Programmatic Parties and the Politics of Bureaucratic Reform

Cesi Cruz
University of California, San Diego

Philip Keefer
The World Bank

pkeefer@worldbank.org
May 14, 2012

Abstract: In many countries, politicians neglect the basic financial and personnel management systems that are essential to political oversight of bureaucratic performance. To explain this, we present a new perspective on the political economy of bureaucracy, arguing that politicians organized into programmatic political parties have stronger incentives to pursue public policies that require a well-functioning public administration. They are also better able to act collectively to demand that the executive provide such an administration. We find robust support for this argument with novel evidence: ratings of 511 World Bank public sector reform loans in 109 countries are systematically higher in countries with programmatic political parties.

Programmatic Parties and the Politics of Public Sector Reform

In many countries, politicians appear to be entirely indifferent to basic financial and personnel management systems that enable politicians to exercise oversight over the public administration. They neither insist that public sector financial management systems allow them to track agency spending, nor care that personnel management practices give them scant control over the number and quality of staff responsible for public administration. Instead, public administration is characterized by lax financial controls and significant leakage, on the one hand, and patronage-based personnel systems and weak capacity, on the other. This phenomenon remains an important puzzle that the large literature on the political economy of public sector organization has only partially addressed. The analysis here develops an original explanation of the political decision to neglect financial and personnel management in the public administration: where politicians are not collectively organized, they have neither the incentive to pursue broad public policies that are particularly sensitive to the quality of public administration, nor are they capable of obliging the executive to implement strong financial controls and meritocratic personnel systems. We test this argument using unique data, evaluations of the performance of World Bank loans to support public sector reform.

Much of the literature, particularly US-focused, takes for granted that politicians care intensely about shirking in the public administration and probes the instruments that they use to control shirking. Shefter (1994) concludes that the strength of bureaucracies shapes political incentives, implying that politicians can do little to control shirking by strong bureaucracies and implicitly raising the question of why they would ever create such bureaucracies in the first place. Tilly (1990) and others address precisely this question:

confronted with dramatic threats, particularly war, politicians support the emergence of a strong, autonomous public administration.

In contrast to these lines of research, the analysis here offers a new explanation, not rooted in war, of why politicians care about bureaucratic shirking more in some settings than in others: organized politicians, able to act jointly in pursuit of their collective political interests, are more likely to insist on transparent, rule-bound systems of public sector management. It provides the first systematic evidence that sophisticated systems of financial and personnel management are associated with political organization. Moreover, the evidence distinguishes two possible causal paths that could explain this association. One is the contention of Shefter (1994) and Chandra (2004), that strong bureaucracies make it more difficult for unorganized politicians to deliver on clientelist promises to their personal constituents. The other, advanced here and supported by the evidence presented below, is that organized politicians are more likely to insist on transparent, rule-bound systems of public sector management.

The new theoretical argument we introduce is straightforward. Financial and personnel management systems reduce the costs to politicians of monitoring the implementation of the public policies that they care about. However, some policies, particularly those that target more citizens and provide public goods, are more difficult for politicians to monitor than others. Politicians can easily verify whether narrowly targeted policies have been implemented, even without strong systems of personnel and financial management. They may even implement such policies themselves, as with constituency development funds (Keefer and Khemani 2009). In contrast, politicians with incentives to pursue public goods and more broadly targeted policies have stronger incentives to invest in a system of transparent and rule-bound public administration.

A large literature examines political preferences for public versus private goods, finding that politicians whose political careers depend only on their personal characteristics have weak incentives to provide public goods and have little capacity to push for administrative reform. In contrast, groups of politicians who are organized to overcome free rider problems can mobilize support on the basis of their group membership and not only their personal characteristics. Such politicians prefer policies that provide public goods and distribute benefits broadly, to build up the group's attractiveness. They are also better able to demand the financial and personnel management systems necessary to oversee them. Finally, programmatic political parties – those with the capacity to convey a credible policy stance to citizens – signal that groups of politicians have organized to overcome free rider problems. As a consequence, countries with programmatic political parties should be more likely to adopt and exhibit stronger systems of financial and personnel management.

Using new and unique data on public administration reform projects from the World Bank, and data on programmatic parties from the Database of Political Institutions, we find strong support for this argument. Our evidence demonstrates that public administration reforms are more likely to succeed in countries with programmatic political parties (or "party" in the case of ruling parties in non-democracies). These results are robust to numerous specifications and controls for endogeneity.

The next section details the argument more thoroughly. A significant contribution of the paper is then to introduce a new approach to testing the politics of public sector reform using World Bank lending data; these data are described in the following section. The remainder of the paper reports numerous tests using these data and outlines their implications for research on the politics of public sector reform.

Public administration and programmatic political parties

The administrative reforms that we examine empirically involve the installation of management systems that ensure that personnel and spending decisions in the public administration are transparent and rooted in the overall policy objectives that politicians set for agencies. The systems impede decisions based on *ad hoc* criteria that are at odds with statutory objectives (e.g., hiring based on cronyism, spending divorced from budget allocations). In the presence of these rules, departures from established policy are more transparent, facilitating political oversight of the public administration.

Consistent with most of the literature, this section focuses on the incentives of non-executive politicians to adopt public sector reforms; the next extends these arguments to executive politicians, whether elected or unelected. The essential argument is the following. Transparent, rule-based systems of financial and personnel management reduce the costs to politicians of identifying shirking by the public administration in the implementation of public policies. Second, it is more costly for politicians to identify shirking in the implementation of policies that affect more citizens and provide public goods. Third, relative to individual politicians, groups of politicians who are organized to overcome free rider problems both prefer such policies and are better able to demand the financial and personnel management systems necessary to oversee them. Fourth, programmatic political parties – those with the capacity to convey a credible policy stance to citizens – signal the presence of politician organized to overcome free rider problems. As a consequence, countries with programmatic political parties should be more likely to adopt and exhibit stronger systems of financial and personnel management.

It is straightforward that transparent and rule-bound public sector management systems reduce the costs to politicians of observing administrative shirking. One form that shirking takes is the diversion of financial resources away from politician priorities. Financial management systems make this easier to observe by tracking financial flows to projects and inputs.[1] A second form of shirking is simply not to exert effort to implement those priorities. This type of shirking is more difficult for politicians to identify when public employees are of lower capacity: they cannot distinguish whether performance failures are due to shirking or lack of competence. As the model of Huber and McCarty (2004) shows, civil service reforms that establish transparent hiring and promotion standards and require employees to have expertise and training in the pursuit of politicians' objectives reduce this potential ambiguity.

