical models. Moreover, when we replace our text-dependent variable with the count of conditions, our models replicate the common finding that powerful principals affect condition count. This suggests that theories of principal influence and theories of a bureaucratic Bank are complementary. Future research can study both aspects of conditionality and account for the empirical patterns we identify.

Second, to promote research on conditionality language, we provide a new dataset of 3,641 publicly available World Bank loan agreements from 1995 to 2021. Studies frequently rely on a dataset that began in 2004, the World Bank's Development Policy Action Database (Clark and Dolan, 2021, 40). We thus add over a decade of empirical material and a new measure of conditionality stringency across all loans for future research. We hope that these data incentivize further analysis of conditionality content.

2. World Bank conditionality: theories and expectations

When explaining World Bank lending, many find evidence that loans reflect the preferences of powerful member states. The Bank, as a development agency, claims that it adjusts loan conditions based on borrowers' institutional capacities and past loan performance. Finally, others emphasize how the Bank's bureaucratic organization informs its lending.

2.1. Powerful principals and World Bank lending

A sizable literature positions IOs as agents of state principals. Principals impose constraints on agents to ensure they do not shirk mandates (Nielson and Tierney, 2003; Hawkins and Jacoby, 2006) and serve principals' interests (Lall, 2017). Many highlight that principal influence depends on their number and homogeneity (Copelovitch, 2010; Stone, 2011). Single-principal frameworks argue that among Bank borrowers, "friends" of the US receive fewer and more-forgiving loan conditions (Kilby, 2009; Clark and Dolan, 2021). Collective-principal frameworks find similar constraints on Bank lending because coalitions of principals gain influence through Bank voting rules (Lyne et al., 2006, 2009). This leads to the proposition that countries with foreign policy affinity to the US or a coalition of powerful principals obtain fewer and less-stringent loan conditions.

2.2. Bank evaluation: policy, governance, and performance

Although principals try to ensure IOs follow their preferences, IOs inevitably gain some agency (Hawkins and Jacoby, 2006). What do IOs do with their autonomy (Gutner, 2005; Kramarz, 2020)? The Bank claims that it lends according to borrowers' institutional quality and past performance. This claim has historical roots. The first generation of structural adjustment conditionality was designed to help borrowers with global market integration (Rodrik, 2006, 973; Williamson, 1990). In response to criticism, a second generation of conditionality emerged in the early 2000s. Reforms were now also supposed to focus on institutions and governance, ostensibly ensuring that the context for successful implementation of first-generation reforms was present (Koeberle, 2005b; Rodrik, 2006).

One desired consequence was more loans to countries with pre-existing institutional capacity. Since "more conditionality cannot compensate for weak government commitment or implementation capacity […] *selectivity in favor of countries with favorable policy environments*" became good practice, according to Bank insiders (Koeberle, 2005b, 4, emphasis added), lending was also based on past performance: "A country's track record is a better indicator of its determination and effectiveness in implementing a viable development strategy than elaborate promises for future efforts […] *conditionality for good performers can be less prescriptive*" (Koeberle, 2005b, 4, emphasis added).

These reviews of conditionality suggest that it should be less stringent where institutions and past performance are strong, reflecting Bank efforts to emphasize borrower ownership (Best, 2014). But whether this occurs in practice is disputed: Buntaine et al. (2017) find that in the natural resources and environmental management sector, countries with relatively weaker institutions tend to receive comparatively "easy and shallow," "form over function" targets. The effect of institutions on Bank loan conditions is thus an open question.

2.3. The World Bank as a bureaucracy

A third approach theorizes that IO behavior can be understood through the lens of specific bureaucratic traits (Barnett and Finnemore, 1999). Exemplifying Weber's (1946) insights, IO bureaucracies shape internal decision-making by creating incentives that guide staff toward standardized behaviors and procedures (Weaver, 2008). As a mature bureaucracy, the Bank values predictability and measurability of outcomes (Honig, 2018). If staff follow routines to pursue organization-wide targets, they reduce uncertainty (Eckhard and Parizek, 2022), if at the cost of development effectiveness (Honig, 2018, 2019).

The most important here is that such incentives lead to patterns in organizational outputs (such as loan conditions) despite staff having different preferences and working in different contexts (Park and Kramarz, 2019). For example, it is widely accepted that Bank staff have career incentives to extend loans (Weaver, 2007; Safarty, 2012; Briggs, 2021). At the same time, the Bank seeks to reduce the risk of loan failure. With policy loans specifically, failure arises if borrowers do not comply with conditions and funds are not used for the agreed goals. Bank emphasis on using conditionality to reduce this risk is codified in post-loan monitoring by the Bank's internal Independent Evaluation Group (IEG). "Unsatisfactory" IEG scores indicating that loan failures have been linked to "weaker" prior actions and conditions that are not "substantive, viz. policy- and action-oriented […] that lack policy substance and are less action-oriented" (Moll et al., 2015, 9). Successful policy loans, then, depend on "having reforms or policy measures that are actionable and that can indeed lead to tangible results" (Moll et al., 2015, 14). Conditionality is thus a core tool for mitigating organization-wide loan risk while creating space for staff to meet the disbursement imperative.

If Weberian bureaucracy informs conditionality, then conditions should at least in part vary by how "risky" a loan is—that is, how likely the borrower is to not comply with loan aims or use funds in ways that do not meet the Bank's development agenda. For example, borrowers who disagree with

the Bank's economic development prescriptions—and thus pose a greater threat of non-compliance—may face stricter conditionality. Similarly, funds for policy budget support loans are more fungible than for project loans, so may require stricter conditionality. Such links between perceived loan risk and variation in Bank loans have been identified elsewhere, if with different dependent variables. Winters (2010) finds that, given concern about the use of fungible funds, aid recipients with a poor governance record are more likely to receive less-fungible project loans than they are to receive budget support policy loans.

3. World Bank lending processes and effects on conditionality

We expect that when the *stringency of Bank condition text* is in question, bureaucratic theory explains Bank lending. In practice, loan agreements result from project cycles, which are organization-wide bureaucratic processes. Of course, it is not credible to claim that principal constraints are entirely absent. However, with sufficient agent slack, principal constraints may have a relatively small impact on conditionality content compared to the bureaucratic organization Bank staff face in their day-to-day work. Principals may also use their veto power only selectively, leaving bureaucratic factors to have a more common, generalizable effect on loan details than external politics (McLean and Schneider, 2014).

Bank loans are the output of a bureaucratic cycle in which projects are identified, appraised, negotiated, approved, implemented, and then evaluated. These cycles are managed by Task Team Leaders (TTLs), who have strong career incentives to make as many loans as possible (Briggs, 2021; World Bank IEG., 2016, 28). TTLs work with recipient governments through a lengthy preparation phase including economic, social, and environmental assessments. Before final approval, projects go through an appraisal stage, in which the Bank and borrower negotiate outcomes, timelines, and evaluation tools, including conditions (World Bank, 2022). Once agreed, both parties sign a loan agreement. Many studies highlight mechanisms through which powerful Bank principals might influence this process (Kilby, 2009). For example, "friends" of the US may have loans expedited, "losing conditions in the process" through a "pleasing" mechanism, where staff either seek to demonstrate that their work supports the US or unconsciously share an American worldview (Clark and Dolan, 2021, 37).

But the details of conditions are not typically a topic of executive board scrutiny. Principals only intervene in important cases (McLean and Schneider, 2014). The Bank's biweekly board meetings provide more evidence of this. Most loans are *not* flagged for discussion but get approved on an "absence of objection basis" or "authorized to proceed on a streamlined basis." As such, the majority of loans—let alone the details of conditionality language—are not the subject of the board meeting discussion.[1] To obtain such easy passes at the board level, staff have incentives to conform to standard operating procedures by submitting uncontroversial and predictable loan conditions (Weaver, 2008; Honig, 2018).

From this, we expect a generalizable bureaucratic effect on the language of conditionality. To be sure, criticism of a lengthy loan preparation process (Ferranti, 2007; Park and Vetterlein, 2010), concerns about over-burdening government agendas (Smets and Knack, 2014), and perhaps competition with Chinese flows (Hernandez, 2017) led the Bank to explicitly reduce the target number of conditions in a loan from twelve to eight in 2012 (Swaroop, 2016, 6). From a scholarly perspective, an organization-wide mandate to reduce condition count in all loans already calls into question using count measures to compare how onerous different loans are. From a practical perspective, the

[1] As an example, publicly available minutes of executive board meetings reveal that, at a 2017 board meeting, over two dozen loans were approved on this passive "absence of objection" basis, and none were actually discussed during the meeting (World Bank, 2017).

persistence of conditionality's central role in Bank lending despite concerns about negative effects highlights how conditionality remains a core tool for risk mitigation while getting money out the door.

