How Women Matter:
Women's Representation and Gender Mainstreaming at the World Bank

Mirko Heinzel (University of Potsdam)
Mheinzel@uni-potsdam.de

Catherine Weaver (University of Texas at Austin)
ceweaver@austin.utexas.edu

Samantha Jorgensen (University of Texas at Austin)
smjorgensen@utexas.edu

Paper presented at the International Political Economy Society Annual Meeting, University of Colorado-Boulder, October 23, 2021. DRAFT paper - please do not quote. Feedback welcome!

Abstract:
Research on the decision-making of International Organizations (IOs) has increasingly focused on the development of policy norms and scripts within IOs. Much less is known about the degree to which these policy norms are applied by the bureaucrats shaping the day-to-day work of IOs. We aim to fill this gap by focusing on the application of the Gender and Development norm in the World Bank. Particularly, we probe how skewed gender within staff (only roughly 35% of the people in charge of projects are women) may translate into variations in the focus on gender issues in the World Bank's work. We test the argument by utilizing novel data on the gender of thousands of World Bank staff members. The empirical part of this paper examines whether patterns of gender representation in World Bank staff translate into the differential application of gender mainstreaming norms in World Bank projects and the degree to which more than 55,000 external stakeholders from more than 130 countries perceive the World Bank as effective at addressing gender equity. The article contributes to debates on the internal workings of IOs by highlighting how differences in representation within IO staff affect concrete organizational outcomes.

Introduction

Scholarship on International Organizations (IOs) has long highlighted that commitment to policy norms and the application of these norms can diverge in different organizational settings (Park and Vetterlein 2010; Kentikelenis, Stubbs, and King 2016; Tallberg et al. 2020).[1] Explanations for divergence between IO's formal policies and actual behavior have emphasized that IOs need to satisfy external actors in their authorizing and task environments by signaling commitment to new agendas valued by these actors. However, when such policy norms conflict with internal procedures or beliefs inscribed in their organizational culture, IOs may disconnect talk from action to simultaneously satisfy external stakeholders and maintain internal behavior. Explanations have focused on the IO-level, highlighting the de-coupling of organizations around issues such as UN peacekeeping (Lipson 2007), World Bank governance and anti-corruption reform (Weaver 2008), IOs commitment to liberal norms (Tallberg et al. 2020), as well as IO voice and vote reforms (Vestergaard and Wade 2013).

Deeper explorations of policy and practice within IOs, however, reveal that policy norms are often taken up variably, or unevenly, *within* the same organizations. This uneven application of policy norms presents a distinct puzzle: what may account for differences we see within IOs regarding the implementation of policy norms? We hypothesize in this paper that such variation is linked to the autonomy, discretion and preferences of key staff. Specifically, we surmise that demographic characteristics (the identities) of staff matter in explaining variation in organizational outputs, because these identities shape individual and group decisions on key actions such as the allocation of scarce funds, composition of project-level activities, selection of implementation teams, and other central factors that indelibly shape the degree to which emerging policy norms are realized in IO practice.

The argument aligns with extant observations from principal-agent and constructivist work on IOs, which place staff at the center of analysis of organizational behavior and outcomes (Barnett and Finnemore 2004; Hawkins et al. 2006;Weaver 2008; Chwieroth 2013; Eckhard and Ege 2016; Bayerlein, Knill, and Steinebach 2020). More specifically, however, our hypotheses are informed by the insights of representative bureaucracy theory - a body of work rooted in public administration literature and only more recently applied to the study of IOs. Bureaucratic representation theory, developed for the study of national administrations, emphasizes the importance of bureaucracies' workforce composition for understanding the decisions taken by them (Kingsley 1944). The central argument is that the composition of the bureaucratic workforce influences those decisions that affect citizens who share key demographic characteristics with them. Discussed demographic characteristics include, for example, bureaucrats' ethnicity, religion, income, or gender (Mosher 1982). Authors commonly differentiate between passive representation—the composition of the bureaucratic workforce itself —and active representation—the degree to which decisions are affected by this demographic make-up of the bureaucracy (Mosher 1982). With respect to gender representation (the central variable in this study), there is significant evidence from the comparative politics literature as well that indicates that women decision-makers increase the benefits of policy outcomes for other women (Dietrich, Hayes, and O'Brien 2019; Holman 2017; Keiser et al. 2002; Sanghee Park 2013; Wilkins and Keiser 2006). These studies show, for example, that greater numbers of women in legislatures decrease discriminatory tax policies (Betz, Fortunato, and O'Brien 2021), women judges are

[1] We define policy norms here to be the "shared expectation for all relevant actors within a community about what constitutes appropriate behavior, which is encapsulated in policy." (Park and Vetterlein 2010, 4).

more sympathetic towards discrimination cases brought by women (Boyd, Epstein, and Martin 2010) or that a greater number of women inside national bureaucracies increase the educational attainment of women (Sanghee Park and Liang 2021). As Mansbridge (2005, 622) has put it: "descriptive representation by gender improves substantive outcomes for women in every polity for which we have a measure."

Whether and how passive or descriptive representation actually translates into active representation in *international* organizations remains an open question. To date, the preponderance of studies have focused on passive, not active, representation in IOs. For example, there are numerous recent studies on the UN, EU and other institutions that have provided statistical descriptions on IO workforce composition of IOs, with a focus on staff nationality (Badache 2020; Christensen 2020; Novosad and Werker 2019; Parizek 2017; Ruggeri, Ruffa, and Bove 2020; Christensen 2020; Eckhard 2021). With respect to gender, feminist research on IOs has long discussed the normative importance of *passive* representation of women on decision-making (Hafner-Burton and Pollack 2002; Griffin 2009; Hannan 2012; Arora-Jonsson and Sijapati 2018; Birken and Cigna 2018; Abels and Mushaben 2020; Barraza Vargas 2019; Haack 2014a; MacRae 2012; O'Rourke 2012). This literature provides convincing evidence that women are pervasively underrepresented in leadership positions of IOs (Haack 2014b; Haack, Karns, and Murray 2020). Some studies have gone one step further in attempting to assess the impact that IO leaders like the President of the European Commission, Ursula Van der Leyen, or the Managing Director of the International Monetary Fund, Christine Lagarde, had or will potentially have on the gender-focus in these organizations decisions (Abels and Mushaben 2020; Barraza Vargas 2019; Blackmon 2020)- an approach akin to a "great women" theory of politics. However, whether increased representation of women in *staff* more generally leads to substantive representation of women in organizational decisions remains insufficiently understood (Dolan 2009).

Active gender representation is especially interesting in the study of IO bureaucracies because so many of these organizations have adopted progressive gender mainstreaming agendas that have espoused strong commitments to enhancing gender representation and empowerment in all aspects of organizational work (Haack, Karns, and Murray 2020). With respect to internal management, IOs have made increasing efforts in the past few decades to diversify their staff and hire more women (Haack, Karns, and Murray 2020; Karim and Beardsley 2016; MacRae 2012). One of the hopes of these initiatives is that women will bring a different perspective into the work of IOs and represent the interests of women – an approach that asserts that enhanced representation of women in IO decision-making is critical to realizing the goals of gender mainstreaming in IO practices (Hannan 2012). Gender mainstreaming initiatives thus often target the double-goal of increasing passive representation of women—the number of women in their staff—and active representation of women—the consideration of gender equality, empowerment, and related goals in their operations. However, the mere presence of more women in organizational decision-making roles may not guarantee that these women are able or willing to proactively promote gender mainstreaming policy norms; nor may it be the gender (versus nationality, education, or other factors) of female staff which is the primary identity upon which they define and pursue their interests. We thus seek in this paper to better understand whether and how this relationship between passive and active representation exists through a study of the World Bank and gender mainstreaming.

To this end, our paper explores evidence for "if and how women matter" by attempting to empirically observe and measure the link between passive and active gender representation in IOs, using the case of the

World Bank. Specifically, we look at the diffusion and internalization of gender and development norms at the project level at the World Bank and link this staffing composition. In line with the hypotheses of representative bureaucracy theory, we posit that women staff members act upon their identity as women to apply gender mainstreaming norms more proactively and consistently than male staff. As such, we expect that the increasing presence of women in IOs, especially in key positions of project authority and decision-making (e.g. task team leaders), increases the extent to which gender mainstreaming shows up as essential components in sectoral policy approaches. Hence, greater passive representation of women in IO staff increases the active representation of women beneficiaries directly in the decisions overseen by women.

Empirically, we utilize novel individual-level data from 2000-2020 on the gender of thousands of World Bank staff members in positions of authority—Country Directors and Task Team Leaders—who are the key actors in decision-making at the level of project content formation and implementation. Evidence from our statistical tests reveal two findings on the impact of the passive representation of women on gender mainstreaming in World Bank projects: First, women staff members oversee projects with a greater focus on gender issues, on average. This observation may appear because women could be more likely to be appointed to manage projects with gender mainstreaming or because women Country Directors and Task Team Leaders are more likely to pursue gender-related activities within projects if they have the discretion to do so. Evidence from an instrumental variable analysis utilizing plausible exogenous variation in the number of TTLs staffed in a project implies that findings are driven, at least in part, by women integrating more gender mainstreaming into their projects. Second, incorporating more women staff seems to change organizational decision-making at the project level (ie, a diffusion effect upon men who work with them). When more women are appointed in the same sectoral global practices, women *and* men staff incorporate gender mainstreaming more in the projects under their discretion. Third, enhanced representation of women in Bank staff appears to shape external perceptions of the Bank's progress on gender mainstreaming. Our analysis of existing World Bank survey data from more than 55,000 stakeholders in more than 130 countries (World Bank 2021) implies that a greater number of women staff seems to increase the degree to which the World Bank can translate effective operations into perceived contributions to gender equity.

Our paper proceeds in five steps. First, to explain our selection of the World Bank as our case study, we discuss the establishment of gender mainstreaming norms at the Bank and show that the pace and extent of mainstreaming have varied considerably within the organization over time and across units. Second, we present our explanation for three causal mechanisms that inform our hypothesis for how, when and why women matter. Third, we discuss our research design utilizing individual-level data on thousands of World Bank staff members working on projects between 2000-2020, project-level evaluations of gender mainstreaming, and survey evidence collected by the World Bank from 55,000 stakeholders from more than 130 countries between 2012 and 2020. Fourth, we present evidence from a range of statistical tests that aim to identify the impact of women and minimize the influence of selection bias that often plague research on the relationship between passive and active representation. Finally, we conclude by highlighting how the representational dynamics we observe positively contribute to the representation of women and problematize how the reliance on staff from high-income countries and the narrow focus on economic growth in gender strategies can continue to marginalize the voices of women beneficiaries within the World Bank.