Such standards have two other effects that limit shirking. Employees with more expertise and training in the pursuit of politician objectives may be more intrinsically motivated to pursue them. In addition, standards limit the ability of individual politicians to influence employee selection to pursue politicians' private objectives. For example, standards that require a high level of academic achievement or experience restrict the types of clients that politicians can attempt to place in the public sector.

These financial and administrative systems are more useful to politicians for those public policies in which shirking is more difficult for politicians to detect. The costs of detecting shirking rise with the number of intended beneficiaries of a policy, and with the extent to which the policy is intended to provide public goods rather than private transfers. It is easy to see that broadly targeted policies present greater monitoring challenges. Policy success is judged by the number of individuals served; when the number of targeted

---

[1] On May 25, 2012, the director of the United States National Weather Service, an agency in the Department of Commerce, resigned after an audit found that the NWS had diverted approximately $30 million per year to 122 weather offices around the country without notifying Congress. Senator Olympia Snowe called the report "deeply troubling." (*The Washington Post*, "Weather Service director retires after critical investigation of the agency," p. A-2).

beneficiaries is large, politicians require correspondingly more information to judge whether the policy is successful. Individual politicians have easy access to information on the policy benefits received by a relatively narrow group of individuals with whom they are in regular contact. They cannot easily infer from the experience of this small group whether a program targeted at many citizens has achieved its goals.

The implementation of public goods is also more difficult for politicians to observe. First, by definition, the welfare effects of public good provision depend on the number of citizens who benefit from the public good. Once again, therefore, politicians cannot rely on the experiences of a narrow group to infer how well a public good has been implemented. Second, in practice, public goods tend to target outcomes that are more likely to be the product of multiple influences that are difficult to observe and unrelated to the effort of public employees. Student learning is a function not only of the effort of public employees, but also of unobservable family and student effort; pollution control depends on public employees' efforts to protect the environment, but also on unobservable factors that influence environmental outcomes; national defense depends not only on the effort of a nation's security forces, but also on the actions of a nation's enemies. In contrast, cash transfers are almost solely a function of employee effort. To assess public employees' efforts to provide public goods, politicians are therefore more likely than in the case of private goods to require a system that aggregates the outcomes experienced by citizens and disentangles the contribution that public sector effort plays in generating those outcomes.

The analysis in Johnson and Libecap (1994), to explain the professionalization of the federal bureaucracy of the United States, is consistent with this argument. Following the Civil War, the scope and complexity of government activity increased substantially and the federal labor force grew correspondingly. Legislators in the United States Congress found it difficult to oversee government employees, leading them to approve measures that increased the professionalization of the bureaucracy and restricted legislators' ability to use the bureaucracy as a vehicle for patronage.

The two aspects – number of beneficiaries and public goods – are related. It is well-known that public goods can be a more efficient mode than private goods of delivering benefits to larger groups (Persson and Tabellini 2000). Money spent on a public health program to control infectious disease, for example, can raise average citizen welfare by much more than if politicians use the same resources to deliver small cash transfers to each citizen. However, it is more difficult for politicians to observe the implementation of the public health program than of the small transfers.

Politicians with incentives to pursue policies that provide public goods to large numbers of citizens are therefore more likely to prefer transparent and rule-bound financial and personnel management systems. An extensive and diverse literature analyzes precisely the conditions under which legislators seek to appeal to broader constituencies and provide public goods. One central condition is the degree to which politicians' careers depend on their personal characteristics or those of the group of politicians with which they are affiliated (e.g., McCubbins and Rosenbluth 1995). Keefer and Vlaicu (2008) argue that political appeals to larger groups depend on the degree to which politicians can make broadly credible pre-electoral commitments. "Clientelist" politicians, those who can make credible pre-electoral commitments only to narrow constituencies, prefer narrowly targeted policies to public good provision, in contrast to politicians who can make credible pre-electoral commitments to broad groups of citizens.

The ability of a group of politicians to campaign on their collective achievements, including their ability to make credible commitments to large groups of citizens, depends on whether they are sufficiently well-organized to solve a dual collective action problem (Keefer 2011). On the one hand, citizens must be capable of acting collectively to punish politicians who renege on their commitment to provide public goods. However, individually they have incentives to free-ride since their access to the public good is not conditioned on whether they join in the collective action to discipline politicians. On the other hand, politicians must be capable of acting collectively to mobilize group constituencies. Individual politicians in the group have an incentive to invest effort on behalf of their personal constituencies, free riding on the efforts of other group politicians to mobilize a collective constituency for the group. Political organizations that curb free-riding by politicians on efforts to mobilize group constituencies are therefore an essential foundation of political incentives to provide public goods to large groups of citizens.

Such organizations also allow non-executive politicians to act collectively to impose administrative reforms on executive branch politicians who may be reluctant to accept them. Moe (1987) and Cox and McCubbins (1994) both emphasize that collective action by legislators is key if they are to resist executive offers of redistributive transfers to any subset of them in exchange for their support for looser executive oversight or the provision of fewer public goods. Absent the ability of non-executive politicians to act collectively, it is straightforward for the executive to use threats of selective punishment to persuade individual legislators to give up their proposals to reform the public administration. If politicians can act collectively, they can resist threats of selective punishment by withdrawing their collective support from the executive; they can respond to selective inducements by punishing the politician who accepts them.

This argument predicts that politicians who have succeeded in overcoming their collective action problems are more likely to commit to the provision of public goods to broad groups of citizens and, as a consequence, more likely to support transparent and rule-bound systems of public administration. The tests below take this prediction to the data using, as a proxy for politician organization, the presence of programmatic political parties in countries. This section concludes, therefore, by arguing that parties capable of making credible policy commitments to large groups of voters – programmatic parties – address both obstacles to collective action.

Parties capable of making credible programmatic policy promises to broad groups of voters recruit like-minded citizens as members and invest in mobilizing the political participation of these citizens, facilitating citizen coordination. They also solve politicians' collective action problems, applying litmus tests to potential candidates to ensure that they share common policy preferences, imposing sanctions on those who deviate from the party line, and facilitating collective action by party members to remove leaders who shirk in the pursuit of the party's programmatic objectives (Keefer 2011).