We focus on bureaucratic theory and studies of Bank operations to understand why loan language would be associated with factors other than powerful principals' interests. To be sure, borrower preferences are also important to loan outcomes (Cormier, 2024), as are factors that may empower borrowers during negotiations—for example, access to other development lenders such as China (Hernandez, 2017). To preview findings, we find mixed evidence of significant relationships in these areas. This does not mean that borrower interests do not affect final conditionality text in practice and evidence of such influence has been shown in certain instances (Gould and Winters, 2007, 12). But as with selective principal influence over text details, such effects may be too inconsistent or moderated by Bank staff to be identifiable in quantitative models.

4. The stringency of World Bank conditionality

To test our theory regarding loan conditionality text, we first define stringency in loan agreement texts. We then illustrate the variation in stringency and show that it is *not* correlated with the number of conditions in a loan.

4.1. Stringency

Loan conditions are part of the legal terms of a Bank loan, negotiated between the Bank and borrower, that a borrower must fulfill to receive the entirety of funds. The length and complexity of a single condition can vary significantly: one demanding condition may require more adjustment than ten less-demanding conditions (Koeberle, 2005b, 62–66).

To define *stringency,* we rely on analyses of conditionality language by Bank researchers and advisors. Some define stringency as the degree of "prescription" in a condition. Since the 1980s, the Bank has mixed highly demanding and specific "reform" measures alongside "nonbinding or [merely] desired" actions (Koeberle, 2005b, 64–66). The former involves "measurable, objective indicators such as spending allocations" while the latter is "inherently subjective [as it] involves *judging* the relevance and effectiveness of a country's policy choices" in response to a loan condition (Koeberle, 2005b, 68 emphasis added).

Other studies of the Bank make similar distinctions between quantitative targets that are objectively identifiable and vague conditions that require subjective assessments about implementation (Kapur and Webb, 2000, 4). Less-prescriptive language includes governments merely needing to "*assess, authorize, build upon, complete, continue, define, ensure, expand, establish, examine, fill, introduce, improve, increase commitment to, mobilize, organize, prepare, pursue, redefine, reform, streamline, strengthen, study, support, update, upgrade*" or otherwise achieve some difficult-to-quantify aim (Kapur and Webb, 2000, 4). Such language "suffer[s] from definitional, strategic and operational ambiguities" where "a government may promise to do a, b, c, etc., but the consequences of not doing so are unclear […] are these conditionalities or simply banalities" (Kapur and Webb, 2000, 4, 8–10)? A variety of behaviors could be seen as meeting these conditions, increasing the space for borrowers to comply without consequential adjustment. In contrast, more-stringent conditions include precise and objective language that requires verifiable and consequential reform (for example: enacting a law, reallocating budget resources, or privatizing a sector).

These language distinctions remain relevant in later studies: "prior actions or 'conditions' agreed upon with the country authorities that are substantive, viz. policy- and action-oriented, are more likely to attain their expected goals […] prior actions that lack policy substance and are less action oriented are less likely to attain their goals. For example, *focusing on process-oriented steps, such as setting up 'task forces,' doing studies, and issuing 'plans' for the future, may be less likely to attain tangible*

results" (Moll et al., 2015, 9, emphasis added). Others give an example of how merely "establishing regional water boards" is a relatively easy-to-meet target, "while reducing the amount of water lost during distribution" is a quantifiable and thus more demanding target (Buntaine et al., 2017, 472).[2]

Across assessments by Bank insiders and scholars, then, there is consensus that not all conditionality is equally stringent. More-stringent conditions prescribe quantifiable, objectively identifiable, and consequential policy reforms. Less-stringent conditions prescribe unquantifiable, subjectively assessed aims that can be met with a variety of behaviors, reducing how consequential the condition is.

4.2. Condition count versus content

Examples of conditionality text illustrate (1) variation in stringency and (2) that the number of conditions does not necessarily capture the degree of adjustment required of a borrower.

For example, a 2017 loan to Serbia (P149751) included a single, yet stringent, condition prescribing a quantified privatization reform:

- *The borrower, through its privatization agency, issued public announcements for at least twenty (20) public bids for PA [privatization agency] companies that were not in restructuring as of August 13, 2014.*

Decades before and in a different context, a 1994 loan to Burkina Faso (P035593) included only three conditions. Yet each required objectively identifiable adjustments with far-reaching and likely unpopular consequences:

- *Establish a national price list for essential and generic drugs.*
- *Abolish the system of margin controls for all drugs.*
- *Abolish price controls established after Jan. 12, 1994, except for petroleum products, utilities, and school supplies.*

Conversely, the count of conditions may be inflated with comparatively vague prescriptions. A 2014 Argentina loan (P083982) included over 60 conditions,[3] but many did not specify objective actions with quantifiable adjustments:

- *Progress in strengthening the regulatory framework in the public service and infrastructure sectors in a manner that is consistent with the renegotiated concessions.*
- *Increasing policy coordination at the executive level by designating the sub-secretariat of trade policy and management as a coordinating entity for foreign trade policies.*
- *Increasing transparency in the judicial branch through the promulgation of a decree in June 2003 modifying the system of designating judges to the Supreme Court by allowing civil society to opine about the pre-selected candidates.*

A variety of behaviors could be seen as meeting these conditions—the types of conditions that the studies cited above suggest are likely to be inconsequential. In short, neither a direct reading of conditions nor the Bank's own studies suggest that the *number* of conditions indicates loan stringency as much as the text.

[2]To validate our measure, in the online appendix, we apply our scaling measure to their data and find substantial alignment.
[3]Notably, *after* the 2012 Bank mandate to reduce conditionality count.

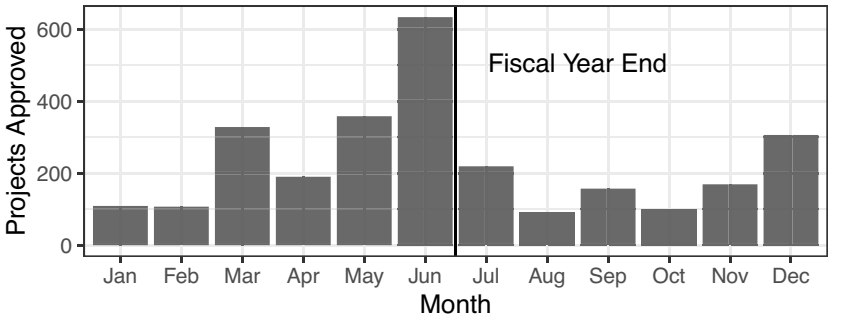

Figure 1. Project approvals over the fiscal cycle.

5. Empirical approach

Even before statistical analysis, *prima facie* evidence of lending patterns is hard to square with a powerful principal explanation rather than bureaucratic tendencies. Considering project approval timing, the World Bank's fiscal year ends on June 30. A common feature of organizations is that budgets need to be exhausted lest they be cut in the next funding round (Vaubel, 2006, 131–32). We see a sharp increase in the number of Bank loans approved in the last quarter of the fiscal year, and a noticeable drop-off at the beginning of a new fiscal year (Figure 1). In the following, we explore this suggestive relationship further.

5.1. Operationalizing stringency: Latent Semantic Scaling

To operationalize conditionality stringency, we use a second-generation text scaling model. These models place a text along a continuum, allowing researchers to compare texts across user-defined dimensions. The most recent class of models, including Latent Semantic Scaling (LSS) (Watanabe, 2021; Trubowitz and Watanabe 2021), combines the flexibility of unsupervised models with the benefit of pre-specifying theoretically relevant keywords (Gallagher et al., 2017; Eshima et al., 2020). In the first (unsupervised) stage of LSS, we train a set of word embeddings. This encodes information about the context in which words are used. In the second (supervised) stage, we measure the distance of each term to a set of hand-coded "seed words." We then use the relative distance, called the "polarity" of terms, to calculate document stringency along a continuum. See the online appendix for further details.

5.2. Text corpus

We extract conditionality texts from World Bank loan agreements, scraped using the Bank documents' Application Programming Interface (API). The Bank classifies loans as investment project, policy, or program-for-results financing. Our analysis includes all of these, but to align with the existing literature, we also show separate results for policy lending only. Conditionality text in these documents uses a common format and terminology, allowing us to extract conditionality content from 3,641 loans going back to 1995, distributed across the Bank's borrowers (Figure 2).