Why the World Bank?

Over the past five decades, gender and development has become a central goal of global development aid.[2] The World Bank, by natural effect of its lending power and distinct role of knowledge leader in the development space, has played an outsized role vis-a-vis its institutional peers in shaping the objectives and rationale for gender mainstreaming - often in very controversial ways. Yet, while the gender mainstreaming era in the World Bank has progressed considerably over these fifty years, the application of gender mainstreaming norms in its internal human resources and operational work has been uneven (Moser and Moser 2005; Griffin 2009; Meyer and Prügl 1999; Kenny and O'Donnell 2016; World Bank 2018). Such outcomes have been linked by both scholars and internal gender advocates to the number and positionality (rank and authority) of gender specialists within the Bank over time and the Bank's slow and uneven progress in mainstreaming gender programs. Moreover, in contrast to other organizations such as the UN, the Bank's embrace of gender and development largely conformed to an approach that fit with the predominantly mainstream economist culture of the Bank and notions of "gender development as smart economics" versus contending feminist frames of women's empowerment as a basic human right. This history, briefly outlined below, informs our driving questions around women's representation in the Bank and its effect on gender mainstreaming. Simply put, *what are the links between the hard-won increases in the presence of gender specialists and women more generally in the ranks of the Bank's task team leaders and organizational managers and the outcomes we can observe with respect to gender mainstreaming in Bank's projects and perceived performance?*

Early efforts to institutionalize gender and development goals were driven by a small group of women policy norms entrepreneurs (Weaver 2010). In 1972, an informal working group came together to pressure the Bank's management to formally institutionalize its commitment to women's development - an effort that eventually led to the creation of the first Women in Development (WID) advisor position in 1977. The WID Advisor, however, was allocated minimal staff, was largely unfunded, and had only a limited mandate to raise awareness inside the Bank at a time when women were substantially underrepresented in staff and management. The WID adviser was seen as an outsider by Bank staff, and the policy prescriptions proposed did not fit with the Bank's operational culture. Nevertheless, from 1979-1985 the Bank published its first major gender publication, completed 35 case studies of gender issues in the Bank projects, and held five gender-related workshops for staff (Moser and Moser 2005; Winters et al. 2018).

The Bank's first gender-related directive, which asked staff to consider gender issues in the project cycles, was adopted in 1984. In practice, the directive was merely a guideline and lacked enforcement mechanisms. It also only applied to projects where women were considered important beneficiaries or recipients (Winters et al. 2018). In addition, no training or additional staff resources were provided to help ensure compliance with the new policy. As a result, there were only infrequent applications of newly developing gender policy norms in the projects and operations of the Bank. "According to a 1994 World Bank report, only about 11 percent of the Bank's lending portfolio in the early 1980s contained projects with gender-related action– mostly in rural development, education, and health projects" (Weaver 2010, 75).

[2] The World Bank's gender policy can be succinctly categorized into four main phases: Women in Development (WID) from 1972-1984, the Institutionalization of WID from 1985-1994, Gender Mainstreaming from 1995-2004, and gender equality as 'smart economics' from 2005-present. For a great overview of this history, see Winters et al 2018.

The 1987 re-organization of the World Bank under then-President Barber Conable nearly halted gender mainstreaming. The WID unit was folded into the Population and Human Resources department, and for the next several years gender issues were primarily confined to the social protection sector in Bank operations. Modest progress was made in influencing operational policy changes in 1990 and 1994, highlighting the necessity of including gender equity in the knowledge and preparatory work surrounding World Bank projects. However, monitoring and enforcement of the new directive largely rested in the operational units and progress remained uneven across regions and sectors. As a result, progress on gender mainstreaming remained highly uneven between operational sub-units. The intellectual and operational culture at the Bank further undermined the gender mainstreaming agenda. Gender advocates sought traction in Bank operations by framing gender in terms of economic efficiency—an approach that was amenable to (male) economists who dominated staff and to Bank leaders and who feared that a more explicit acknowledgment of women's development as human rights (a UN framing) would be seen by political principals as violating the Bank's Article IV mandate against using political factors in lending decision. Gender mainstreaming remained largely relegated to "soft social sectors" like health or education (Kenny and O'Donnell 2016).

The 1995 Beijing Fourth World Conference on Women served as a critical "norm tipping point" (Finnemore and Sikkink 1998), spurring two subsequent gender mainstreaming action plans at the Bank in 2002 and 2006. For the first time, gender mainstreaming advocates—mostly women recruited from the UN into the Bank—had visible support from the Bank's top leadership. Then-President James Wolfensohn was a vocal advocate of gender mainstreaming at the Bank and attended the 1995 Beijing Conference. Women's Eyes on the Bank, a coalition of NGOs, used the Beijing conference to mobilize external pressure for gender mainstreaming in the Bank (Weaver 2010). After returning from Beijing, Wolfensohn created an External Gender Consultative Group, composed of fourteen members from various organizations, NGOs, and academic institutions. The newfound focus on gender issues was institutionalized through a reorganization in 1997, under the Strategic Compact reform program, when Gender and Development (GAD) was moved to the newly formed Poverty Reduction and Economic Management thematic network. This new Gender and Development Unit was assigned numerous staff and given the responsibility for "knowledge management, monitoring and reporting on the status of policy implementation, as well as building staff capacity for gender analysis" (Weaver, 2010). Regional officers were also asked to prepare gender action plans, and gender "focal points" were assigned to operational units to advise on integrating women's issues into project goals and activities. The framing of the gender and development agenda continued to avoid any semblance of political overtones by emphasizing the economic dividends of investment in women's socio-economic development—a compromise that many internal advocates at the time argued was critical to sustaining the goodwill of powerful economists (Weaver 2010).

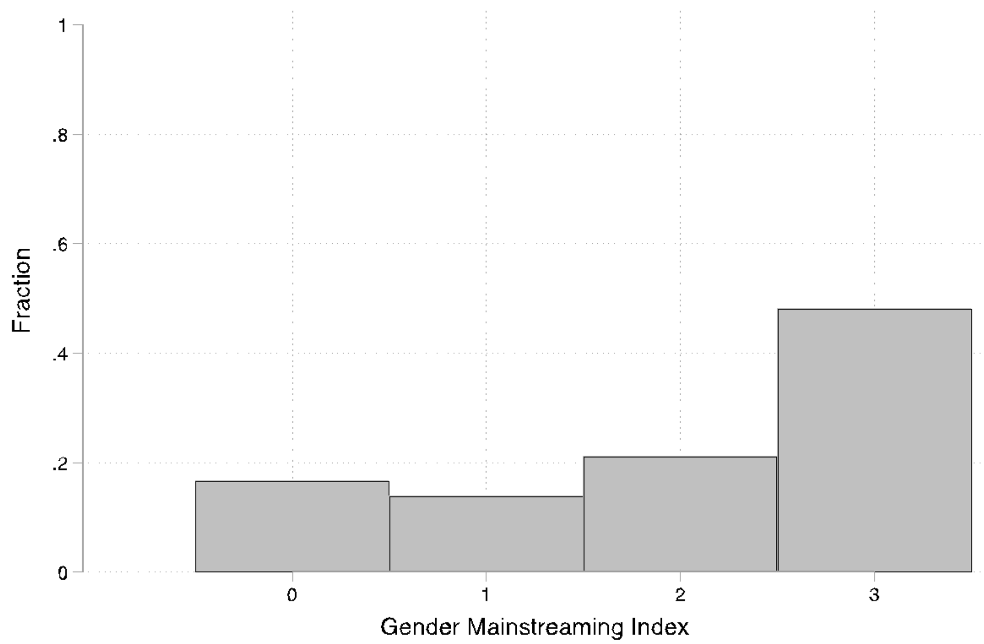
This shift in the attitudes of top leadership and the increase in staff members (mostly women) delegated with the responsibility of promoting gender mainstreaming appears to have had an impact. The proportion of projects that included some consideration of gender issues in their design almost doubled between 1995 and 2001, to nearly 40%. The Bank's operational policy on gender considerations was also expanded during this time to include all programs, not just social sector programs (Winters et al. 2018). However, while organization-wide progress was made, internal application of gender mainstreaming remained uneven within the Bank. GAD's new home in the Poverty Reduction and Economic Management thematic network did provide it with more authority than it had before, but it also limited its ability to influence other networks. Where units lacked significant numbers of female staff and gender advocates, such as Middle Eastern operations, there was little progress. On the other hand, in other units which had more female staff and gender specialists, such as Latin America, gender mainstreaming appears to have progressed at a much faster pace.

The most recent gender policy period embodied an approach labeled “gender equality as smart economics” and began in 2005. According to key informant interviews with numerous gender specialists (all women) within the Bank, gender advocates had enough voice and influence to carve out space for more attention to gender, but not enough to gain the acquiescence of influential decision-makers (mostly men) in the Bank’s middle management, who continued to be concerned that the gender agenda would conflict with the Bank’s apolitical mandate and create tension with key country partners. As a result, while gender advocates had a stronger presence and platform within the Bank, they were compelled to stick with a conceptual framing that was “more palatable to the powerful male economists and the economistic culture of the Bank”.[3] This was clearly reflected in the Bank’s 2007 gender mainstreaming strategy report, *Gender Equality as Smart Economics*, wherein the investment in projects focused on women and girls’ development was justified narrowly as a means of enhancing economic growth and the socio-economic development of children and families. The concept of gender equality as a human right and the explicit language of women’s empowerment were largely omitted. In the 2012 *World Development Report*, likely due to criticism, the Bank reasserted that it believed gender equality was an “objective in its own right,” but continued to use the language of smart economics (Winters et al. 2018).