Such parties make it easier for citizens to hold politicians collectively accountable for their actions, encouraging politicians to care less about the particularistic benefits they deliver to voters and more about broadly distributed benefits (e.g., Aldrich 1995). The sanctions imposed by such parties are meaningful, because these parties offer voters credible reasons to take into account both party affiliation and candidate characteristics in making their electoral decisions. Expulsion, for example, costs such politicians the support of all those citizens who value the party's programmatic stance (see, e.g., Snyder and Ting 2002).

Machine parties also solve the dual collective action problem, but with a different effect. These parties recruit candidates on the basis of their demonstrated loyalty to the party itself rather than to a policy program supported by the party. They coordinate citizen support through the direct allocation, by the party, of private benefits that the party itself strategically targets to citizens. Citizens who fail to support the party lose access to these benefits; politicians who fail to support the party must mobilize citizen support without the benefit of the party machinery. However, machine parties, relative to programmatic parties, have less interest in improving the quality of the state public administration for the purpose of providing public goods; they are more concerned with the administrative apparatus of the party and with building up the party's own capacity to deliver benefits to voters.[2] Hence, the comparison in the analysis below is between countries with programmatic political parties and all others.

Programmatic parties and executive incentives to reform public administration in democracies and non-democracies

Most analyses of the political economy of public administration focus on politicians outside of the executive branch, typically the legislative branch of government in democracies. This work is primarily concerned with the challenge for non-executive politicians of monitoring the executive, of which the public administration is simply an extension. Executives themselves, though, whether elected or unelected, also confront challenges in overseeing the bureaucracy over which they preside. They also have a choice about whether to meet these challenges by adopting transparent, rule-bound financial and personnel management systems. The presence of programmatic parties influences these incentives. On the one hand, executives cannot implement such policies without the cooperation of the legislature; legislative interest in such policies is lower in the absence of

---

[2] Machine parties would, however, be interested in the efficiency of revenue collection, to finance the party apparatus. Eaton's (2003) discussion of tax administration in Argentina supports the general prediction that legislative oversight is more vigorous when legislators are better organized. He concludes that strong party discipline (of the Peronist party machine) in Argentina enabled the legislature to exercise significant collective oversight of the tax administration. Siavelis (2000) argues that congressional oversight of the public administration in Chile is weak. However, parties are programmatic in Chile and Chile has the most well-regarded public administration in Latin America. Siavelis' analysis focuses on constitutional changes that weakened Congressional influence over the budget, giving legislators, no matter how well-organized, more limited influence over the executive.

programmatic parties. On the other hand, if executives promise public goods and renege on them, the absence of a programmatic party makes it more difficult for citizens or other politicians to act collectively to remove them from office.

Like non-executive politicians, then, in the absence of programmatic parties, elected executives are more likely to rely on narrowly targeted policies to retain power and are less likely to demand sophisticated management systems to supervise the public sector. Executives' political success, in the absence of programmatic parties, depends on their personal characteristics and ability to mobilize electoral support through patron-client networks, much the same as for their counterparts among non-executive politicians. Transparent and rule-bound management systems restrict their ability to engage in these strategies. Joaquín Balaguer, for example, the president of the Dominican Republic for many years, did not represent a programmatic party and personally directed the spending of as much as 50 percent of the national budget (Keefer and Vlaicu 2008).

A similar logic extends to unelected executives who attempt to remain in power by committing to policies that benefit broad constituencies, raising the costs of revolt by challengers who cannot credibly match this commitment. The ruler commitment to broad policy benefits, though, is not credible unless rulers can be punished for reneging on the commitments. Rulers can solve their credibility problem by linking the compensation of members of the ruling party to their success in implementing the strategy, and allowing ruling party members to organize collectively to impose sanctions on them if they renege on these commitments.

Gehlbach and Keefer (2010) model this process in the case of rulers who choose to promote private investment and job creation through the ruling party. Rulers facilitate collective action by ruling party members, allowing them to punish rulers who expropriate members' investments, yielding higher private investment. They illustrate their argument with evidence from China. Following the death of Mao, who took measures that undermined the ability of Communist Party members to organize collectively, Deng Xiaoping undertook numerous reforms to increase that ability. These included personnel reforms that increased the transparency of leader decisions towards members: systematic standards of promotion for party cadres that both explicitly linked their performance to broader party objectives and made arbitrary treatment of cadres by leaders more transparent.

Testable Hypotheses

The theory predicts, first, that in countries with programmatic political parties, we should observe greater emphasis on public policies that benefit citizens broadly. Keefer (2010) finds evidence for this conclusion: programmatic political parties give stronger incentives to both legislators and executives (elected or unelected) to pursue broadly targeted policies. Second, politicians in countries with programmatic parties should prefer financial and personnel management systems that facilitate politician oversight of the public administration. A corollary of this second hypothesis, the focus of the empirical tests below, is that efforts to introduce these systems into a country are more likely to succeed in the presence of programmatic parties.

The theory points to a causal effect of programmatic parties on the presence of transparent, rule-bound administrative practices. Shefter (1994) and Chandra (2004) make the reverse argument, that parties turn to programmatic appeals in countries where bureaucracies are independent, well-functioning and resist the implementation of arbitrary, clientelist policies. Chandra (2004), for example, observes that in India patronage is easier for politicians to dispense when the civil service is less professional. These arguments, consistent with Tilly (1990), have as their underlying assumption that well-functioning bureaucracies are able to block politicians from implementing clientelist policies, encouraging them to form programmatic parties.

The empirical tests below provide the first systematic evidence of an association between programmatic parties and transparent and rule-bound administrative systems. However, while these results, showing an association between programmatic parties and bureaucratic reform, fill an important gap in the literature, sorting out issues of causal inference remains critical. One reason for this is the policy imperative. If, for example, strong bureaucracies come first, then donor activities to reform bureaucracy are likely to have more far-reaching effects than if bureaucratic reform is a consequence of political change.

There is ample reason to believe that political incentives might dictate bureaucratic institutions. For example, despite the strong bureaucratic legacy of British colonial rule in India and Kenya, clientelist and sectarian or ethnic appeals (detached from credible policy commitments to these groups) are pervasive. Van de Walle (2003) argues that in the 1990s, parties in African democracies switched to programmatic appeals not because strong bureaucracies pushed them away from clientelist strategies, the explanation that flows from Shefter's arguments, but because politicians themselves found it difficult to manage clientelist handouts. Their efforts to move to programmatic policy promises were not particularly successful, though. The argument here attributes the lack of success to the weak incentives of party leaders to institutionalize the programmatic agenda inside their parties and to encourage the emergence of a strong bureaucracy that could implement their policy agenda.