We draw seed words from frequently occurring terms in condition text, manually coding them as either *stringent* (1) or *lenient* (−1). Seed words were identified and coded based on the Bank's own research. As detailed above, this body of research argues that conditionality is less stringent when a variety of behaviors could be interpreted as meeting the condition, making them easier to satisfy and thus less consequential. The complete seed word list is in the online appendix, with examples in Table 1. Terms like "facilitate" or "taking steps" exemplify conditions that can be satisfied without consequential adjustments. In contrast, stringent seed words were coded as such because they indicate

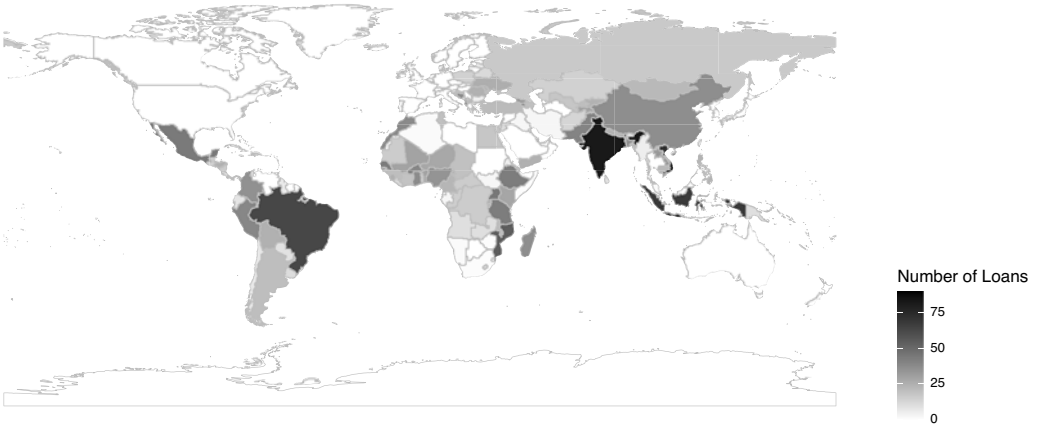

Figure 2. Loans per country, 1995-2021.

Table 1. Example seed words

Lenient (−1)	Stringent (1)
facilitate	audit
take_steps	implement
foster	allocate
enhance	enact
guidelines	budget
recommend	privatize
study	remove
committee	law
promote	tender

that it is possible to make objective assessments about whether reforms were implemented or not, with clearer consequences than lenient terms.

Each non-seed term's polarity is measured by the proximity between its embedding and the set of seed words. Proximity is measured as the cosine similarity between the non-seed and the seed word embeddings. These similarities are multiplied by the polarity of each term and then averaged. Terms used in similar contexts as stringent (lenient) seed words will have a positive (negative) polarity score. Of course, most terms communicate nothing about stringency—so have scores near zero. Figure 3 shows the polarity of several frequently occurring terms, with the words identified as most-polarizing in bold.

5.3. Document scores

We then use the polarity of terms in loan conditions to measure overall loan stringency. We scale each sentence and average the term polarity of each sentence to identify each condition's stringency. Table 2 shows an example sentence from each quintile of polarity. We then sum the polarity of all sentences in loan condition sections to identify the loan's overall stringency. This is our dependent variable. In the appendix, we use alternative strategies for the stringency measure to ensure that our results do not depend on any particular choice about how to aggregate condition polarity.

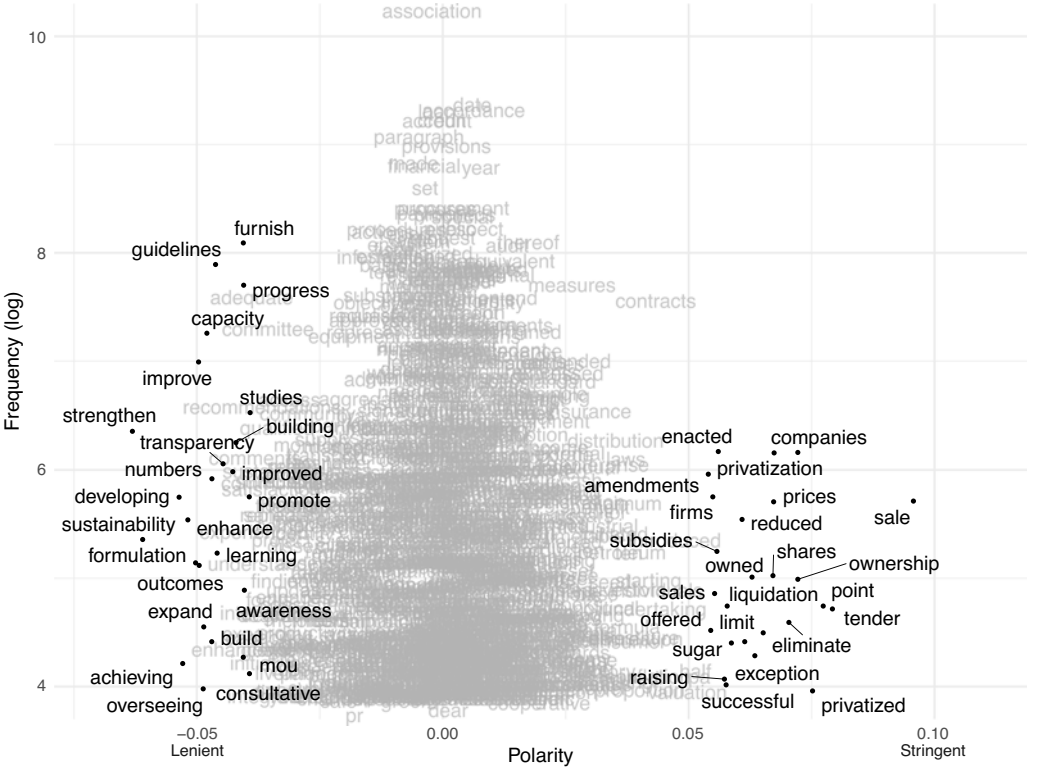

Figure 3. Term polarity, most polarizing words in bold.

Table 2. Condition stringency examples

Country	Project	Condition text	Polarity	Percentile
Kenya, 2001	P001333: Kenya Sexually Transmitted Infections Project	Representatives of other NGOs, and other Ministries and specialized divisions of MOH may be invited, as necessary, to attend the meetings of the Steering Committee.	−2.25	2%
Brazil, 2008	P080830: BR Maranhao Integrated Program: Rural Poverty Reduction Project	The Borrower has adopted the Operational Manual; and the Borrower has furnished to the Bank the terms of reference for the baseline evaluation study and the overall Project evaluation framework, as required for the impact evaluation to be carried out under Part D of the Project.	−0.32	31%
Armenia, 1996	P042793: Structural Adjustment Credit (SAC)	The Borrower has finalized the report presenting the status of primary, secondary, and technical education facilities.	−0.01	52%
Gambia, 2012	P123679: Budget Support—DPL	The Recipient has improved public financial management accountability, through the appointment of five judges for its tax and customs tribunal, as required in the Income and Sales Tax Act and the Customs and Excise Tax Act.	0.30	71%
Pakistan, 2012	P113372: Pakistan Poverty Reduction and Economic Support Operation (PRESO)	The Recipient has adopted and begun implementation of a plan to eliminate power sector subsidies by June, consistent with the Recipient's budget for the fiscal year.	2.05	97%