During the 2014-2017 replenishment of the Bank’s concessional lending arm, President Jim Yong Kim and Bank leaders reinforced the importance of gender within the institution by declaring gender a cross-cutting solution and integrating greater private sector involvement. The most recent gender strategy (2016-2023) finally adopts in a limited fashion the UN framing of women’s rights, voice, and agency as human development, yet continues to emphasize the perceived economic dividends. While the majority of World Bank projects now include some kind of gender component today, the depth and quality of gender mainstreaming across operational units and sectors still varies considerably (Kenny and O’Donnell 2016; Winters et al. 2018). Figure 1 displays the average gender mainstreaming in projects approved by the World Bank between 2009 and 2017 based on a Gender Mainstreaming Index collected by the organization (World Bank 2018). The data (described in more detail in the research design section) illustrates the continued uneven application of gender mainstreaming norms in World Bank projects. Around half of the projects fully incorporate gender mainstreaming. However, the other half only apply gender mainstreaming in some dimension, and nearly twenty percent do not incorporate gender mainstreaming at all.

[3] This section draws heavily from primary interviews conducted within the Bank by one of this paper’s authors in 2009 (see Weaver 2010).

Figure 1: Percentage of all projects approved by the World Bank between 2009 and 2017 that have 0-3 gender mainstreaming components



Where is the Bank today in terms of gender mainstreaming? Internal and external stakeholders have often criticized in very general terms the uneven application of gender mainstreaming in Bank operations. Many gender projects are purportedly still concentrated within the 'soft sectors' and recent studies have shown that Bank rhetoric and publicly released data on gender project financing are inconsistent.

In the remainder of this paper, we attempt to delve deeper into the reasons behind this uneven performance. In doing so, we draw from representative bureaucracy theory the key hypothesis that the uneven attention and resources to gender mainstreaming activities may be linked to the representation of women within the Bank. Simply put, where women have stronger presence within ranks of authority and decision-making, we would expect that this passive representation will be positively correlated with active representation; that is where women call the shots, there will be more evidence of gender mainstreaming at the level of project activities and spending.

Women's Representation and the Application of Gender and Development Norms in the World Bank

In this section, we argue that the pervasive underrepresentation of women in the staff of the World Bank has slowed down and incited unevenness in the application of gender mainstreaming in its lending projects. Three central mechanisms can link women's presence with changes in decision-making outputs in IOs: representation, expertise, and bureaucratic sub-cultures.

First, a representational perspective would argue that women bureaucrats act on behalf of other women to further collective women's interests. Whether women represent other women in IO decision-making depends on whether there is a collective women's interest that can be represented (Sapiro 1981). This question has been debated extensively by feminist scholars (Childs and Krook 2009; Diamond and Hartsock 1981). Many authors have argued that collective interests are based on shared socialization and discriminatory experiences (Diamond and Hartsock 1981; Mansbridge 2005; Sapiro 1981). In contrast, intersectional feminists have highlighted the differential experiences of different groups of women (Acker 2012; Riegle-Crumb and Humphries 2012). These authors have shown that the experiences of women differ markedly and women do not necessarily represent the interests of all women all the time (Atkins and Wilkins 2013). Nevertheless, the representational bureaucracy approach typically takes the former view that women have certain identifiable collective interests in specific policy areas and represent these issues on average (Keiser et al. 2002; Wilkins 2007; Wilkins and Keiser 2006). The argument is that for a policy issue to become important, it must be salient for the demographic group in question (Meier 1992; Wilkins and Keiser 2006). In other words, it depends on whether the policy area has become a gendered policy issue.

Poverty alleviation at the World Bank can be seen as a gendered policy area due to the considerable political efforts to establish gender mainstreaming. Keiser et al. (2002) argue that a policy issue becomes gendered if the policy directly affects women as a group because, in some issue areas, women can work differently with clients of the bureaucracy and because the issue has been gendered in political discussions. Given the World Bank's mandate to alleviate poverty and the extensive gaps in the economic fortunes of different genders globally, it is clear that poverty alleviation directly affects women as a group. However, World Bank bureaucrats are typically not affected by poverty directly, as they are highly educated individuals with high salaries, and many come from high-income countries. Nevertheless, poverty alleviation has been gendered through the World Banks' political process. The repeated initiatives aimed at mainstreaming gender into all sectors and operations of the Bank fulfill the criteria of a definition of gender as a relevant component of World Bank policy-making and implementation and the burden of representation of gender issues within the Bank has often fallen primarily on women.

A second perspective highlights differences in the expertise women and men bureaucrats bring to their work. The argument focuses on differential gendered experiences that shape bureaucrat's knowledge about gender issues. The main observable implication—passive representation of women increases decision-making outcomes for women—does not differ between the first and the second perspective. However, the causal pathway between the passive and active representation of women differs. In the expertise-based perspective, women's distinctive life experiences shape the extent to which they are mindful of gender issues (Boyd, Epstein, and Martin 2010; Voeten 2020). Women will have experienced gender-based discrimination in their careers and are thus more sensitive to the problem than men. Hence, they may be more inclined to look at the distinct impact of their project on women. Women bureaucrats might also seek to build expertise on gendered differences in poverty within their area due to these life experiences. Even in professional fields that are not distinctly focused on gender, women are more likely to emphasize gender issues than men. For example, women economists (co-)author more than 50% of studies on gender in top development economics journals (*American Economic Review*, *Quarterly Journal of Economics*, and *Journal of Development Economics*), while they only (co-) author around 29% of all other articles on average (Jayachandran and Daubenspeck 2018). A greater number of women have built expertise on gender issues throughout their professional life than men. The observable implication of both the representational and the expertise-based perspective is that women bureaucrats would be more inclined to apply gender mainstreaming thoroughly in their decision-making. Hence, the first hypothesis posits:

H1: If women staff oversee projects, these projects will have stronger gender mainstreaming components

A third perspective goes beyond a focus on individual behavior and highlights the impact of greater numbers of women on bureaucratic culture. The argument is that the participation of women in decision-making alters the norms and policy paradigms of men and women working in their bureaucratic sub-units. IOs are organizations that are influenced by shared norms and policy paradigms (Weaver and Nelson 2016). Individuals select into organizations that share their outlook on key policy issues (Hooghe 2005; Momani 2007), and they get socialized into prevailing norms over time (Murdoch et al. 2019). Bureaucratic representation theory argues that increasing numbers of people sharing a certain demographic characteristic change the prevailing policy norms on salient issues. Applied to gender representation, outcomes become less ignorant of women's interest as the aggregate number of women in staff increases relative to men (Kesier et al 2002; Wilkins and Keiser 2006). These patterns emerge because a greater number of women change deliberations on gender issues and women colleagues are more likely to hold men accountable for ignorance towards gender discrimination (Dietrich, Hayes, and O'Brien 2019; Mendelberg, Karpowitz, and Goedert 2014). For example, Karim and Beardsley (2016) show that sexual exploitation and abuse in peacekeeping missions decrease as the number of women peacekeepers increases.

Does this imply that a critical mass of women is needed to affect organizational change? The idea of a "threshold effect" has been controversial in feminist scholarship (Grey et al. 2006). Nonetheless, even without reaching a particular threshold, this literature suggests that bureaucratic sub-cultures evolve to be more mindful of gender discrimination as the number of women working within them increases. In this perspective, the institutionalization of gender mainstreaming norms at the World Bank becomes stronger over time as the underrepresentation of women becomes less stark. Those departments where more women work would also be where gender mainstreaming is more incorporated into decision-making. Crucially, this application of the policy norm would happen irrespective of whether the staff members making the actual decisions are women or men. Hence, a bureaucratic culture perspective would expect policy outputs to incorporate gender mainstreaming more if bureaucratic units have more staff. The corresponding second hypothesis claims:

H2: If more women are working in a bureaucratic unit, projects have stronger gender mainstreaming components

Differences in the organizational culture of IOs create critical scope conditions for the impact of women on policy outcomes. Keiser et al. (2002) highlight hierarchy, stratification, and professionalization as central organizational factors enhancing or suppressing the degree to which passive representation of women translates into active representation. Hierarchy crucially impacts the degree to which staff can shape decision-making (Honig 2019). In hierarchical organizations, where decision-making is centralized at the top, only top-level managers may be able to alter organizational outputs. When hierarchy is more dispersed, women in lower-level positions may possess the necessary autonomy and discretion to affect change within their portfolios (Keiser et al. 2002). Additionally, stratification matters because individual bureaucrats are not acting in a vacuum. Instead, whether their supervisors and managers are sympathetic towards gender mainstreaming may shape the degree to which they are willing and able to affect change in decisions (Keiser et al. 2002). Finally, professionalization creates external sources of policy norms and paradigms. Professional socialization and training affect the policy beliefs of individuals, and these beliefs affect the interpretations and decisions of individual bureaucrats working in IOs (Chwieroth 2010). If an IO is populated with individuals whose professional background does not include a specific gender focus, the translation of passive to active representation is much more difficult (Keiser et al. 2002).

Research Design

To understand whether women staff differ in their focus on gender equality in decision-making, we draw on novel data on the gender of World Bank staff in charge of individual projects and country portfolios. We then regress project-level data on gender mainstreaming on measures of the involvement of women staff.

Dependent variables

In a first step, we focus on gender mainstreaming in World Bank lending. For our dependent variable, we draw on data collected by the World Bank to identify the degree to which nearly 3,000 projects approved between 2009 and 2017 included gender mainstreaming. The data are taken from the World Bank's *Monitoring Gender Mainstreaming in the World Bank Lending Operations Database* (World Bank 2018). The indicator measures the extent to which gender considerations are incorporated into (1) analysis, (2) actions, and (3) monitoring and evaluation of World Bank projects.

The analysis dimension refers to how the World Bank's analysis of the recipient country and sectoral issues is prepared throughout the project and how their stakeholder consultations incorporate gender-related issues. To be coded as one, the project documents need to either specifically analyze gender issues, discuss gender diagnostics, undertake a gender assessment, include results of consultations with gender-focused NGOs, or gender-focused consultations with beneficiaries. If it does not satisfy these criteria, the project is coded as zero. For example, the Uzbekistan Health System Improvement Project explicitly discusses how certain diseases are more prevalent for women than men in Uzbekistan and that the country has exceptionally high maternal mortality rates compared to other countries in Central Asia (World Bank 2020).

The actions indicator requires specific gendered actions in the context of a project. This means that a project needs to include targets that address the needs of women, girls, men, or boys, have gender-specific environmental or social safeguards, or discuss explicitly how the targets address gender disparities. For instance, the Zhejiang Qiantang River Basin Small Town Environment Project finances specific training programs for women and ensures that women and men share equally in compensation contracts (World Bank 2017).