The tests below have several attributes that provide empirical support for the causal path going from programmatic parties to bureaucratic reform. First, the presence of programmatic parties is coded at the beginning of the bureaucratic reform efforts analyzed here – they cannot be the product of these efforts. Second, the strong bureaucracies that Shefter or Tilly have in mind are largely absent in the sample. Indeed, this is one reason why the World Bank extends loans to them to encourage public sector reform. It is, therefore, implausible that strong bureaucracies prompted politicians to pursue public sector reform in these countries. Third, our results are robust to the use of instrumental variables that exclude the influence of bureaucracies on programmatic parties.

Measuring public sector reform: Evaluations of World Bank public sector loans

The large literature examining both the causes and effects of bureaucratic quality (see, e.g., Keefer and Knack 1995, Rauch and Evans 2000) has relied on subjective and broad evaluations of bureaucratic quality, particularly measures of corruption and bureaucratic quality in *International Country Risk Guide*. Keefer (2011) shows that these measures are, in fact, associated with programmatic parties. The inherent limitations of standard measures of bureaucratic quality mean, however, that this finding sheds little light on the origins of administrative reform and on the degree to which politicians support system improvements in order to boost administrative performance.

One limitation is that broad, subjective measures of bureaucratic quality change little over time, making causality difficult to assess. Another is that the measures look at performance (e.g., whether bureaucratic performance is predictable and efficient). This also has implications for causality: deep social and historical circumstances could explain both bureaucratic performance and the degree to which politicians are organized.

The focus on performance means that evaluators do not explicitly rate or even characterize the bureaucratic systems that underlie performance: whether public sector financial management systems are capable of tracking spending, or whether personnel management systems are designed to attract and retain the best and brightest candidates for public sector work. These, however, are the objects of interest in the political economy of public administration: under what conditions do politicians have incentives to introduce these arrangements into the public sector?

This analysis exploits a new source of data that give a precise characterization of the public sector reforms undertaken by countries. The data are evaluations of World Bank public sector reform loans, produced by the Independent Evaluation Group, an independent unit of the Bank that reports to the Bank Board of Executive Directors and not to Bank management. IEG evaluations are commissioned by the Board or decided upon unilaterally by the IEG. Bank management makes a formal response to IEG reports that is printed together with the report; it has no right, however, to require any changes in the wording or conclusion of the IEG document itself or its underlying ratings.

The IEG selected 467 loans that began in 1987 or later and ended in 2006 or earlier. World Bank loans often have multiple, cross-sectoral objectives. IEG therefore classified as "public sector" (as opposed to education, infrastructure, health, etc.) those projects that met two criteria. First, at least 25 percent of the loan amount had to be related to achieving PSR-related themes or sectors. Second, because loans are generally disbursed when countries demonstrate that they are meeting development goals established in the conditions of the loan, the IEG required, in addition, at least three of these conditions to be in PSR focus areas.

For projects active since December 2001, the Bank's management information system classifies projects according both to the sectors (public sector, economic policy, health, education, energy and mining, etc.) and up to five sub-sector themes. Prior to December 2001, only sectoral classifications are available; since December 2001, project managers sometimes assign only three or four rather than the possible five sub-sector themes. The most frequently assigned sub-sector theme in this group of loans is public expenditure and financial management, appearing in 307 of the 467 projects. The next most commonly-assigned are administrative and civil service reform (249); state enterprise reform (144); and tax policy and administration (120). A large number of other themes appear, but none more than 90 times. Among these, the most important are regulation and competition policy (90); debt management and fiscal sustainability (87) and macroeconomic management (83).

Public sector reform in the World Bank is broadly defined and includes decentralization, anti-corruption, state enterprise reform, and judicial reform. Many of these themes fit only loosely with the arguments here. In our base regressions we therefore restrict the analysis only to those loans for which at least one sub-sector theme was public expenditure and financial management; administrative and civil service reform; or tax policy and administration.

The public expenditure and financial management theme is defined as budget planning and execution (financial management information systems and medium-term expenditure frameworks), procurement, auditing, and monitoring and evaluation. Administrative and civil service reform involves all aspects of the management and organization of personnel. It includes programs to downsize the civil service and reforms to the personnel information system (including civil service censuses), establish career paths, decompress pay grades and improve other aspects of the incentive system, and the organization of ministries. Finally, tax administration is defined as including the key aspects of revenue administration, particularly the institutional setting and development of operational processes, including automation and interaction with taxpayers (World Bank

2008, 4-7).[3] These particular public sector reform areas closely match the foregoing theoretical arguments. However, results reported below are robust to alternative samples, including the whole sample of loans (based on the argument that the World Bank classifies all of these as public sector loans); or after excluding tax policy and administration (since, although tax administration itself is entirely within the scope of our analysis, loans to assist in the revision of the tax code itself are less so).

Having chosen the 467 (511 in the updated database used in this study) public sector reform loans, the IEG subjected them to several evaluations, as it does to all loans that it evaluates. Two are most appropriate for the analysis here. First, the IEG ratings evaluate institutional development impact. This is the extent to which a project improves the ability of a country or region to make more efficient, equitable and sustainable use of its human, financial, and natural resources. For loans in most sectors, the main objectives of the projects can be achieved with little attention to institutional development: dams can be built successfully even if the ministry in charge of them remains a shambles; schools can be built and teachers hired even if the administrative apparatus for managing education resources remains weak. For these loans, institutional development might be secondary.

In the case of public sector reforms, however, institutional development is precisely the point. Moreover, in loans with mixed objectives, some public sector and some not, the institutional development rating is likely to weight the public sector components most highly. Among those loans that received an institutional development rating, five received a high rating; 136 a substantial rating; 172 a modest rating; and 44 the lowest rating, negligible.

---

[3] Downsizing might seem to be unrelated to the earlier arguments. However, one consequence of hiring based on patronage is a large civil service (Keefer 2007 shows that younger democracies – which tend to be more clientelist – spend more on public employment as a fraction of GDP). Another consequence is a civil service that is staffed inappropriately for achieving the statutory objectives of the public sector. Downsizing is a potentially necessary policy response if there is political support for more programmatic goals.