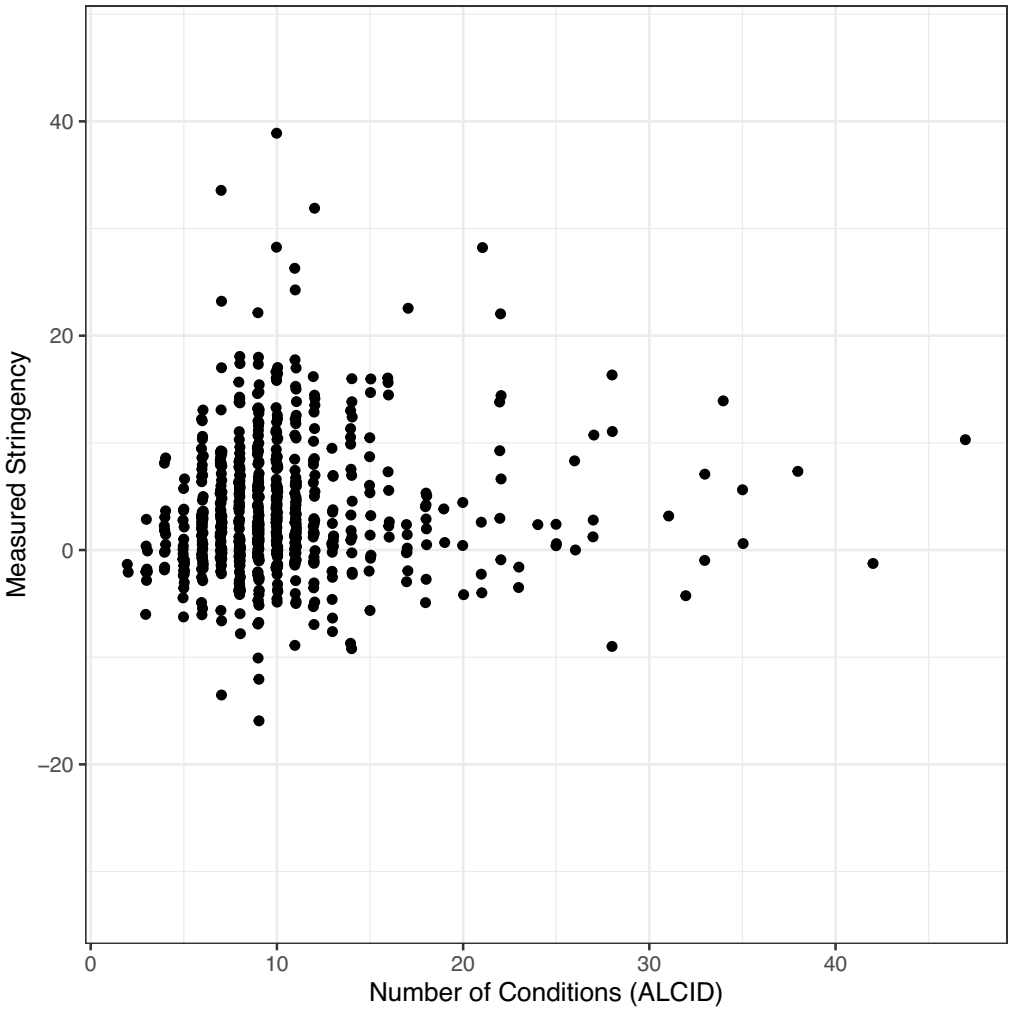

Figure 4. Condition count versus stringency.

As a preliminary test, we visualize the correlation between the number of conditions reported in the Development Policy Action Database used in most previous studies and our measure of loan stringency, finding no systematic relationship between count and stringency (Figure 4).

5.4. Data and specifications

Explanatory variables draw on the most commonly used variables in tests of principal influence and our own operationalization of bureaucratic behaviors.

5.4.1. Powerful principals

Many use UN voting to test whether countries close to the US receive fewer conditions, focusing on "important" votes on which Congress is briefed (Clark and Dolan, 2021). We include *UNGA distance* between borrowers and the US on important votes (Bailey et al., 2017). We alternatively include *UNGA distance* between borrowers and the G5 on important votes (US, UK, France, Germany, and Japan) to account for collective principal effects (Copelovitch, 2010). Greater voting distance should

be associated with more stringency, insofar as "foes" are treated less favorably than "friends" (Clark and Dolan, 2021). As an alternative measure of UNGA alignment, we use *UNGA agreement* on all votes for both US and G5, rather than just important ones.

UN Security Council (UNSC) temporary members may also benefit from tit-for-tat relationships with permanent UNSC members and thus major Bank principals (Dreher et al., 2009a, 2009b). A dummy indicates whether a borrower held *UNSC Membership* the year a loan agreement was signed. The expectation is a negative coefficient (less stringency).

5.4.2. Staff and bureaucratic incentives

We operationalize bureaucracy with variables reflecting internal procedures and staff incentives. A dummy indicates whether a loan was approved in the *Final Fiscal Quarter* of the Bank's fiscal year (April–June). We expect lenient conditions when staff are pressured to conclude projects before year-end (as suggested above in Figure 1). Staff may adjust conditionality to expedite negotiations or expedite board approval.

To control for country's economic fundamentals, we include *Economic Growth, Inflation*, and *GDP per capita*. If a country is in or near economic crisis, we expect stricter conditions. A dummy indicates if the borrower was also drawing on conditional IMF resources via an *IMF Program*, given evidence of coordination between both institutions (Woods, 2000). If a country's *Credit Rating* is weak, we also expect stricter conditionality.

Staff may perceive borrowing governments as less likely to comply with conditions if their policies do not align with Bank preferences. Insofar as the Bank privileges liberal economic policy reforms, we expect borrowing governments supported by the working classes (Cormier, 2021, 2024) to receive stringent conditions. Using V-Party (Lindberg et al., 2022), we code when borrowing governments depend on urban or rural *Working Class Support*.

Logged *Total Loan Size* captures larger projects, which carry more risk for staff. We include the number of loans a borrower received in the last five years (*NumLoans5Yrs*), as repeat clients may either represent less risk due to strong working relationships or represent recidivism and get stronger conditionality (Easterly, 2005, 4; Graham et al., 2004). To preserve its role in the aid landscape, the Bank may be lenient with borrowers using other official creditors (Hernandez, 2017), in particular where China is also lending (Humphrey and Michaelowa, 2019). *Chinese Lending* as a share of national GDP, averaged over the last three years, measures a borrower's use of Chinese finance (Custer et al., 2021; Dreher et al., 2022; Cormier, 2023).

We account for different Bank lending operations. International Bank for Reconstruction and Development (*IBRD*) controls for loans rather than International Development Association (IDA) grants. *Policy* lending, including development policy loans (DPLs), sector loans, structural loans, and programmatic-structural adjustment loans likely increase risk due to questions about fund fungibility and "impact and fiduciary soundness" (Koeberle et al., 2006, 1). Since stringency might differ across substantive areas, we control for the loan sector by running a semi-supervised topic model over the condition text and assigning the loan to the sector with the highest probability (Cormier and Manger, 2022).[4]

5.4.3. Effectiveness and past performance

Bank reviews suggest that conditionality should vary by past borrower performance. The Bank's IEG measures performance in previous loans on a six-point scale (Highly satisfactory = 5, Unsatisfactory = 0). We use a borrower's average score over IEG reports released in the three preceding years (*IEG Evaluations*).

Bank reviews also suggest that strong institutions affect loan success (Koeberle, 2005a). We control for *Democracy* because, insofar as democracies are more transparent (Hollyer et al., 2011), staff may perceive democracies as less risky. We also control for *Corruption* since staff perceptions of loan risk

[4]See the online appendix for the model and output.

may vary by how corrupt (or perceived to be) a regime is while corrupt borrowers may not perceive stringency in the same way as less-corrupt borrowers (Coppedge et al., 2022).

5.4.4. Bank policy shifts

We include a dummy for the period before 2005, when the Bank shifted from Structural Adjustment Lending to Development Project Financing, and a dummy for the period after 2012, when the Bank limited the number of conditions that should be in a loan.

5.4.5. Models

The unit of observation is a loan. In our first set of models, we use ordinary least squares (OLS) with fixed effects for the country and loan sector:

$$(1)\ Stringency_{its} = Intercept + PrincipalControls_{its} + BureaucracyControls_{its} + PerformanceControls_{its} + CountryEffects_i + SectorEffects_s + Pre2005 + Post2012 + \varepsilon_{its}$$

Not all countries agree to Bank loans each year, leading to possible selection bias since stringency is only observable in years when countries receive loans. Accordingly, we next estimate a two-stage model in which the first stage models the hazard of obtaining a Bank loan in the first place (Stone, 2008; Clark and Dolan, 2021) and include this estimated hazard in the second stage.

The two-stage approach requires an instrument that impacts the probability of selection (relevance) but not the dependent variable (excludability). Lang (2021) estimates IMF program effects with a shift-share approach, interacting the level of IMF liquidity with a country's likelihood of receiving a loan. We adapt this to the World Bank. We use the volume of IDA replenishments as a constraining liquidity factor that may affect the probability of getting a loan. We interact *IDA replenishment* volume with the base *probability* of receiving Bank loans, defined as the proportion of years from a country's independence until the observation year in which a country had received a loan.