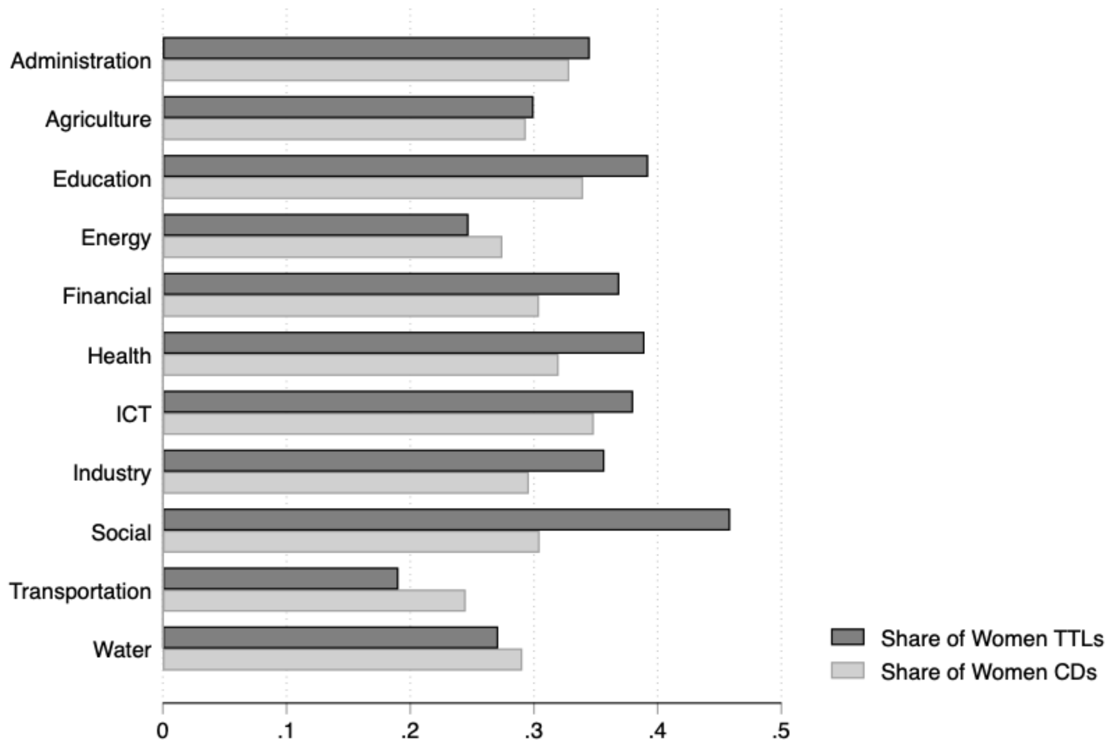
The evaluation dimension measures whether projects include monitoring and evaluation frameworks that specifically focus on gendered outcomes. The indicator is scored as one if projects incorporate gender-disaggregated targets in their result frameworks or gender issues in their evaluation strategies. For example, the Mauritania Mining Sector Capacity Building Project set targets of 30% women in vocational training programs and 70% women as micro-grant beneficiaries (World Bank 2005). World Bank staff score a project on whether it includes gender considerations into each of the three dimensions. The resulting scores are then added to an index that ranges from 0 (no gender mainstreaming in any dimensions) to 3 (gender mainstreaming in all dimensions).

Independent variables

To measure the gender of individuals in charge of specific projects and country portfolios, we extracted the names of World Bank staff from different sources. Our analysis focuses on the primary staff in charge of decision-making on individual projects. Specifically, we collect data on Country Directors and Task Team Leaders (TTLs). Country Directors are in charge of the overall project portfolio for a given country (Honig 2020; Weller and Yi-chong 2010). They are also supposed to play an essential role in ensuring that gender issues are considered within the portfolio (Kenny and O'Donnell 2016). Task Team Leaders are the main staff members in charge of individual projects and are generally considered the most critical staff affecting decision-making in individual World Bank projects (Briggs 2021; Bulman, Kolkma, and Kraay 2017; Denizer, Kaufmann, and Kraay 2013). The names of Country Directors and Task Team Leaders were extracted from the World Bank website and World Bank documents. First, the names of Country Directors were collected by hand from project approval documents, country assistance strategies, and the World Bank website. Second, the names of TTLs in charge of implementation were scraped from the publicly available World Bank API. However, World Bank TTLs rotate every 3-7 years across duty stations, which means that staff in charge of approval is often different from staff in charge of implementation (Heinzel and Liese 2021). Therefore, we extracted the names of the TTLs in charge of approval from project approval documents available from the World Bank website. Through these procedures, we attained the names of 3,619 TTLs overseeing 86% (8,506) of World Bank projects between 2000 and 2020 as well as 196 Country Directors in charge of more than 95% of the project portfolio of the World Bank during the same time.

To classify individuals' binary gender, we combine automated methods and hand-coding. In line with recent data collection of individual-level data (Nyrup and Bramwell 2020), we predicted the gender of staff using *genderize.io,* which classifies the gender of individuals based on millions of self-reported names and gender from social media profiles. To validate the data, we hand-coded the gender of 981 World Bank management named in World Bank annual reports based on their use of gendered pronouns. The algorithm correctly classified 92.6% of these staff members. However, the correct classification decreases to 65% when considering only individuals with less than 75% of social media profiles listed as one particular gender (e.g., Andrea). To increase the accuracy of our measurement of gender further, we hand-code every individual where *genderize.io* reports a distribution of less than 75% or fails to classify the name. When doing so, we can increase the accuracy of our gender variable to around 98-99%. Figure 2 displays the share of women staff in charge of projects by different sectors. The highest share of women TTLs work on Social, Education, and Health projects, while the least women work in Transportation and Energy projects.

Figure 2: Women staff in different sectors



Control variables

We employ a range of control variables to minimize the prospect of omitted variable bias. The biggest threat to inference is that projects would be pre-determined by someone else to incorporate gender mainstreaming, and women staff is simply selected into these projects more often. Hence, we seek to control for alternative sources for gender mainstreaming.

Three groups of actors are involved in World Bank decision-making: donors, recipients, and the World Bank. The first group of actors are donor principals. Research has repeatedly shown that the World Bank is influenced by the preferences of its major shareholders (Clark and Dolan 2021; Lyne, Nielson, and Tierney 2009). We would face reverse causality if gender mainstreaming was pushed more by certain World Bank shareholders and women staff preferred working in these projects. In this case, a greater gender mainstreaming focus would be caused by the preferences of shareholders and would cause the selection of women staff into the project. Therefore, we control for the average gender-focused lending of the five most important World Bank shareholders (USA, UK, Germany, France, and Japan). Data on gender-focus is taken from the OECD Creditor Reporting System (OECD 2020). The indicator is calculated by taking the average share of projects the five donors approved in a given year for a recipient with a gender marker.

Additionally, heterogeneity in gender mainstreaming could be driven by differential gender equality in the recipient country. Gender equality could impact both the degree to which recipients demand gender mainstreaming in operations (Weaver 2010) and the urgency the World Bank places on addressing gender inequality. It could also shape the extent to which women staff gets sent into these countries. For example, many countries the World Bank works with are characterized by considerable suppression of women's rights. Such misogyny might deter (or incentivize) women to work in these contexts and affect whether recipients demand loans with a gender focus. We control for four variables that indicate the situation of women in the recipient country. First, we include a measure of women's economic rights to control for the de jure economic situation of women. The indicator is taken from Cingranelli and Richards (2010). Second, we control for the political representation of women in the recipient government by including the share of women ministers in the national government (Nyrup and Bramwell 2020). Third, we control for the socioeconomic fortunes of women by including a measure of the infant mortality of girls and the share of women in vulnerable employment (WDI 2020). Fourth, we control for GDP per capita and population (log) to account for differences in the level of development of recipient countries and differences between countries with larger and smaller populations.

The final set of control variables focuses on differences in projects that may affect the World Bank's decision to incorporate gender mainstreaming on the World Bank side. First, we control for whether the project is an IDA or an IBRD loan and whether the World Bank has listed the country on its list of conflict-affected and fragile countries because the World Bank's policy paradigms differ for these contexts. More difficult contexts have been shown to affect the selection of staff (Limodio 2021). Second, we include a measure of the project amount because projects with higher amounts are seen as more prestigious by Bank staff (Briggs 2021). More prestigious projects generally have more attention from senior management and, consequently, staff members on the ground might have less discretion than in smaller projects. In addition to these control variables, we use various fixed effects to minimize omitted variable bias, which are discussed in more detail when introducing the models below.

Do women oversee more gender-focused operations?

We estimate several regression models to grasp the impact of passive representation on gender mainstreaming in World Bank projects. The unit of analysis is the individual World Bank project. All models cluster standard errors to allow for correlated errors at the country-level. Table 2 displays the results from models regressing gender mainstreaming ratings on the gender of TTLs (Models 1, 2, and 4) and Country Directors (Models 3 and 4). In Model 1, we evaluate the role of a TTLs' gender while holding all control variables constant and including country, sector, and approval year fixed effects. However, one could argue that sector-specific effects, like policy norms or paradigms, change over the time of examination. Similarly, we might omit critical variables at the country-level, like the interests of recipient governments. Hence, Model 2 provides a more stringent test by using country-year and sector-year fixed effects. We drop the country-level control variables from Model 2, as they only vary at the country-year level and are, thus, incorporated in the fixed effects. In Model 3, we re-estimate Model 1 but substitute our TTL variable with our indicator for women Country Directors. Finally, we include both staff indicators on the right-hand side of the equation in Model 4.

The results indicate that women's passive representation in key positions of authority increases the degree to which their projects incorporate gender mainstreaming. The coefficient for women TTL is statistically significant and positive ($p<0.05$) in Model 1 and remains marginally significant when using both country-year and sector-year fixed effects in Model 2 ($p<0.1$). The coefficients are moderate—having at least one woman among TTLs is associated with an average 0.12 increase in gender mainstreaming ratings (on a four-point scale). The coefficient is similar in magnitude to the weight of a 25% increase in the economic rights of women in Model 1. Country Directors are likewise linked with an increase in gender mainstreaming ratings. When the Country Director is a woman, gender mainstreaming ratings of projects increase on average by 0.22 in Model 3 ($p<0.01$). The findings reported in Model 4 indicate that the gender of Country Directors has a more substantial association with gender mainstreaming ratings than the gender of TTLs. The coefficient is nearly double in magnitude, although the 95% confidence intervals overlap slightly. These results imply that the position of women in the organizational hierarchy plays a considerable role when determining their impact on gender mainstreaming in the World Bank.