The second IEG rating used here is that of a loan's overall outcomes, ranging from highly satisfactory (the program achieved at least acceptable progress toward all major relevant objectives and had best practice development impact on one or more of them); to satisfactory (progress on all relevant objectives, none best practice and no major shortcomings) to moderately satisfactory (major progress on most objectives and no major shortcomings) to moderately unsatisfactory (no progress on most major objectives or progress on all, but with a major shortcoming, such as a violation of a safeguard (e.g., such as the improper resettlement of communities in an area to be dammed); and unsatisfactory – did not make acceptable progress on most objectives and produced a major shortcoming. In the sample, 10 received the highest rating, 151 the second-highest; 95 the third highest; 26 the fourth; and 68 the fifth and lowest.

The second, comprehensive evaluation could be a less precise measure of public sector reform success, since its rating is presumably based on all sectors and themes of the loan, public sector or not. Nevertheless, these outcome ratings formed the basis of the IEG's own report on PSR loans and this attracted no dissent from Bank management and staff (e.g., in the Bank management response to the IEG Evaluation). Moreover, the wording of the scoring categories, combined with the criteria used to identify PSR loans, make it unlikely that progress on PSR objectives would not be considered among the most important major objectives of the loans.

A key issue in lending is the effort to ensure that projects and loan conditions are well-designed at the outset, reflecting best practice and special country conditions. The IEG rated quality of projects at entry. We use this variable to test for the robustness of the party effect, to control for the "quality" of loan preparation. This is a difficult and conservative test, since the quality of loan preparation is related not only to Bank inputs, but also to country engagement with the lending process. Engagement, though, should be a function of political incentives, so that quality at entry could also capture variation in loan success ratings that is properly attributed to the presence or absence of programmatic political parties.

Once a loan is signed, Bank staff supervise country progress towards meeting loan objectives. The IEG also rated the quality of Bank supervision. Of those that were rated, 23 received a highly satisfactory rating; 264 a satisfactory rating and 57 an unsatisfactory rating; a handful of other projects received other intermediate ratings (moderately satisfactory, etc.). The IEG also rated overall Bank performance, which takes into account both project design and supervision; the distribution of responses is nearly identical to that for supervision. We use these ratings to show that, although programmatic parties affect loan performance, a product of country effort, they do not affect the Bank's own supervisory effort. This is important, since if IEG evaluations were endogenous to the presence of programmatic political parties, we would expect all IEG ratings to reflect the endogeneity bias. In fact, they do not.

The IEG evaluations of public sector loans have two advantages over the standard, broad and subjective measures of bureaucratic quality. One is exactly that they are precisely-defined indicators of the rules and systems of public administration. They allow us to test the hypotheses advanced here, that specific financial and personnel management reforms depend on the incentives of political actors interested in oversight of executive implementation of their policy priorities.

Another advantage is that they are less subject to endogeneity concerns. Current broad and subjective measures of bureaucratic quality focus on bureaucratic performance. Performance is a function of many factors, including administrative structure, but also historical legacies and the socio-economic environment. These factors, many of which are unobserved, are likely to influence both performance and whether countries have programmatic political parties. In contrast, the decision of countries to accept a loan from the World Bank to pursue specific administrative reforms is less likely to be influenced by these deeper underlying characteristics of countries. It is possible that conjunctural country circumstances prompt both the emergence of programmatic political parties and increased political incentives to improve the bureaucracy. The identification section below addresses this issue. However, it is important to underline that the influence of these circumstances is likely to be consistent with the theory advanced here: politicians confront an unobserved political shock that makes it important for them to provide public goods; to do this, they form programmatic parties, which allows them, in turn, to act collectively to demand improvements in the quality of the bureaucracy.

Other Data

The main correlate of interest, the fraction of parties in a country that are programmatic, is constructed from variables in the *Database of Political Institutions* (DPI, Beck et al. 2001). They identify whether each of the largest three government parties and the largest opposition party are right, left or center in their orientation or whether, on the contrary, their orientation is either not discernible in the sources employed or unrelated to economic policy. Our key variable is the share of these parties that is programmatic (can be coded as right, left or center, or *rlc_avg*). Across all the loans in our sample, only half of parties were programmatic, on average. We use the value for *rlc_avg* that corresponds to the first year of the loan, which is unlikely to be influenced by the successful or failed implementation of the loan and makes it more difficult to attribute our findings to a pattern of causality running from bureaucratic quality to programmatic parties. Results are robust to using other timing assumptions. [4]

Information from other sources supports the interpretation of the DPI indicator as a measure of a party's ability to make credible programmatic appeals to citizens. The Duke political parties database has extensive information on parties in 80 democracies in 2009. Assuming that these party characteristics were constant from 2000, one can compare these variables to the average DPI programmatic variable of sample countries with projects during period since 2000. The comparison points to a high correspondence between *rlc_avg* and the more detailed party characteristics identified by expert respondents in the Duke database.[5] For example, if parties are programmatic, one would expect them to have more stable factional groups in the legislature – party-switching is more difficult for politicians when ideological labels are attached to party affiliation. In the Duke dataset, parties that have factional stability (question A.7) are coded one and unstable parties are coded as two; the correlation with *rlc_avg* is -.42 for the 48 countries that are common to our database and the Duke database over this period. Similarly, programmatic parties should have less need to rely on charismatic leaders to mobilize voters. Consistent with this, the Duke variable indicating whether parties are led by charismatic leaders (E1) is correlated at -0.28 with *rlc_avg*.

IEG ratings could be based on assessments of bureaucratic capacity, which in turn is often linked to income and education: poorer countries with badly educated citizens are less able to accumulate the human capital needed to run an efficient public sector and to

---

[4] Such as averaging *rlc_avg* over the project period or using 3 and 5 year backward averages.