Excludability is plausible because there is no direct path through which IDA replenishments or the number of past Bank projects could affect stringency except via the loan itself. We prefer this IDA instrument rather than using all Bank lending because the Bank can issue bonds to raise IBRD capital (Humphrey, 2014). Even though IBRD lending is limited by capital adequacy considerations, access to bond markets means that the Bank's IBRD capital stock is likely correlated with numerous global economic variables that also directly affect individual borrower economies and thus the conditions they receive. Regardless, we show results using IBRD replenishments, IBRD reserves, and the maximum funding envelope in the Country Partnership Framework as alternative instruments in the online appendix. Results are consistent.

The second question is relevance. The typical Heckman procedure is to use a likelihood ratio test to check if the first-stage coefficients are jointly zero. We follow a similar approach, but given the dependence of errors within countries and years, we bootstrap the standard errors clustered at the country-year level. The F-statistic that the shift-share variable coefficients (the base probability, IDA replenishment volume, and their interaction) are jointly zero is 166.5, much higher than conventional thresholds for instrument relevance.[5] This test is more demanding than merely assessing if the inverse Mills ratio in a two-stage model is statistically significant at the 95 percent level where the null hypothesis is that the hazard is equal to zero. First-stage model results are reported in the online appendix.

The selection specification is:

$$(2)\ SelectionHazard_{ijt} = Intercept + IDAReplenishmentXPastProjectYears + PastProjectYears + IDAReplenishment + TimeSplines + \varepsilon_{its}$$

[5]For a further discussion of the use of an F-test in this context, see Andrews et al. (2019).

$$(3)\ Stringency_{its} = Intercept + PrincipalControls_{its} + BureaucracyControls_{its} + PerformanceControls_{its} + SelectionHazard_{ijt} + CountryEffects_i + SectorEffects_s + Pre2005 + Post2012 + \varepsilon_{its}$$

In the first stage, we use a cubic spline to account for temporal trends (Beck et al., 1998). In the second stage, we include fixed effects for the country and loan sector. Estimates are similar and inferences are unchanged across modeling strategies. To obtain standard errors clustered on country and years, we use a wild bootstrap resampling procedure.

6. Results

Table 3 reports initial OLS estimations. First is a bivariate model of borrower UN voting alignment to the US with only country fixed effects. Models 2–3 include other UN voting alignment measures, UNSC membership, and controls used in previous conditionality research. These measures of principal influence are not significantly associated with Bank loan stringency.

Models 4–9 introduce organizational variables. Across models, these organizational measures are significantly associated with the stringency of Bank loan conditions, while various combinations of principal measures remain insignificant. Models 4–7 include various measures of borrower UN voting alignment with major Bank principals as well as UNSC membership. Models 10–11 restrict the sample to policy lending, which has been the narrower subject of many previous conditionality studies. Across specifications, borrower alignment with powerful principals is not statistically significant. In other words, we find no evidence that a borrower's relationship with principals determines Bank conditionality when the stringency of loan condition text is the dependent variable.

Instead, we find evidence that concern for loan failure gives rise to some bureaucratic effects on Bank conditionality. Across all specifications, loans to working-class-supported governments and larger loans are more stringent. Meanwhile, policy budget support loans and the presence of a conditional IMF program are significantly associated with more stringent conditionality. These effects are only significant at conventional levels in the full sample, but the sign and magnitude are similar for the smaller sample of policy loans. For policy loans, repeat borrowers and those with high proportions of Chinese finance face less stringent conditionality. In addition, lending approved at the end of the fiscal year is less stringent, suggesting staff offer looser conditionality when they must get money out the door to exhaust annual budgets.

Stringency decreased after the Bank shifted to Development Policy Financing in 2005. Conversely, when the Bank decided to limit the number of conditions in each loan in 2012, it appears to have compensated with slightly more stringent conditions. But neither control undermines the finding that stringency is associated with bureaucratic traits rather than principal influence. Interestingly, despite Bank pronouncements about what effective conditionality would be, we do not find a significant association between stringency and borrower institutional quality (*Corruption, Democracy*) or past performance (*IEG evaluations*).

Table 4 shows the same specifications with the two-stage model. Inferences are unchanged. The selection parameter (inverse Mills ratio [IMR]) is not significant, but the value of the first-stage F-test discussed above indicates that the instrument is relevant.

Figure 5 shows a standardized coefficient plot to facilitate the interpretation of the relative effect size in the US, G5, and DPL-only models (Table 3, models 4, 5, and 8, respectively).

The online appendix includes dozens of additional tests. Most importantly, we replace the stringency-dependent variable with the count-of-condition variable used in most studies. We replicate findings that powerful principals do affect conditionality *count* using our models, as borrower alignment with the US and the G5 decreases the number of conditions in a loan. Evidently, arguments about powerful principal constraints on Bank lending appear to depend on using count measures of loan conditionality, but focusing on the content of loan conditions leads to different conclusions. We check the robustness of different methods of aggregating stringency of conditions, year fixed effects,

https://doi.org/10.1017/psrm.2025.19 Published online by Cambridge University Press

Table 3. OLS models

	Dependent variable: World Bank loan condition stringency:								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
UNGA Distance, (US, Important)	−0.327 (0.174)	0.098 (0.207)		0.159 (0.237)				−0.779 (0.688)	
UNGA Distance, (G5, Important)			0.297 (0.277)		0.357 (0.301)				−1.157 (0.800)
UNGA Agreement, (US, All)						−0.589 (1.192)			
UNGA Agreement, (G5, All)							−0.509 (2.288)		
UNSC Membership		0.011 (0.428)	0.008 (0.416)	0.330 (0.417)	0.330 (0.401)	0.329 (0.395)	0.332 (0.394)	0.698 (0.754)	0.691 (0.762)
Working Class Support				1.036* (0.404)	1.017* (0.411)	1.044* (0.414)	1.040* (0.415)	2.248** (0.794)	2.272** (0.794)
Total Loan Size				0.320* (0.124)	0.323* (0.126)	0.319* (0.125)	0.317* (0.125)	1.554*** (0.335)	1.541*** (0.330)
Final Fiscal Quarter				−0.500* (0.233)	−0.495* (0.237)	−0.503* (0.229)	−0.501* (0.221)	−0.459 (0.497)	−0.459 (0.500)
IMF Program				0.699* (0.272)	0.708* (0.285)	0.677* (0.269)	0.678* (0.284)	0.105 (0.656)	0.101 (0.676)
Number Loans 5 Yrs				−0.003 (0.016)	−0.003 (0.016)	−0.002 (0.016)	−0.003 (0.016)	−0.070* (0.033)	−0.071* (0.033)
Credit Rating				−0.150 (0.087)	−0.147 (0.084)	−0.158 (0.088)	−0.156 (0.084)	−0.253 (0.202)	−0.256 (0.192)
Chinese Lending				−0.516 (9.267)	−0.456 (9.704)	−0.482 (9.684)	−0.496 (10.030)	−38.074* (15.771)	−38.334* (15.952)

(*Continued*)

https://doi.org/10.1017/psrm.2025.19 Published online by Cambridge University Press

Table 3. (*Continued.*)

	Dependent variable: World Bank loan condition stringency:								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
IBRD				0.206	0.190	0.227	0.226	2.678*	2.738*
				(0.751)	(0.762)	(0.771)	(0.785)	(1.317)	(1.265)
Policy Lending				3.372***	3.369***	3.353***	3.360***		
				(0.314)	(0.316)	(0.312)	(0.328)		
Economic Growth		−0.059*	−0.059*	−0.044	−0.044	−0.047	−0.045	−0.119	−0.117
		(0.028)	(0.029)	(0.029)	(0.028)	(0.030)	(0.028)	(0.072)	(0.070)
Inflation		−0.002	−0.001	0.0001	0.001	−0.0003	−0.0003	−0.048	−0.052
		(0.009)	(0.009)	(0.008)	(0.007)	(0.008)	(0.008)	(0.037)	(0.039)
ln GDP per capita		1.151	1.281*	1.933*	1.921*	1.899*	1.910*	0.283	0.330
		(0.631)	(0.629)	(0.916)	(0.928)	(0.949)	(0.926)	(2.743)	(2.695)
IEG Evaluations		−0.011	−0.017	0.051	0.051	0.047	0.051	−0.284	−0.265
		(0.145)	(0.147)	(0.143)	(0.146)	(0.148)	(0.139)	(0.330)	(0.323)
Corruption		−1.181	−1.205	1.178	1.157	1.175	1.171	6.278	6.362
		(1.760)	(1.798)	(1.572)	(1.492)	(1.504)	(1.429)	(4.025)	(3.891)
Democracy		3.818*	4.042*	2.157	2.377	1.918	1.953	1.009	0.697
		(1.861)	(1.870)	(1.834)	(1.841)	(1.683)	(1.755)	(4.702)	(4.436)
Post-2012				0.910**	0.946**	0.936**	0.924*	2.464***	2.300***
				(0.352)	(0.339)	(0.351)	(0.380)	(0.645)	(0.629)
Post-2005				−0.958*	−0.916*	−1.058**	−1.074**	−2.709**	−2.629**
				(0.425)	(0.404)	(0.367)	(0.358)	(1.047)	(1.014)
Country and Sector Effects	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	2,470	2,157	2,157	1,891	1,891	1,893	1,893	600	600
R^2	0.137	0.157	0.157	0.249	0.249	0.249	0.249	0.404	0.404