Table 2: regressing gender mainstreaming on women staff

	(1)	(2)	(3)	(4)
Women TTLs	0.1235*	0.0860^{+}		0.1191*
	(0.0501)	(0.0449)		(0.0491)
Women Country Directors			0.2234**	0.2060**
			(0.0703)	(0.0731)
IDA	0.2805*	0.4172**	0.3304**	0.2962*
	(0.1247)	(0.1323)	(0.1234)	(0.1199)
Amount (log)	0.0556^{+}	0.0911**	0.0562^{+}	0.0537^{+}
	(0.0288)	(0.0283)	(0.0295)	(0.0295)
Women ministers	0.2976		0.0078	0.0571
	(0.4914)		(0.4779)	(0.4740)
Women economic rights	0.1381^{+}		0.1141	0.0938
	(0.0739)		(0.0743)	(0.0739)
Women infant mortality	0.0057		0.0028	0.0096
	(0.0122)		(0.0127)	(0.0141)
Women vulnerable employment	-0.0145		-0.0110	-0.0060
	(0.0126)		(0.0125)	(0.0121)
Principals gender lending	0.1532		0.1127	0.1517
	(0.2771)		(0.2710)	(0.2644)
Post-conflict country	-0.3388^{+}		-0.3207	-0.3743^{+}
	(0.2018)		(0.1968)	(0.1942)
GDP per capita	0.0003*		0.0002*	0.0002
	(0.0001)		(0.0001)	(0.0001)
Population (log)	-3.0533*		-3.5027**	-2.9872*
	(1.3456)		(1.2987)	(1.3092)
Country fixed effects	Yes	No	Yes	Yes
Country-year fixed effects	No	Yes	No	No
Sector fixed effects	Yes	No	Yes	No
Sector year fixed effects	No	Yes	No	Yes
Year fixed effects	Yes	No	Yes	No
Observations	2088	2205	2122	2084
R^2	0.384	0.583	0.382	0.443

Clustered standard errors in parentheses; $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

So far, our models proceeded under the assumption of exogeneity conditional on covariates. Our tests controlled for relevant factors on the country- or sector-level. That way, we were able to minimize the likelihood that staff select in certain countries or sectors in which gender projects are also more likely. However, we still might face endogeneity on the project-level. Such endogeneity would occur in two scenarios: first, management could pre-determine that a project should have a greater gender focus and select women to run this project based on gender stereotypes. Second, women staff might be more likely to apply internally for projects with a greater gender focus and, hence, self-select into these projects. In these cases, the association between women and gender mainstreaming ratings would be driven by women selecting into projects rather than designing projects more in line with gender mainstreaming norms. Such pre-determined gender mainstreaming focus is largely unobservable (see robustness check section for alternatives to address this issue). In other words, the coefficients reported in Table 2 would reflect reverse causality. Selection based on pre-determined project objectives is arguably less of an issue for Country Directors, who oversee large, often multi-country, project portfolios and certainly do not select into countries based on specific projects that have not even started when they are hired. However, reverse causality is a considerable concern for TTLs that oversee particular projects.

To address reverse causality, we use an instrumental variable approach in Table 3. The approach utilizes an instrumental variable that is valid if it predicts whether at least one of the TTLs in a project was a woman (relevance criterion) but must not affect gender mainstreaming ratings through any other channel than through increasing the likelihood of selecting at least one woman as a TTL. We use the number of TTLs listed on a project in our database as an instrument.

The instrument is relevant because with every additional TTL, the chances that at least one woman works in a project increase. Women make up between 25% and 45% of staff in our period of interest. Thus, the more hiring decisions are made, the more likely is the selection of women among TTLs because of the supply of candidates. Furthermore, gender balancing norms imply that World Bank practice managers (who make the hiring decisions) should become increasingly likely to select women as TTLs the more selection decisions they make to avoid the impression of misogyny in hiring.

The instrument is also plausibly excludable, conditional on control variables because theoretical arguments do not imply that projects with larger numbers of TTLs lead to a stronger gender mainstreaming focus—except through the greater likelihood that at least one woman is working on the project. One could question the excludability of the instrument from two main angles: first, one could argue that larger projects tend to have more TTLs, and the increased scrutiny in these projects could lead World Bank management to try to ensure that gender mainstreaming is incorporated. However, we account for this argument by controlling for the (log) project amount. A second point of contention could be that sectoral norms on the cooperation of TTLs differ. These sectors could also be where women are more likely staffed, and gender mainstreaming is more prevalent (like social protection or education). Nevertheless, these factors should be controlled for by employing sector or sector-year fixed effects. Hence, we argue that the number of TTLs on a project is a valid instrument for the purposes of this analysis.

Table 3 displays the results from three instrumental variable regressions. Model 5 includes all control variables, country, sector, and approval year fixed effects. In Model 6, we substitute sector with sector-year fixed effects. Finally, Model 7 uses both country-year and sector-year fixed effects. The results further strengthen our confidence in the relationship between the passive representation of women and the application of gender mainstreaming norms in World Bank projects. The coefficient is positive and marginally significant ($p<0.1$) in Model 5 as well as positive and statistically significant in Model 6 and Model 7 ($p<0.05$). Overall, our analysis provides strong evidence that the null hypothesis that passive representation of women and the application of gender mainstreaming norms in World Bank projects are unrelated can be rejected. Therefore, more women in World Bank staff seem to increase whether gender mainstreaming norms are applied in the projects under their authority (Hypothesis 1).

Table 3: instrumental variable estimation of the influence of women TTL on gender mainstreaming

	(5)	(6)	(7)
Women TTL (instrumented)	0.2541^{+}	0.4269^{**}	0.4621^{**}
	(0.1412)	(0.1473)	(0.1530)
IDA	0.2750^{*}	0.2553^{*}	0.4207^{**}
	(0.1232)	(0.1294)	(0.1378)
Amount (log)	0.0535^{+}	0.0476^{+}	0.0829^{**}
	(0.0285)	(0.0274)	(0.0275)
Women ministers	0.2780	0.2501	
	(0.4916)	(0.4242)	
Women economic rights	0.1402^{+}	0.1199^{+}	
	(0.0732)	(0.0689)	
Women infant mortality	0.0044	0.0058	
	(0.0118)	(0.0116)	
Women vulnerable employment	-0.0142	-0.0057	
	(0.0126)	(0.0114)	
Principals gender lending	0.1687	0.2291	
	(0.2757)	(0.2118)	
Post-conflict country	-0.3321^{+}	-0.3812^{**}	
	(0.1976)	(0.1280)	
GDP per capita	0.0002^{+}	0.0002	
	(0.0001)	(0.0001)	
Population (log)	-3.1840^{*}	-2.8583^{*}	
	(1.3418)	(1.1409)	
Country fixed effects	Yes	Yes	No
Country-year fixed effects	No	No	Yes
Sector fixed effects	Yes	No	No
Sector-year fixed effects	No	Yes	Yes
Year fixed effects	Yes	No	No
Observations	2089	2089	2205

Clustered standard errors in parentheses; $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

We now turn to Hypothesis 2, which posited that an increased number of women in staff also changes the overall culture in bureaucratic units. That way, greater numbers of women would impact not only the decisions in their discretion but also those taken by staff with other gender identities that work with them. World Bank TTLs are sectoral staff members who answer to sectoral practice managers and are incorporated in a sectoral organizational structure. Therefore, the World Bank sector is the relevant peer group for TTLs. Hence, we use an additional independent variable that calculates the share of women appointed in a given sector in all other projects approved within three years of the project of interest. The specifications mimic the models discussed above with one substantial modification. One concern is that the variable picks up organizational changes towards gender equality in the sub-unit more generally. These organizational changes could drive both whether women are appointed and the extent to which staff in that sub-unit incorporate gender mainstreaming. Hence, we control for the gender mainstreaming rankings in the same projects within three years.

Table 4 reports the results from four models focusing on tests of Hypothesis 2. Model 8 includes all control variables, country, sector, and approval year fixed effects. In Model 9, we include country-year fixed effects and exclude all country-level variables. We cannot use sector-year fixed effects in these models because our variable of interest does not vary considerably on the sector-year level. Finally, in Models 10 and 11, we replicate Model 8 but further include our measure of Women TTLs and Women Country Directors.

The regression models again lend support to our expectation that passive representation shapes the application of gender mainstreaming norms in the World Bank. The coefficient for Women appointed in the sector within three years of the project of interest is positive and statistically significant ($p<0.05$) in all four models. One standard-deviation (0.09) change in the share of women in a sector is comparable to the coefficient of having at least one Woman TTL overseeing the project of interest. If all TTLs in a given sector were women, the models estimate that gender mainstreaming scores would increase on average by around 1.17 (on a four-point scale). The TTL and Country Director variables similarly hold, even when controlling for the share of women appointed in a sector and the average gender mainstreaming rating in the sector. Overall, the results imply that hiring more women staff seems to alter the behavior of staff working in the sector more generally.