[5] This is a crude comparison, since the Duke database has information on more parties than DPI. Ideally, we would compare individual parties across the two databases, but this effort was stymied by significant differences in party labels across the two databases.

implement public sector reforms successfully. All regressions therefore control for the log of per capita income and the rate of progression to secondary school, both from the *World Development Indicators*. Log of per capita income is averaged over the duration of the loan, and rate of progression to secondary school is a country average.[6]

The size of the loan potentially relates to the resources that are applied to public sector reform, the sectoral diversity of the loan (larger loans are more likely to include non-public sector elements), and, potentially, the strictness and accuracy of IEG ratings (larger loans attract greater attention). All regressions therefore control for the loan amount and the percentage of the loan amount that is allocated for public sector components. They also control for the duration of the loan. Longer duration loans could imply a more conscientious effort by countries to pursue loan objectives (consistent with the frequently-voiced argument that institutional reform takes longer). However, within the World Bank, slow loan disbursements are also a red flag, indicating difficulties with implementation. In either case, duration is a rigorous control variable: we would expect the absence of programmatic political parties to make implementation more difficult and disbursement slower; and the presence of programmatic parties to lead to more conscientious implementation (but then, longer loan duration).

All regressions control for population and land area, from *World Development Indicators*. Larger countries may be more difficult to administer; on the other hand, there may be economies of scale in administering more populous countries (holding area constant). There is, in fact, some evidence that public sector reform is more likely to succeed in countries with

[6] The main results hold even when all variables that are averaged over the project period are measured using project start year values instead. Country averages for progression to secondary school are used because of limited data availability. Results are robust to using similar education variables, such as primary completion rate.

more people. All regressions also control for region dummies, using the UN standard regions.

We test the robustness of our results to many alternative specifications. Countries' initial level of bureaucratic quality may influence their ability to earn a successful rating from the IEG. Moreover, the specific characteristics of public sector reforms may depend on initial bureaucratic quality. We show that the results are robust to controlling for the beginning-of-period evaluation of bureaucratic quality found in Political Risk Services' *International Country Risk Guide*. Results are robust, as well, to controls for IEG's rating of quality at entry, described previously.

A large literature argues that formal political and electoral institutions influence political incentives to pursue particularistic versus public goods. Nearly all of that literature assumes that political parties are able to credibly promise public goods, and are therefore programmatic. The effects of these institutions in settings where political parties are weak and politicians are not collectively organized are less clear. Results are robust to controlling for whether a system is presidential or parliamentary, whether the electoral system is proportional or plurality, and for average district magnitude. These variables are taken from the *Database of Political Institutions*.

Most discussions of public sector organization and conflicts of interest between executive and legislature emphasize information asymmetries.   We use newspaper circulation in countries, taken from *World Development Indicators*, to account for this effect. Gehlbach and Keefer (2010) argue that leader willingness to allow for collective action (by a ruling party or by a bureaucracy) is attenuated when natural resource rents are high. Knack (2008) shows that natural resource exports (and, ironically, foreign aid) weaken government incentives to design efficient tax systems. We therefore control for natural resources as a percent of exports, from WDI.

Our base regressions pool all countries, more and less democratic, since our hypotheses apply equally to both. Results are robust, however, to restricting the sample to countries that have exhibited competitive elections and have had at least one democratic change of government; the definition follows, and the data are taken from, ACLP (Alvarez et. al. 1996, Cheibub et. al. 2009).

Identification

One clear prediction of the literature is that there should be an association between programmatic political parties and the success of personnel and financial management reforms of a country's public administration. The association might emerge because bureaucratic strength conditions political strategies, as in the arguments of Shefter, or because organized politicians are more likely to insist on a transparent and rule-bound public administration, the argument here. The tests below offer the first systematic evidence of this prediction, strongly rejecting the null hypothesis that such an association does not exist.

However, several characteristics of our tests allow us to go further and identify a causal effect of programmatic political parties on the incentives of politicians to adopt these reforms. One is that the reforms that we consider are technical – the introduction of management systems – that are unlikely to precipitate political change. A second is that the beginning-of-period measure of programmatic parties that we use is more likely to *contribute to* than to be a *function of* reform success. We measure programmatic parties the year that the public sector reform project was approved, well in advance of the start of the project and disbursement of funds. It is therefore difficult to argue that the reforms are a contributing factor to whether parties are programmatic.

Third, adopting a common strategy in the literature on political institutions, we instrument the programmatic party variable with country latitude (distance from the equator). Latitude is one of a number of possible instruments used by researchers to control for the possible endogeneity of institutions. It meets the three criteria for a reasonable instrument for the presence of programmatic political parties. One is that latitude is a strongly significant predictor of the presence of these parties (unlike all other instruments commonly used in the institutional literature, such as settler mortality and colonial history). In addition, it is theoretically reasonable that latitude should have a causal effect on the formation of programmatic parties. For example, lower latitude (proximity to the equator) is associated with more infectious disease, which should influence inhabitants' incentives to solve collective action problems. In areas with endemic disease and high mortality risk, for example, it might be more difficult for members of a political party to delegate to leaders the responsibility to prevent free-riding. Faced with higher mortality risks, leaders prefer to renege on their commitments to the party in order to accumulate rents, even if this means that they are expelled from the party.

Finally, it is plausible that latitude only influences bureaucratic reform through its influence on the emergence of programmatic political parties and, therefore, is reasonably excluded from the second stage regression. Moreover, to the extent that the exclusion assumption is not fully justified, it is likely that exclusion generates a bias against our hypothesis that programmatic parties promote the success of public sector reform. Continuing with the infectious disease example, to the extent that it is a greater hazard closer to the equator, the demand for public health programs should also be greater closer to the equator. This should create voter pressure for more effective government administration of these programs closer to the equator. However, countries closer to the equator are actually *less* likely to exhibit programmatic parties. To the extent that latitude influences public sector reform directly, this direct effect will therefore be absorbed by the programmatic party variable when latitude is excluded from the main regression. This leads to a downward bias in the programmatic party coefficient.[7]

Our timing assumptions and use of instrumental variables techniques suggest that omitted variables are not driving a spurious association between programmatic parties and bureaucratic reform, and that our results are not driven by the effect of bureaucratic reform on the emergence of programmatic parties. Measurement error is another possible source of endogeneity bias. World Bank evaluations may themselves be endogenous to the presence of programmatic political parties. [8] If evaluators increased success ratings when they observed programmatic parties, we should observe such a bias across all loan ratings. We therefore re-ran our models on ratings of Bank supervision and overall performance. Unlike the ratings of project institutional development impact and overall outcome, these ratings are given based on the quality of the Bank's project design and supervisory effort. According to our theory, programmatic parties should have no effect on these ratings. However, if Bank evaluations are indeed endogenous to programmatic parties, then we would expect that all of the ratings are subject to this bias. In fact, programmatic parties are insignificant predictors of these ratings.