Note: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

https://doi.org/10.1017/psrm.2025.19 Published online by Cambridge University Press

Table 4. Two-stage selection models

	Dependent variable: World Bank loan condition stringency								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
UNGA Distance, (US, Important)	−0.229 (0.178)	0.128 (0.207)		0.154 (0.235)				−0.835 (0.693)	
UNGA Distance, (G5, Important)			0.331 (0.278)		0.353 (0.296)				−1.181 (0.806)
UNGA Agreement, (US, All)						−0.585 (1.200)			
UNGA Agreement, (G5, All)							−0.549 (2.270)		
UNSC Membership		0.011 (0.422)	0.006 (0.422)	0.331 (0.386)	0.331 (0.393)	0.331 (0.401)	0.334 (0.411)	0.695 (0.768)	0.690 (0.749)
Working Class Support				1.040* (0.407)	1.022* (0.414)	1.049* (0.416)	1.045* (0.423)	2.282** (0.805)	2.303** (0.798)
Total Loan Size				0.320* (0.128)	0.323* (0.132)	0.319* (0.128)	0.318* (0.126)	1.563*** (0.326)	1.548*** (0.340)
Final Fiscal Quarter				−0.500* (0.231)	−0.495* (0.235)	−0.503* (0.234)	−0.501* (0.232)	−0.457 (0.484)	−0.457 (0.514)
IMF Program				0.705* (0.284)	0.714** (0.271)	0.685* (0.277)	0.685* (0.279)	0.149 (0.683)	0.143 (0.691)
Number Loans 5 Yrs				−0.003 (0.016)	−0.002 (0.016)	−0.002 (0.016)	−0.002 (0.016)	−0.069* (0.033)	−0.070* (0.032)
Credit Rating				−0.149 (0.088)	−0.147 (0.084)	−0.156 (0.082)	−0.155 (0.087)	−0.233 (0.199)	−0.238 (0.202)
Chinese Lending				−0.362 (9.900)	−0.310 (9.377)	−0.307 (10.006)	−0.321 (9.910)	−37.615* (16.426)	−37.919* (15.912)

(*Continued*)

https://doi.org/10.1017/psrm.2025.19 Published online by Cambridge University Press

Table 4. (*Continued.*)

	Dependent variable: World Bank loan condition stringency								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
IBRD				0.211	0.195	0.232	0.232	2.686*	2.741*
				(0.764)	(0.772)	(0.758)	(0.767)	(1.236)	(1.253)
Policy Lending				3.373***	3.370***	3.354***	3.361***		
				(0.313)	(0.317)	(0.330)	(0.323)		
Economic Growth		−0.059*	−0.059*	−0.045	−0.044	−0.047	−0.045	−0.118	−0.117
		(0.028)	(0.028)	(0.027)	(0.028)	(0.029)	(0.029)	(0.072)	(0.070)
Inflation		−0.001	−0.001	−0.0001	0.001	−0.0005	−0.0005	−0.048	−0.052
		(0.009)	(0.009)	(0.008)	(0.008)	(0.008)	(0.007)	(0.039)	(0.039)
ln GDP per capita		1.044	1.155	1.943*	1.931*	1.911*	1.925*	0.402	0.448
		(0.640)	(0.659)	(0.942)	(0.884)	(0.909)	(0.953)	(2.718)	(2.775)
IEG Evaluations		−0.004	−0.010	0.050	0.051	0.047	0.050	−0.303	−0.282
		(0.155)	(0.148)	(0.145)	(0.141)	(0.143)	(0.145)	(0.339)	(0.324)
Corruption		−1.244	−1.285	1.185	1.165	1.184	1.179	6.202	6.292
		(1.777)	(1.855)	(1.517)	(1.517)	(1.567)	(1.574)	(3.737)	(3.999)
Democracy		3.694	3.894*	2.185	2.407	1.958	1.997	0.980	0.713
		(1.887)	(1.937)	(1.787)	(1.759)	(1.778)	(1.766)	(4.627)	(4.502)
IMR	−1.280	−0.768	−0.838	0.326	0.312	0.369	0.390	2.692	2.429
	(0.807)	(1.060)	(1.053)	(0.916)	(0.895)	(0.897)	(0.895)	(3.295)	(3.216)
Post-2012				0.913*	0.947**	0.938**	0.924*	2.504***	2.324***
				(0.368)	(0.346)	(0.339)	(0.361)	(0.645)	(0.629)
Post-2005				−0.939*	−0.896*	−1.033**	−1.047**	−2.625*	−2.527*
				(0.418)	(0.403)	(0.383)	(0.379)	(1.045)	(0.997)
Country and Sector Effects	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	2,470	2,157	2,157	1,891	1,891	1,893	1,893	600	600
R^2	0.138	0.157	0.157	0.249	0.249	0.249	0.249	0.404	0.405

Note: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

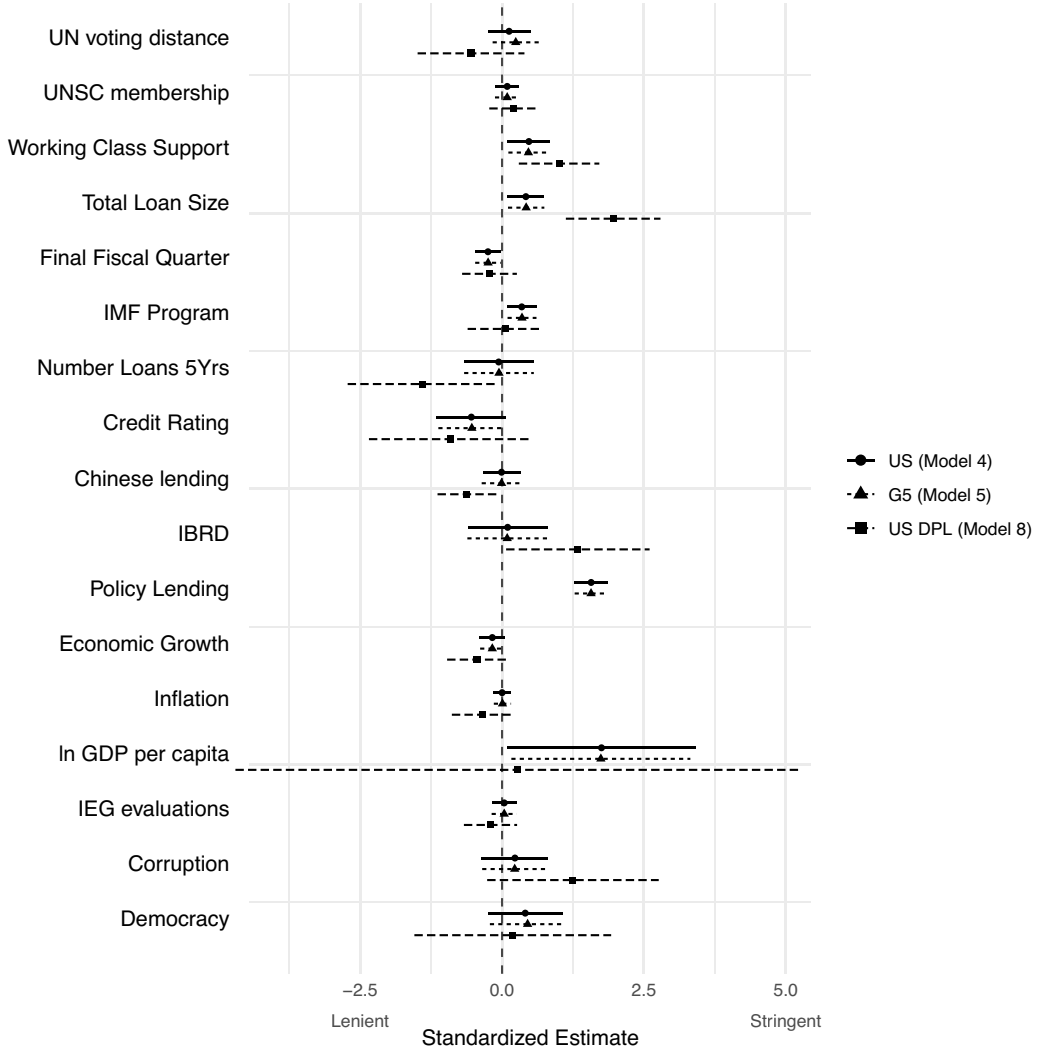

Figure 5. Standardized coefficient plot (95% confidence intervals).

and alternative measures of principal interests: all UN votes, UNGA movement toward US (Kilby, 2009), UNSC voting alignment, bilateral aid, foreign direct investment, and US troop presence (Allen et al., 2022). We add alternative controls (other corruption measures, current account balance, and total past loans) and sample subsets (post-2005 only). Results are consistent.