Table 4: OLS-models regressing gender mainstreaming on average women appointed in sector

	(8)	(9)	(10)	(11)
Women appointed in sector	1.1789*	1.1626*	1.2205*	1.2099*
(within 3 years)	(0.5556)	(0.4874)	(0.5589)	(0.5548)
Women TTL			0.1312**	0.1297*
			(0.0489)	(0.0494)
Women Country Director				0.1919**
				(0.0701)
Gender Mainstreaming in sector	0.6518***	0.5296***	0.6580***	0.6439***
(within 3 years)	(0.1012)	(0.1162)	(0.1046)	(0.1050)
IDA	0.3056*	0.4850***	0.2891*	0.3179**
	(0.1222)	(0.1127)	(0.1205)	(0.1193)
Amount(log)	0.0549^{+}	0.0937**	0.0540^{+}	0.0544^{+}
	(0.0295)	(0.0278)	(0.0285)	(0.0283)
Women ministers	0.2915		0.2691	0.0199
	(0.4738)		(0.4745)	(0.4678)
Women economic rights	0.1307^{+}		0.1273^{+}	0.1052
	(0.0744)		(0.0725)	(0.0710)
Women infant mortality	0.0039		0.0083	0.0087
	(0.0128)		(0.0122)	(0.0125)
Women vulnerable employment	-0.0074		-0.0100	-0.0092
	(0.0124)		(0.0124)	(0.0122)
Principals gender lending	0.1543		0.1590	0.1253
	(0.2751)		(0.2742)	(0.2698)
Post-conflict country	-0.3723^{+}		-0.3722^{+}	-0.3539^{+}
	(0.2024)		(0.1983)	(0.1904)
GDP per capita	0.0002*		0.0002^{+}	0.0002^{+}
	(0.0001)		(0.0001)	(0.0001)
Population (log)	-2.6552^{+}		-2.7346*	-3.2218*
	(1.3395)		(1.3113)	(1.2579)
Country fixed effects	Yes	No	Yes	Yes
Country-year fixed effects	No	Yes	No	No
Sector fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	No	Yes	Yes
Observations	2127	2252	2088	2083
R^2	0.396	0.551	0.401	0.403

Clustered standard errors in parentheses; $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

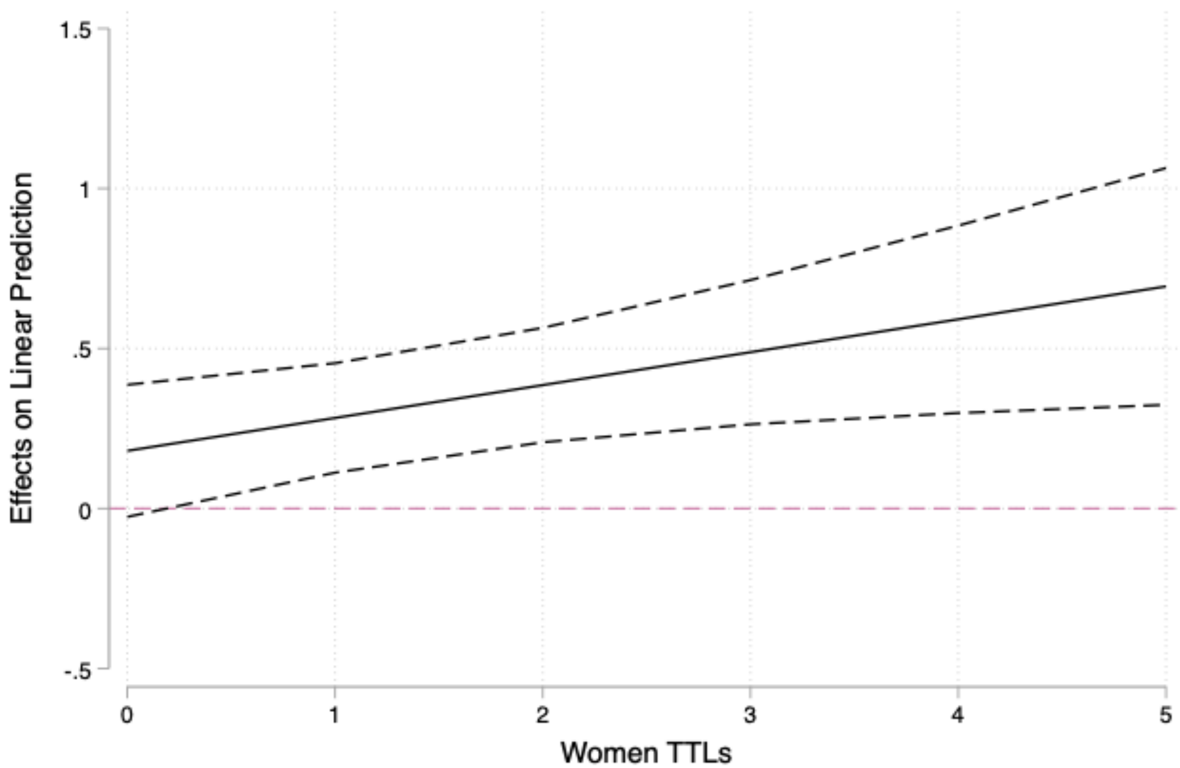
So far, our results show that passive representation seems to increase the degree to which the World Bank considers gender mainstreaming in its operations. In a final step, we now evaluate whether the passive representation of women also translates into improved performance for gender equity in recipient countries. Our extension draws on data on the evaluations of more than 55,000 stakeholders collected in the context of the World Bank Country Opinion Survey Programs. Data stem from 299 country surveys conducted between 2012 and 2020 in more than 130 countries. The World Bank identifies key stakeholders in government, civil society, academia, business, implementing agencies, and the media to conduct the surveys. As such, surveys should not be seen as representative of the broader population in recipient countries. However, they provide an important snapshot into the views of those stakeholders the World Bank Public Opinion Research Group deems crucial in the respective recipient countries. We do not have strong assumptions on the distribution of the population of stakeholders and, consequently, do not weigh responses. The dependent variable of interest for our study is a question asking to what extent the World Bank effectively addresses gender equity in the recipient country. Respondents identify the primary sector of their work. Hence, we can build an aggregated database on the country-sector-year level that allows us to link the gender of staff working in these country-sectors to the perceptions of key stakeholders.

Gender mainstreaming implies the consideration of gender issues in the overall work of the World Bank. As discussed extensively above, the idea is to ensure that the benefits stemming from World Bank projects are not withheld from women in recipient countries. Therefore, effective gender mainstreaming would imply that the effectiveness of the World Bank's work translates into projects that effectively address gender equity. The corresponding hypothesis would posit that women staff maximize the gender effectiveness of projects. To mimic this theoretical expectation in our models, we interact perceptions of the overall effectiveness of the World Bank in a given country-sector-year with the number of Women TTLs that have worked in this country-sector within three years. Therefore, we test whether larger numbers of women staff can better translate World Bank country-sector operations deemed effective by key stakeholders into perceptions of effectiveness for gender equity.

Table 5 displays the results from the analysis of the perceptions of the World Bank's effectiveness in addressing gender equity. Since most countries have only a few country surveys and we likely face many unobservable idiosyncrasies at the level of the country survey, we employ country-year fixed effects in all models. That way, we only compare respondents' perceptions across sectors based on data from the same surveys. In Model 12, incorporate the interaction, controlling for country-year and sector fixed effects and a range of control variables. For example, we control for the number of TTLs to ensure that results are not simply driven by larger numbers of staff working in a country-sector. Additionally, the models hold the share of IDA projects and the average project amount constant. We also control for gender mainstreaming rankings in the past three years. Model 13 further contains sector-year fixed effects. In Models 14 and 15, we further include four control variables taken from the surveys. First, to ensure that we are not simply capturing the degree to which recipients perceive the World Bank's work as aligned with their preferences, we control for perceptions of alignment of the World Bank with recipient's views of the main development priorities for their country. Second, we hold constant the share of respondents that are (self-reported) experts for gender equity to account for heterogeneity in the expertise of respondents. Finally, we hold differences in respondents' relationship with the World Bank constant by controlling for perceptions of familiarity with the World Bank's work and the share of respondents that collaborated with the World Bank directly at the time of the survey.

The coefficient for the interaction between the number of women staff and the effectiveness ratings of the World Bank is positive and statistically significant in three of the four models (at least $p<0.05$). This finding implies that the more women that are staffed in a particular country-sector, the more the World Bank can translate effectively perceived operations into operations that are also perceived as effective in addressing gender equity. Figure 2 illustrates the interaction based on Model 14. The figure shows that an increase from zero to five women staff in a country sector nearly triples the association of overall effectiveness perceptions with effectiveness perceptions for gender equity.

Figure 3: Linear predictions of gender performance based on Model 14, 95% Confidence Intervals



Note: Plot shows the interaction between Women TTLs and perceptions of effective World Bank projects (with 95% confidence intervals). The association of perceptions of the effectiveness of the World Bank in general with perceptions of effectiveness in addressing gender equality increases with each additional Women TTL staffed in a country-sector.

Table 5: Passive representation and stakeholder perceptions of World Bank gender performance

	(12)	(13)	(14)	(15)
Women TTLs (past 3 years)	-0.4575	-0.5379^{+}	-0.5543^{+}	-0.6664*
	(0.3058)	(0.3205)	(0.2922)	(0.3001)
Performance	0.0762^{+}	0.0877^{+}	0.0874*	0.1033*
	(0.0452)	(0.0482)	(0.0433)	(0.0457)
Women TTLs # Performance	0.3586***	0.3229**	0.2018*	0.1701
	(0.0905)	(0.0977)	(0.1021)	(0.1062)
TTL count	0.0021	0.0046	0.0036	0.0063
	(0.0151)	(0.0171)	(0.0141)	(0.0164)
Project amounts (log)	-0.0141	-0.0099	-0.0072	0.0056
	(0.0725)	(0.0759)	(0.0704)	(0.0735)
IDA share	0.0082	-0.0284	0.0116	-0.0340
	(0.1250)	(0.1301)	(0.1223)	(0.1272)
Gender Mainstreaming	-0.0010	-0.0052	0.0032	-0.0051
	(0.0541)	(0.0604)	(0.0572)	(0.0640)
Alignment			0.2513**	0.2845***
			(0.0767)	(0.0784)
Gender experts			-1.9449***	-1.6991*
			(0.5448)	(0.8296)
Collaboration			-0.4135	-0.4359
			(0.2776)	(0.2718)
Familiarity			-0.0213	-0.0863
			(0.0813)	(0.0784)
Country-year fixed effects	Yes	Yes	Yes	Yes
Sector fixed effects	Yes	No	Yes	No
Sector-year fixed effects	No	Yes	No	Yes
Observations	843	834	842	833
R^2	0.493	0.540	0.509	0.558

Clustered standard errors in parentheses; $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

To summarize, the quantitative analysis has yielded three key results. First, women staff members (Country Directors and TTLs) seem to run projects with stronger gender mainstreaming components. Second, when more women staff are appointed to a sector, the extent of gender mainstreaming increases in projects of that sector more generally. Third, when more women are appointed, the translation of World Bank performance into the effectiveness of the World Bank in addressing gender inequality increases in the eyes of key stakeholders.