[7] Lower latitude leads to more demand for effective administration of the public sector to curb infectious disease. Since it is also associated with fewer programmatic parties, the observed relationship between programmatic parties, instrumented by latitude, and public sector reform success, will be biased downwards. Fewer programmatic parties, instrumented by latitude, will be associated, through the possible direct effects of latitude, with more, rather than less, public sector reform.

[8] This is unlikely. No project or other Bank document that we reviewed, including the IEG Report (2008), even mentioned political parties.

Table 1: Summary Statistics

Variable	Obs.	Mean	Std. Dev.	Min	Max
Programmatic Parties (Full Sample)	415	0.5171888	0.3828382	0	1
Programmatic Parties (Democracies)	214	0.651947	0.3531845	0	1
Programmatic Parties (Non-democracies)	164	0.3974594	0.3534068	0	1
Log GDP per capita	418	1.271757	1.659389	0.0825486	9.024541
Log Population	418	16.59085	1.469631	12.9496	20.94271
Log Land Area	426	12.63474	1.644888	5.828946	16.61196
Progression to Secondary School (%)	380	75.74639	21.62246	25.98069	99.80688
Funding Amount	439	99.25689	198.3292	0.4933589	2525.25
Project Duration	439	3.776765	2.533845	1	12
Loan % Allocated to PSR	439	44.64237	21.00635	11	100
IEG Quality at Entry	305	4.213115	1.420188	1	6
Initial ICRG Bureaucratic Quality	354	2.512307	1.134971	0	6
PR	401	0.535469	0.4824109	0	1
Political System	423	0.3904903	0.742379	0	2
Mean District Magnitude	397	31.6449	107.382	0.7	888
Democracy (ACLP)	423	0.5462203	0.4815177	0	1
Newspaper Circulation	321	34.54603	42.17872	0.09	200.1508
Natural Resource Exports (% GDP)	365	2.691767	5.38174	0	49.19149

One final endogeneity concern relates to the fact that we are investigating reform episodes. This raises several issues. First, it could be the case that legislative processes for adopting World Bank loans differ systematically between countries with programmatic parties and those without. For example, countries that have programmatic parties may require legislative approval for Bank loans, while those without may not require such approval. This would imply that only loans that enjoy wide support receive approval in countries with programmatic parties and enter our database. This would create a potentially spurious bias in favor of finding that programmatic parties lead to public sector reform success.

In fact, though, all the countries in our sample share a similar legislative starting point. Legal authority to sign loan agreements with the World Bank rests with the executive branch and, within the executive, with the responsible sectoral minister and the finance minister. In the case of a health loan, for example, the minister of health and the minister of finance are the usual Bank counter-parties. Very rarely do national legislatures, or even cabinets, have to approve World Bank loans. For essentially all of the sample countries and loans, therefore, loan agreements have proceeded after similar levels of support have been exhibited by officials occupying similar positions.

On the other hand, our fundamental argument is that politicians in countries with programmatic parties should be more enthusiastic about adopting public sector reform loans than politicians that lack such parties. This again suggests that any bias due to our focus on reform countries is likely to work against our hypothesis. We are only comparing countries that agree to pursue public sector reform. Our theory predicts that countries without programmatic parties should be reluctant to make such an agreement. This means, however, that our sample of countries with non-programmatic parties has unobserved characteristics that make them more likely to apply for a public sector reform loan than the average country without programmatic parties. Those same unobserved characteristics might make them more likely to succeed in implementing the loan than other countries without programmatic parties. The fact that we are focusing only on countries that have applied for a public sector reform loan therefore biases our results towards a finding of no difference in their success with respect to public sector reform.

Specification

We use a logistic specification to estimate the following,

$IEG_i = \beta_0 + \beta_1 * (programmatic\ parties) + X_i \beta_2 + \varepsilon_i$. Each observation *i* is a loan. The variable IEG_i is the IEG rating of either institutional development or overall loan success. The original scales were used to create indicators of project success. The institutional development indicator variable takes the value of 1 when the institutional development impact of the project is "substantial" or higher; 0 otherwise. The overall outcome indicator variable takes the value of 1 when the outcome of the project is "satisfactory" or higher; 0 otherwise. All of the ratings indicator variables follow the standards used by the IEG and World Bank for determining project success. The main results are robust to alternative thresholds for project success. Furthermore, the main results are robust to ordered logit regressions using the full ratings scales, for all specifications that meet the proportional odds assumption.

Since the evaluation is categorical and we cannot be sure that differences between categories are constant across categories, the logistic specification is appropriate. We estimate only clustered standard errors, assuming that errors are distributed independently across countries, not across loans within countries. The programmatic party variable was described earlier. $\boldsymbol{X}_i$ is a vector of covariates; ε_i is an error term; and β_0, β_1, and β_2, are (vectors of) parameters to be estimated.

Results

The results indicate that programmatic political parties are significant determinants of the success of World Bank public sector reform projects, both in terms of the project's long term institutional development impact and overall outcome. Table 2 below reports the results of logistic regression for the baseline and full models. The baseline control variables include country characteristics (GDP per capita, population, land area, progression to secondary school) and project characteristics (project duration, funding amount, loan percentage allocated to public sector reform), and the full model adds IEG quality at entry and initial bureaucratic quality to the base model. All tables report odds ratios, the ratio of the odds that a project earns a successful rating to the odds that it earns an unsuccessful rating, adjusted for the simultaneous effect of the other explanatory variables.[9] Odds ratios greater than one imply that increases in the independent variable increase the probability of success; less than one, reduce it.

The odds ratios for programmatic politics in the base model are 4.02 for institutional development impact and 2.92 for overall project outcome. This implies that a one unit change in the proportion of programmatic parties increases the odds of a project earning a successful institutional development impact rating by a factor of 4.02, and a successful overall outcome rating by a factor of 2.92. When controlling for initial bureaucratic quality and the quality of the project design in the full model, the odds ratios are slightly larger: 5.89 for institutional development impact and 3.5 for overall project outcome. In terms of probabilities, holding all other variables at the mean, increasing the proportion of programmatic political parties from 0 to 1 increases the probability of institutional development impact success from .240 to .560 in the base model and from .197 to .591 in the full model. For overall project outcome ratings, the probability of earning a successful rating increases from .275 to .525 in the base model and from .188 to .449 in the full model.