7. Conclusion

A large literature debates the politics of Bank lending, and a prominent finding is that conditionality reflects the preferences of its most powerful principals. We present evidence that reframes the debate. We first highlight the importance of scrutinizing the *content* of Bank conditionality rather than merely the *count* of conditions. We emphasize the variability of loan *stringency*: the degree to which conditions require quantifiable, objectively verifiable, and thus relatively demanding actions by the borrower vis-à-vis the degree to which conditions involve vague, subjective, and thus less-demanding

actions by the borrower. We show that stringency is not correlated with the number of conditions in a loan. When stringency is the subject of analysis, Bank lending is not determined by powerful principals. Instead, the stringency of Bank loan conditions is primarily associated with loan risk and bureaucratic tendencies. Accordingly, future quantitative studies should use bureaucratic and risk measures to ensure they do not underplay the importance of organizational factors in World Bank lending.

As Bank insiders argue, what is most impactful on borrowers is stringency rather than the number of conditions (Kapur and Webb, 2000; Koeberle, 2005b). Prioritizing the study of conditionality text is a policy-relevant path forward for researchers. To aid this effort, our replication data provide the largest dataset of Bank loan conditions available to date, covering the text of 3,641 World Bank loans since 1995. For comparison, the publicly available and frequently used Development Policy Action Database only goes back to 2004.

Lastly, our findings reframe the debate about principal control and agent autonomy in research on the World Bank, IMF, and IOs. We present evidence that findings about powerful principals appear to depend on the measures used. This study only hints at the reasons for this. Section 2 highlighted that Bank staff have incentives to minimize the risk of loan failure, and conditionality language is a core tool for managing risk while meeting the disbursement imperative. This may incentivize focus on the detail of conditionality text rather than the number of conditions. In contrast, there is some evidence that principals use count as a heuristic for risk management—as noted in Section 2, Bank-wide mandates are based on conditionality count rather than content. Carefully parsing which outcomes are subject to which theories and why, at the World Bank and other IOs, is another important path forward for future research.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/psrm.2025.19. To obtain replication material for this article, https://doi.org/10.7910/DVN/AF3EZP.

References

Allen MA, Flynn ME and Martinez Machain C (2022) US global military deployments, 1950–2020. *Conflict Management and Peace Science* **39**(3), 351–370. doi:10.1177/07388942211030885

Andrews I, Stock JH and Sun L (2019) Weak instruments in instrumental variables regression: theory and practice. *Annual Review of Economics* **11**, 727–753. doi:10.1146/annurev-economics-080218-025643

Bailey MA, Strezhnev A and Voeten E (2017) Estimating dynamic state preferences from United Nations voting data. *Journal of Conflict Resolution* **61**(2), 430–456. doi:10.1177/0022002715595700

Barnett MN and Finnemore M (1999) The politics, power, and pathologies of international organizations. *International Organization* **53**(4), 699–732. doi:10.1162/002081899551048

Barnett MN and Finnemore M (2004) *Rules for the World*. Cornell University Press: Ithaca, NY.

Beck N, Katz JN and Tucker R (1998) Taking time seriously: time-series-cross-section analysis with a binary dependent variable. *American Journal of Political Science* **42**(4), 1260–1288. doi:10.2307/2991857

Best J (2014) *Governing Failure: Provisional Expertise and the Transformation of Global Development Finance*. Cambridge: Cambridge University Press.

Bird G, Hussain M and Joyce JP (2004) Many happy returns? Recidivism and the IMF. *Journal of International Money and Finance* **23**(2), 231–251. doi:10.1016/j.jimonfin.2003.12.002

Briggs RC (2021) Why does aid not target the poorest? *International Studies Quarterly* **65**(3), 739–752. doi:10.1093/isq/sqab035

Buntaine MT, Parks BC and Buch BP (2017) Aiming at the wrong targets: the domestic consequences of international efforts to build institutions. *International Studies Quarterly* **61**(2), 471–488. doi:10.1093/isq/sqx013

Chwieroth JM (2013) The silent revolution: how the staff exercise informal governance over IMF lending. *The Review of International Organizations* **8**(2), 265–290. doi:10.1007/s11558-012-9154-9

Clark R and Dolan LR (2021) Pleasing the principal: U.S. Influence in World Bank policymaking. *American Journal of Political Science* **65**(1), 36–51. doi:10.1111/ajps.12531

Copelovitch MS (2010) Master or Servant? Common Agency and the Political Economy of IMF Lending. *International Studies Quarterly* **54**(1), 49–77. doi:10.1111/j.1468-2478.2009.00577.x

Coppedge M, Gerring J, Knutsen CH, Lindberg SI, Teorell J, Alizada N, Altman D, et al. (2022) *V-Dem Dataset V12*. Gothenburg: Varieties of Democracy (V-Dem Project). doi: 10.23696/vdemds22.

Cormier B (2021) Interests over institutions: political-economic constraints on public debt management in developing countries. *Governance* **34**(4), 1167–1191. doi:10.1111/gove.12551

Cormier B (2023) Chinese or western finance? Transparency, official credit flows, and the international political economy of development. *The Review of International Organizations* **18**, 297–328. doi:10.1007/s11558-022-09469-x

Cormier B (2024) *How Governments Borrow: Partisan Politics, Constrained Institutions, and Sovereign Debt in Emerging Markets*. Oxford: Oxford University Press.

Cormier B and Manger MS (2022) Power, ideas, and World Bank conditionality. *The Review of International Organizations* **17**(3), 397–425. doi:10.1007/s11558-021-09427-z

Custer S, Dreher A, Elston T-B, Fuchs A, Ghose S, Lin JJ, Malik AA, et al. (2021) *Tracking Chinese Development Finance: An Application of AidData's TUFF 2.0 Methodology*. Version 2.0. Williamsburg, VA: AidData at William & Mary.

Dreher A, Fuchs A, Parks BC, Strange AM and Tierney MJ (2022) *Banking on Beijing: The Aims and Impacts of China's Overseas Development Program*. Cambridge: Cambridge University Press.

Dreher A, Sturm J-E and Vreeland JR (2009a) Development aid and international politics: does membership on the UN security council influence World Bank decisions? *Journal of Development Economics* **88**(1), 1–18. doi:10.1016/j.jdeveco.2008.02.003

Dreher A, Sturm J-E and Vreeland JR (2009b) Global horse trading: IMF loans for votes in the United Nations Security Council. *European Economic Review* **53**(7), 742–757. doi:10.1016/j.euroecorev.2009.03.002

Easterly W (2005) What did structural adjustment adjust? The association of policies and growth with repeated IMF and World Bank adjustment loans. *Journal of Development Economics* **76**(1), 1–22. doi:10.1016/j.jdeveco.2003.11.005

Eckhard S and Parizek M (2022) Policy implementation by international organizations: a comparative analysis of strengths and weaknesses of national and international staff. *Journal of Comparative Policy Analysis: Research and Practice* **24**(3), 254–270. doi:10.1080/13876988.2020.1813032

Eshima S, Imai K and Sasaki T (2020) "Keyword assisted topic models." arXiv Preprint. Cambridge, MA: Harvard University. http://arxiv.org/abs/2004.05964.