Robustness checks and extensions

In addition to the robustness checks reported throughout the article's main body, we perform additional analyses. First, we further aim to minimize the danger of reverse causality. The World Bank categorizes projects by (non-exhaustive) themes. One of these themes is gender. These projects focus directly on gender equality as one of the main goals of the project. Since this paper focuses on gender mainstreaming—incorporating a gender lens in all operations—rather than only gender projects, we did not focus on the gender theme here. However, the gender theme can serve as a useful control variable to minimize reverse causality. While the specific design of projects, including gender mainstreaming components, is often not determined yet when TTLs are hired for projects, the project's overall goals are much more set. If women staff select into projects with a greater gender focus, we would most likely see this pattern emerge for projects with a gender theme. Therefore, we control for a binary indicator that measures whether a given project had a gender theme. The results are robust to the inclusion of this additional control variable.

Second, the discretion of staff varies in different types of loans. While TTLs have considerable discretion in investment policy loans, their impact is much lower in development policy loans that include prior action conditionality negotiated at a higher political level. Indeed, research has shown that TTLs have a lower impact on DPF compared to IPF (Heinzel and Liese 2021). Consequently, we would expect that women TTLs fail to impact gender mainstreaming in DPF, while women Country Directors impact gender mainstreaming in both IPF and DPF. When disaggregating the sample by loan types, we indeed find this pattern. The coefficient for TTLs fails to attain statistical significance at any conventional threshold for the DPF sub-sample, while the coefficient for women Country Directors remains positive and statistically significant.

Third, we aim to ensure that idiosyncrasies with our model specifications do not drive our results. The gender mainstreaming indicator may be more accurately modeled as count data, as each component is scored and the indicator is the sum of these scores. Hence, we re-estimate models using Poisson-Pseudo-Maximum-Likelihood to estimate count data with a large number of fixed effects.

Fourth, we test the robustness to using alternative dependent variables. Hence, we disaggregate the indicator and estimate separate linear probability models for each of the three components. We also use a dependent variable focusing simply on the presence of any gender mainstreaming components instead of the additive indicator. The findings show that projects overseen by women are not necessarily more likely to include gender mainstreaming per se. The results are mixed when using the binary indicator. Instead, the depth of commitment to gender mainstreaming seems to increase when women oversee operations. This finding is consistent with expectations on de-coupling in the literature. With the increasing institutionalization of gender mainstreaming as a policy norm, the expectation that all staff incorporate it into operation increases.

However, women staff members seem to treat gender mainstreaming less as a box-checking exercise. They seem to increase the depth of commitments to gender mainstreaming in World Bank projects.

Conclusion

The literature on IOs often cannot explain uneven patterns in policy norm mainstreaming because analyses rarely account for the factors internal to organizational cultures that may be tied directly to staff composition. We address this gap by highlighting an explanation based on bureaucratic representation theory. Individuals in IO staff can possess considerable autonomy and discretion, and under these conditions, pursue interests and goals aligned with their identities. We develop this argument based on the case of the World Bank and gender mainstreaming. Empirically, we presented considerable evidence that implies that such active representation is present and meaningful: women staff in the Bank incorporate gender mainstreaming more in projects, a greater number of women in organizational sub-units increases the gender-focus of decisions taken within this unit, and greater representation of women increases the extent to which IOs can translate positive performance evaluations into being perceived by external principals as effective in addressing gender equity.

Together these findings imply that more attention should be paid to the workforce composition of IOs and the staff representation of key demographic attributes (including not only gender but also nationality, ethnicity, or disciplinary expertise). To understand decisions taken in IOs, one needs to go beyond studying the interests of member states or the collective interests of IO bureaucracies. The background of individual decision-makers within IO staff seems to affect decisions and needs to be appreciated more fully if one aims to understand decision-making in IOs.

In turn, our research also has important real-world implications. Given the pervasive gendered poverty gaps and the demonstrated impact that women staff seem to have on the gender focus of projects, hiring and placing more women into key decision-making positions seems crucial to achieving the Bank's gender and development goals. The patterns of representation we describe thus call for increased scrutiny on the hiring and promotion decisions of IOs (and likely other organizations). It is essential for IOs, like the World Bank, to reach gender parity in their workforce composition – not a merely symbolic representation on executive boards or as a minority percentage of the overall staff. Moreover, we must acknowledge that the highly skewed representation of nationals among the staff of IOs like the World Bank, which is dominated by staff from high-income countries, likely also means that the voices of women from countries where the World Bank runs most of its projects remain marginalized in the organization.

Additionally, hiring more women should not be treated as a silver bullet. Studies of the World Bank's gender mainstreaming have repeatedly highlighted that the framing of gender goals within the development context can easily be skewed by factors other than gender, including the dominance of orthodox economic ideas and perceived political resistance (Berik 2017; Elson 2009; Ferguson and Harman 2015; Weaver 2010). These limitations cannot be overcome by simply increasing the number of women in their staff. Instead, the World Bank would need to hire and strategically place staff members with explicit gender expertise that can problematize these dynamics and highlight how the World Bank's gender mainstreaming strategy can go beyond simplistic conceptions of the origins and consequences of gender inequality.

Bibliography

Abels, Gabriele, and Joyce M. Mushaben. 2020. "Great Expectations, Structural Limitations: Ursula von Der Leyen and the Commission's New Equality Agenda." *Journal of Common Market Studies* 58(S1): 121–32.

Acker, Joan. 2012. "Gendered Organizations and Intersectionality: Problems and Possibilities." *Equality, Diversity and Inclusion* 31(3): 214–24.

Arora-Jonsson, S. and B.B. Sijapati. 2018. "Disciplining Gender in Environmental Organizations: The Texts and Practices of Gender Mainstreaming," Gender, Work & Organization

Atkins, Danielle N., and Vicky M. Wilkins. 2013. "Going beyond Reading, Writing, and Arithmetic: The Effects of Teacher Representation on Teen Pregnancy Rates." *Journal of Public Administration Research and Theory* 23(4): 771–90.

Badache, Fanny. 2020. "A Representative Bureaucracy Perspective on Workforce Composition in International Organizations: The Case of the United Nations Secretariat." *Public Administration* 98(2): 392–407.

Barnett, Michael N, and Martha Finnemore. 2004. *Rules for the World. International Organizations in Global Politics.* Cornell: Cornell University Press.

Barraza Vargas, Constanza. 2019. "Women in Command: The Impact of Female Leadership on International Organisations." *Global Society* 33(4): 541–64. https://doi.org/10.1080/13600826.2019.1642858.

Bayerlein, Louisa, Christoph Knill, and Yves Steinebach. 2020. *A Matter of Style. Organizational Agency in Global Public Policy.* Cambridge: Cambridge University Press.

Betz, Timm, David Fortunato, and Diana Z. O'Brien. 2021. "Women's Descriptive Representation and Gendered Import Tax Discrimination." *American Political Science Review* 115(1): 307–15.

Birken, Marie-Anne and Gian Piero Cigna. 2018. "Gender Diversity on Boards: A Cause for Multilateral Organizations," in *Good Governance and Modern Financial Institutions*. Beijing: Asian Infrastructure Investment Bank; 25-43.

Blackmon, Pamela. 2020. "The Lagarde Effect: Assessing Policy Change Under the First Female Managing Director of the International Monetary Fund (IMF)." *Global Society* First view: 1–20. https://doi.org/10.1080/13600826.2020.1763925.

Boyd, Christina L., Lee Epstein, and Andrew D. Martin. 2010. "Untangling the Causal Effects of Sex on Judging." *American Journal of Political Science* 54(2): 389–411.

Briggs, Ryan C. 2021. "Why Does Aid Not Target the Poorest?" *International Studies Quarterly*: 1–14.

Bulman, David, Walter Kolkma, and Aart Kraay. 2017. "Good Countries or Good Projects? Comparing Macro and Micro Correlates of World Bank and Asian Development Bank Project Performance." *Review of International Organizations* 12: 335–63.

Childs, Sarah, and Mona Lena Krook. 2009. "Analysing Women's Substantive Representation: From Critical Mass to Critical Actors." *Government and Opposition* 44(2): 125–45.

Christensen, Johan. 2020. "Representative Bureaucracy, International Organizations and Public Service Bargains." *Public Administration* 98(2): 408–23.

Chwieroth, Jeffrey M. 2010. *Capital Ideas: The IMF and the Rise of Financial Liberalization*. Princeton: Princeton University Press.

———. 2013. "'The Silent Revolution:' How the Staff Exercise Informal Governance over IMF Lending." *Review of International Organizations* 8(2): 265–90.

Cingranelli, David L., and David L. Richards. 2010. “The Cingranelli and Richards (CIRI) Human Rights Data Project.” *Human Rights Quarterly* 32(2): 401–24.
Clark, Richard, and Lindsay R. Dolan. 2021. “Pleasing the Principal: U.S. Influence in World Bank Policymaking.” *American Journal of Political Science* 65(1): 36–51.
Denizer, Cevdet, Daniel Kaufmann, and Aart Kraay. 2013. “Good Countries or Good Projects? Macro and Micro Correlates of World Bank Project Performance.” *Journal of Development Economics* 105: 288–302.
Diamond, Irene, and Nancy Hartsock. 1981. “Beyond Interests in Politics: A Comment on Virginia Sapiro’s ‘When Are Interests Interesting? The Problem of Political Representation of Women.’” *American Political Science Review* 75(3): 717–21.
Dietrich, Bryce J., Matthew Hayes, and Diana Z. O’Brien. 2019. “Pitch Perfect: Vocal Pitch and the Emotional Intensity of Congressional Speech.” *American Political Science Review* 113(4): 941–62.
Dolan, Julie. 2009. "Representative Bureaucracy and Gender Consciousness: A Framework for Investigating Gendered Policy Outputs." APSA 2009 Conference Paper, Available at https://ssrn.com/abstract=1450767.
Eckhard, Steffen. 2021. “Bridging the Citizen Gap: Bureaucratic Representation and Knowledge Linkage in (International) Public Administration.” *Governance* 34(2).
Eckhard, Steffen, and Jörn Ege. 2016. “International Bureaucracies and Their Influence on Policy-Making: A Review of Empirical Evidence.” *Journal of European Public Policy* 23(7): 960–78.
Grey, Sandra et al. 2006. “Do Women Represent Women? Rethinking the ‘Critical Mass’ Debate.” *Politics and Gender* 2(1): 119–28.
Griffin, Penny. 2009. *Gendering the World Bank : Neoliberalism and the Gendered Foundations of Global Governance*. Palgrave Macmillan.
Finnemore, Martha and Kathryn Sikkink. 1998. “International Norm Dynamics and Political Change.” *International Organization* 52(4): 887-917.
Haack, Kirsten. 2014. “Gaining Access to the ‘World’s Largest Men’s Club’: Women Leading UN Agencies.” *Global Society* 28(2): 217–40. http://dx.doi.org/10.1080/13600826.2014.887558.
Haack, Kirsten, Margaret P. Karns, and Jean Pierre Murray. 2020. “The United Nations at Seventy-Five: Where Are the Women in the United Nations Now?” *Ethics and International Affairs* 34(3): 361–71.
Hafner-Burton, Emilie and Mark A. Pollack. 2002. "Gender Mainstreaming and Global Governance," *Feminist Legal Studies* 10: 285-298.
Hannan, Carolyn, 2012. "Feminist Strategies in International Organization: the United Nations Context," in Feminist Strategies in International Governance. London: Routledge.
Hawkins, Darren G., David A. Lake, Daniel L. Nielson, and Michael J. Tierney. 2006. *Delegation and Agency in International Organizations*. Cambridge: Cambridge University Press.
Heinzel, Mirko, and Andrea Liese. 2021. “Managing Performance and Winning Trust: How World Bank Staff Shape Recipient Performance.” *Review of International Organizations* First view.
Holman, Mirya R. 2017. “Women in Local Government.” *State and Local Government Review* 49(4): 285–96.
Honig, Dan. 2019. “When Reporting Undermines Performance : The Costs of Politically Constrained Organizational Autonomy in Foreign Aid Implementation.” *International Organization* 73: 171–201.
———. 2020. “Information, Power, & Location: World Bank Staff Decentralization and Aid Project Success.” *Governance* 33(4): 749–69.
Hooghe, Liesbet. 2005. “Several Roads Lead to International Norms, but Few via International