Ferranti DD (2007) The World Bank and middle income countries. In Birdsall N (ed), *Rescuing The World Bank*. Washington, DC: Brookings Institution Press, pp. 133–151.

Gallagher RJ, Reing K, Kale D and Ver Steeg G (2017) Anchored correlation explanation: topic modeling with minimal domain knowledge. *Transactions of the Association for Computational Linguistics* **5**, 529–42. doi:10.1162/tacl_a_00078

Gould JA and Winters MS (2007) An obsolescing bargain in Chad: shifts in leverage between the government and the World Bank. *Business and Politics* **9**(2), 1–34. doi:10.2202/1469-3569.1199

Gutner T (2005) Explaining the gaps between mandate and performance: agency theory and World Bank environmental reform. *Global Environmental Politics* **5**(2), 10–37. doi:10.1162/1526380054127727

Hawkins DG and Jacoby W (2006) How agents matter. In Hawkins DG, Lake DA, Nielson DL and Tierney MJ (eds), *Delegation and Agency in International Organizations*. Cambridge: Cambridge University Press, pp. 199–228.

Hernandez D (2017) Are 'new' donors challenging World Bank conditionality? *World Development* **96**, 529–549. doi:10.1016/j.worlddev.2017.03.035

Hollyer JR, Rosendorff BP and Vreeland JR (2011) Democracy and transparency. *The Journal of Politics* **73**(4), 1191–1205. doi:10.1017/S0022381611000880

Honig D (2018) *Navigation by Judgment: Why and When Top-Down Management of Foreign Aid Doesn't Work*. Oxford: Oxford University Press.

Honig D (2019) When reporting undermines performance: the costs of politically constrained organizational autonomy in foreign aid implementation. *International Organization* **73**(1), 171–201. doi:10.1017/S002081831800036X

Humphrey C (2014) The politics of loan pricing in multilateral development banks. *Review of International Political Economy* **21**(3), 611–639. doi:10.1080/09692290.2013.858365

Humphrey C and Michaelowa K (2019) China in Africa: competition for traditional development finance institutions? *World Development* **120**, 15–28. doi:10.1016/j.worlddev.2019.03.014

Kapur D and Webb R (2000) Governance-related conditionalities of the international financial institutions. *G-24 Discussion Paper Series* No. 6. Geneva: United Nations Conference on Trade and Development.

Kilby C (2009) The political economy of conditionality: an empirical analysis of World Bank loan disbursements. *Journal of Development Economics* **89**(1), 51–61. doi:10.1016/j.jdeveco.2008.06.014

Koeberle S (2005a) *Conditionality Revisited: Concepts, Experiences, and Lessons*. Washington, DC: The World Bank.

Koeberle S (2005b) Conditionality: under what conditions? In Koeberle S, Bedoya H, Silarsky P and Verheyen G (eds), *Conditionality Revisited*. Washington DC: The World Bank, pp. 57–83.

Koeberle S, Walliser J and Stavreski Z (eds.) (2006) *Budget Support as More Effective Aid? Recent Experiences and Emerging Lessons*. Washington, DC: The World Bank.

Kramarz T (2020) *Forgotten Values: The World Bank and Environmental Partnerships*. Cambridge, MA: MIT Press.

Lall R (2017) Beyond institutional design: explaining the performance of international organizations. *International Organization* **71**(2), 245–280. doi:10.1017/S0020818317000066

Lang V (2021) The economics of the democratic deficit: the effect of IMF programs on inequality. *The Review of International Organizations* **16**(3), 599–623. doi:10.1007/s11558-020-09405-x

Lindberg SI, Düpont N, Higashijima M, Berker Kavasoglu Y, Marquardt KL, Bernhard M, Döring H, et al. (2022) *Varieties of Party Identity and Organization (V–Party) Dataset V2*. Version 2. Varieties of Democracy (V–Dem) Project.

Lyne MM, Nielson DL and Tierney MJ (2006) Who delegates? Alternative models of principals in development aid. In Hawkins DG, Lake DA, Nielson DL and Tierney MJ (eds), *Delegation and Agency in International Organizations*. Cambridge: Cambridge University Press, pp. 41–76.

Lyne MM, Nielson DL and Tierney MJ (2009) Controlling coalitions: social lending at the multilateral development banks. *The Review of International Organizations* **4**(4), 407–433. doi:10.1007/s11558-009-9069-2

McLean E and Schneider C (2014) Limits of informal governance? The scope of conditionality in the World Bank. *Paper Prepared for the 7th Annual Conference of the Political Economy of International Organizations*, 30.

Moll P, Geli P and Saavedra P (2015) Correlates of success in World Bank development policy lending. *World Bank Policy Research Working Paper* 7181. Washington, DC: The World Bank.

Nelson SC (2017) *The Currency of Confidence: How Economic Beliefs Shape the IMF's Relationship with Its Borrowers*. Ithaca, NY: Cornell University Press.

Nielson DL and Tierney MJ (2003) Delegation to international organizations: agency theory and World Bank environmental reform. *International Organization* **57**(2), 241–276. doi:10.1017/S0020818303572010

Park S and Kramarz T (2019) *Global Environmental Governance and the Accountability Trap*. Cambridge, MA: MIT Press.

Park S and Vetterlein A (eds.) (2010) *Owning Development: Creating Policy Norms in the IMF and the World Bank*. Cambridge: Cambridge University Press.

Rodrik D (2006) Goodbye Washington consensus, hello Washington confusion? A review of the World Bank's economic growth in the 1990s: learning from a decade of reform. *Journal of Economic Literature* **44**(4), 973–987. doi:10.1257/jel.44.4.973

Safarty G (2012) *Values in Translation: Human Rights and the Culture of the World Bank*. Stanford, CA: Stanford University Press.

Smets L and Knack S (2014) World Bank lending and the quality of economic policy. *World Bank Policy Research Working Paper* 6924. Washington, DC: The World Bank.

Stone RW (2008) The scope of IMF conditionality. *International Organization* **62**(4), 589–620. doi:10.1017/S0020818308080211

Stone RW (2011) *Controlling Institutions: International Organizations and the Global Economy*. Cambridge: Cambridge University Press.

Swaroop V (2016) World Bank's experience with structural reforms for growth and development. *World Bank MFM Global Practice Discussion Paper No. 11*. Washington, DC: The World Bank.

Trubowitz P and Watanabe K (2021) The geopolitical threat index: a text-based computational approach to identifying foreign threats. *International Studies Quarterly* **65**(3), 852–865. doi:10.1093/isq/sqab029

Vaubel R (2006) Principal-agent problems in international organizations. *The Review of International Organizations* **1**(2), 125–138. doi:10.1007/s11558-006-8340-z

Watanabe K (2021) Latent semantic scaling: a semisupervised text analysis technique for new domains and languages. *Communication Methods and Measures* **15**(2), 81–102. doi:10.1080/19312458.2020.1832976

Weaver C (2007) The world's bank and the bank's world. *Global Governance* **13**(4), 493–512. doi:10.1163/19426720-01304005

Weaver C (2008) *Hypocrisy Trap: The World Bank and the Poverty of Reform*. Princeton, NJ: Princeton University Press.

Weber M (1946) *From Max Weber: Essays in Sociology*. New York: Oxford University Press.

Williamson J (1990) What Washington means by policy reform. In Williamson J (ed), *Latin American Adjustment: How Much Has Happened?* Washington, DC: Institute for International Economics, pp. 7–20.

Winters MS (2010) Choosing to target: what types of countries get different types of World Bank projects. *World Politics* **62**(3), 422–458. doi:10.1017/S0043887110000092

Woods N (2000) The challenge of good governance for the IMF and the World Bank themselves. *World Development* **28**(5), 823–841. doi:10.1016/S0305-750X(99)00156-4

World Bank (2017) *Minutes of Meeting of the Executive Directors of the Bank and IDA and the Boards of Directors of IFC and MIGA Held in the Board Room on March 28, 2017 at 10:03 A.M*. Washington, DC: The World Bank.

World Bank (2022) World Bank project cycle. https://projects.worldbank.org/en/projects-operations/products-and-services/brief/projectcycle (accessed 5 May 2023).

World Bank IEG (2016) *Behind the Mirror: A Report on Self-Evaluation Systems of the World Bank Group*. Washington, DC: The World Bank.

Cite this article: Winter J, Cormier B, Kramarz T and Manger MS (2025) Reassessing World Bank conditionality: beyond count measures. *Political Science Research and Methods*, 1–21. https://doi.org/10.1017/psrm.2025.19