Socialization: A Case Study of the European Commission.” *International Organization* 59(4): 861–98.
Jayachandran, Seema, and Jamie Daubenspeck. 2018. “Gender Differences in What Development Economists Study.” *World Bank Blogs*. https://blogs.worldbank.org/impactevaluations/gender-differences-what-development-economists-study.
Karim, Sabrina, and Kyle Beardsley. 2013. “Female Peacekeepers and Gender Balancing: Token Gestures or Informed Policymaking?” *International Interactions* 39(4): 461–88.
———. 2016. “Explaining Sexual Exploitation and Abuse in Peacekeeping Missions: The Role of Female Peacekeepers and Gender Equality in Contributing Countries.” *Journal of Peace Research* 53(1): 100–115.
Keiser, Lael R., Vicky M. Wilkins, Kenneth J. Meier, and Catherine A. Holland. 2002. “Lipstick and Logarithms: Gender, Institutional Context, and Representative Bureaucracy.” *American Political Science Review* 96(3): 553–64.
Kenny, Charles, and Megan O’Donnell. 2016. “Do the Results Match the Rhetoric? An Examination of World Bank Gender Projects.” (March). www.cgdev.orgwww.cgdev.org.
Kentikelenis, Alexander E., Thomas H. Stubbs, and Lawrence P. King. 2016. “IMF Conditionality and Development Policy Space, 1985–2014.” *Review of International Political Economy* 23(4): 543–82. http://dx.doi.org/10.1080/09692290.2016.1174953.
Kingsley, J. Donald. 1944. *Representative Bureaucracy: An Interpretation of the British Civil Service*. Ohio: The Antioch Press.
Limodio, Nicola. 2021. “Bureaucrats Allocation in the Public Sector: Evidence from the World Bank.” *The Economic Journal* First view.
Lipson, Michael. 2007. “Peacekeeping : Organized Hypocrisy ?” *European Journal of International Relations* 13(1): 5–34.
Lyne, Mona M., Daniel L. Nielson, and Michael J. Tierney. 2009. “Controlling Coalitions: Social Lending at the Multilateral Development Banks.” *Review of International Organizations* 4(4): 407–33.
MacRae, Heather. 2012. “Double-Speak: The European Union and Gender Parity.” *West European Politics* 35(2): 301–18.
Mansbridge, Jane. 2005. “Quota Problems: Combating the Dangers of Essentialism.” *Politics and Gender* 1(4): 622–38.
Meier, Kenneth J. 1992. “Representative Bureaucracy: A Theoretical and Empirical Exposition.” In *Research in Public Administration*, ed. James Perry. Greenwich: JAI Press, 1–35.
Mendelberg, Tali, Christopher F. Karpowitz, and Nicholas Goedert. 2014. “Does Descriptive Representation Facilitate Women’s Distinctive Voice? How Gender Composition and Decision Rules Affect Deliberation.” *American Journal of Political Science* 58(2): 291–306.
Meyer, M.K. and E. Prügl. 1999. *Gender Poltics in Global Governance*. Rowman & Littlefield.
Moser, Caroline and Annalise Moser. 2005. "Gender Mainstreaming since Beijing: A Review of Success and Limitations in International Institutions," *Gender & Development*, 13(2): 11-22.
Mosher, Frederick. 1982. *Democracy and the Public Service*. Oxford: Oxford University Press.
Murdoch, Zuzana, Hussein Kassim, Sara Connolly, and Benny Geys. 2019. “Do International Institutions Matter? Socialization and International Bureaucrats.” *European Journal of International Relations* 25(3): 852–77.
Neumayer, Eric, and Thomas Plümper. 2007. “The Gendered Nature of Natural Disasters: The Impact of

Catastrophic Events on the Gender Gap in Life Expectancy, 1981-2002." *Annals of the Association of American Geographers* 97(3): 551–66.
Novosad, Paul, and Eric Werker. 2019. "Who Runs the International System? Nationality and Leadership in the United Nations Secretariat." *Review of International Organizations* 14(1): 1–33.
Nyrup, Jacob, and Stuart Bramwell. 2020. "Who Governs? A New Global Dataset on Members of Cabinets." *American Political Science Review* 114(4): 1366–74.
OECD. 2020. "Creditor Reporting System." https://stats.oecd.org/index.aspx?DataSetCode=CRS1# (January 5, 2020).
Parizek, Michal. 2017. "Control, Soft Information, and the Politics of International Organizations Staffing." *Review of International Organizations* 12: 559–83.
Park, Sanghee. 2013. "Does Gender Matter? The Effect of Gender Representation of Public Bureaucracy on Governmental Performance." *American Review of Public Administration* 43(2): 221–42.
Park, Sanghee, and Jiaqi Liang. 2021. "A Comparative Study of Gender Representation and Social Outcomes: The Effect of Political and Bureaucratic Representation." *Public Administration Review* 81(2): 321–32.
Park, Susan, and Antje Vetterlein. 2010. *Owning Development. Creating Policy Norms in the IMF and the World Bank.* Cambridge: Cambridge University Press.
Riegle-Crumb, Catherine, and Melissa Humphries. 2012. Exploring Bias in Math Teachers' Perceptions of Students' Ability by Gender and Race/Ethnicity. *Gender and Society* (26)
Ruggeri, Andrea, Chiara Ruffa, and Vincenzo Bove. 2020. *Composing Peace: Mission Composition in UN Peacekeeping*. Oxford: Oxford University Press.
Sapiro, Virginia. 1981. "When Are Interests Interesting? The Problem of Political Representation of Women." *American Political Science Review* 75(3): 701–16. https://doi.org/10.2307/1960962.
Tallberg, Jonas, Magnus Lundgren, Thomas Sommerer, and Theresa Squatrito. 2020. "Why International Organizations Commit to Liberal Norms." *International Studies Quarterly* 64(3): 626–40.
Vestergaard, Jakob, and Robert H Wade. 2013. "Protecting Power : How Western States Retain The Dominant Voice in The World Bank ' s Governance." *World Development* 46: 153–64. http://dx.doi.org/10.1016/j.worlddev.2013.01.031.
Voeten, Erik. 2020. "Gender and Judging: Evidence from the European Court of Human Right." *Journal of European Public Policy* 0(0): 1–21.
WDI. 2020. "World Development Indicators (Version April)."
Weaver, Catherine. 2008. *Hypocrisy Trap. The World Bank and the Poverty of Reform*. Princeton and Oxford: Princeton University Press.
———. 2010. "The Strategic Social Construction of the World Bank's Gender and Development Policy Norm." In *Owning Development. Creating Policy Norms in the IMF and the World Bank.*, eds. Susan Park and Antje Vetterlein. Cambridge: Cambridge University Press, 70–90.
Weaver, Catherine, and Stephen C. Nelson. 2016. "Organizational Culture." In *The Oxford Handbook of International Organization*, 920–39.
Weller, Patrick, and Xu Yi-chong. 2010. "Agents of Influence. Country Directors at the World Bank." *Public Administration* 88(1): 211–31.
Wilkins, Vicky M. 2007. "Exploring the Causal Story: Gender, Active Representation, and Bureaucratic Priorities." *Journal of Public Administration Research and Theory* 17(1): 77–94.
Wilkins, Vicky M., and Lael R. Keiser. 2006. "Linking Passive and Active Representation by Gender: The Case of Child Support Agencies." *Journal of Public Administration Research and Theory* 16(1):

87–102.
Winters, Janelle, Genevie Fernandes, Lauren McGivern, and Devi Sridhar. 2018. "Mainstreaming as Rhetoric or Reality? Gender and Global Health at the World Bank." *Wellcome Open Research* 3(0): 18.
World Bank. 2005. *Mauritania - Mining Sector Capacity Building Project. Implementation Completion and Results Report.* Washington, D.C: World Bank.
———. 2012. *World Development Report 2012: Gender Equality and Development*. Washington, D.C: World Bank.
———. 2017. *China - CN-Zhejiang Qiantang River Basin Small Town Environment Project. Implementation Completion and Results Report.* Washington, D.C: World Bank.
———. 2018. *Monitoring Gender Mainstreaming in the World Bank Lending Operations Database*. Washington, D.C: World Bank. https://documents.worldbank.org/en/publication/documents-reports/documentdetail/207551599142919624/uzbekistan-health-system-improvement-project.
———. 2020. *Uzbekistan - Health System Improvement Project. Implementation Completion and Results Report.* Washington, D.C: World Bank.
———. 2021. "The World Bank Group Country Opinion Survey Program (COS)." *World Bank Microdata*. https://microdata.worldbank.org/index.php/catalog/COS/about.