The only baseline control variable significant across both models is funding amount: larger loans are associated with a lower probability of loan success. This is not due to a dilution effect – that public sector reforms play a smaller role in larger loans – because we control for the percent of the loan that the Bank attributes to the PSR component. Of the two additional variables in the full model, the IEG rating for quality at entry is significant and positive for both the project outcome and institutional development impact ratings. On the other hand, initial bureaucratic quality is associated with a *decrease* in the probability of a

---

[9] For all countries in the sample, the odds of success relative to the odds of failure is .653 for institutional development impact and .797 for overall outcome. The corresponding probability that a project will earn a successful rating is .395 for institutional development impact and .445 for overall outcome.

project earning a successful outcome rating. This estimated effect is consistent with at least three possibilities: that reform is more difficult in more advanced bureaucracies; that the Bank may hold countries with lower quality bureaucracies to more lax standards; and that public officials in more institutionalized bureaucratic settings may have greater ability to resist reform.

The results in Table 2 are robust to a wide range of robustness tests. One question, for example, is whether they are robust to outliers. In fact, if anything, the exclusion of outliers strengthens the results. This was true excluding those observations that had the largest residuals after plotting actual versus predicted values; those that, according to Cook's D, had the greatest overall influence on the regression results; and those with large dfBetas, which had the largest influence specifically on the programmatic party coefficient (results not reported).

Table 3 demonstrates robustness to the inclusion of a broad range of additional control variables. The main results hold even when controlling for newspaper circulation and natural resource exports, and for the inclusion of political controls (whether the electoral system is proportional representation; whether the political system is presidential, semi-presidential or parliamentary; mean district magnitude; and democracy, as determined by ACLP).

The main results are also robust to restricting the sample to countries rated as democratic by ACLP (Table 4). Within democracies and controlling for IEG quality at entry rating and initial bureaucratic quality, a unit increase in the proportion of programmatic parties is associated with an increase in the odds of successful institutional development impact rating by a factor of 10.32 and an increase in the odds of successful overall outcome rating by a factor of 11.69.

The estimates in Table 5 provide further support for our claim that the earlier results are not the spurious product of endogeneity. If IEG based its evaluations on the presence of programmatic parties, or other country characteristics besides the actual performance of the public sector loan, those omitted variables would likely affect all of the loan ratings given by IEG. However, the Table 5 results show that programmatic parties are *not* significant determinants of IEG ratings of World Bank supervision and overall performance. That is, the programmatic party variable is not related to ratings with which it should, in fact, have no association. This reduces the concern that the relationship between programmatic parties and project success is driven by some unobserved factor or IEG bias that should influence all ratings.

Table 6 also addresses endogeneity issues. First, we find evidence that the initial logistic regressions are already appropriate. The Wald statistic for the test of exogeneity is insignificant: there is insufficient information in the sample to reject the null hypothesis that programmatic parties are exogenous. This test implies that instrumenting for programmatic parties to control for endogeneity might lead to a loss of efficiency with no corresponding increase in consistency. Nevertheless, because of the limitations of the Wald test for exogeneity, we estimate an instrumental variables probit specification, using latitude as our instrument for programmatic parties, to address the potential concern that programmatic parties are endogenous to unobserved conjunctural political circumstances that might also influence support for bureaucratic reform.

The programmatic parties result holds in both models, with odds ratios of 8.725 and 4.322 for institutional development impact and overall outcome, respectively. In addition, the first stage Wald test is statistically significant for both models, with a value of 78.48 for institutional development impact and 26.97 for overall outcome, indicating that explanatory variables are jointly significant. Other first stage tests, including over- and under-identification and the Stock-Yogo test of weak instruments, all indicate that the instrument is valid.

Discussion and implications for the literature on the political economy of public sector reform

The analysis offers a new perspective on the political economy of politicians' oversight of the bureaucracy. Prior research (e.g., McCubbins and Schwartz, 1984 or Calvert, McCubbins and Weingast, 1989) focuses on *how* legislators control administrative decisions of the executive branch, asking whether politicians rely on *ex ante* prescriptions (e.g., precise directions to the bureaucracy on what decisions to make under a range of specific circumstances) or *ex post* supervision (e.g., oversight committees and judicial review of administrative decisions). All of these control mechanisms share the common characteristic that they are systemic and assume that politicians can act collectively to enforce them. In the absence of programmatic political parties, however, none of these methods are credible; legislators have limited ability to act collectively to punish executives who ignore them.

For example, McCubbins and Schwartz (1984) focus on two types of *ex post* controls used by the US Congress. "Fire alarms" allow supporters to alert politicians to adverse administrative decisions, giving politicians an opportunity to take credit for redressing grievances. They are facilitated by notice and comment procedures that regulators are required to use when they develop regulations to implement legislation. Legislators can also use "police patrols", in which Congressmen oversee agency performance themselves (e.g., by convening regular hearings at which agency officials appear).

Calvert, McCubbins and Weingast (1989) discuss a variety of *ex ante* controls on agency discretion: agreements between the legislature and executive on "the denomination of [the agency's] powers and jurisdiction, the specification of administrative procedures to be followed, and the type of personnel with which the agency is to be staffed (lawyers, economists, engineers, generic civil servants, etc.)" (p. 604). They conclude that politicians should almost always prefer to rely on *ex ante* institutional rules to guide agency actions, buttressed by *ex post* fire alarms to take into account unforeseen contingencies. Epstein and O'Halloryn (1994) point out that restricting discretion reduces the ability of agencies to use their expertise to respond to new information. Their evidence shows that the US Congress is most willing to limit discretion when preference differences with the public administration are greatest.

Thies (2001) also finds significant evidence of *ex ante* control devices (e.g., junior ministers from one party placed with senior ministers of another party) in parliamentary settings, where intra-coalition monitoring is the central challenge. In these cases, politicians from one party can trust that their representative (e.g., the junior minister) will protect their collective interests when monitoring the other party's actions.

The arguments here indicate considerably less interest in *ex ante* institutional rules among politicians who are not organized into programmatic parties. Legislators are more likely to rely on personal contacts with affected constituents to find out whether adverse regulatory decisions have been made and to use their personal contacts with the bureaucracy, rather than formal legislative action, to monitor implementation and seek redress. In coalition settings, politicians will trust oversight of the bureaucracy by their personal representatives, but not representatives of their party. In these environments, politicians are likely to expend little effort building up a meritocratic recruitment system in the public administration or elaborate financial management systems to monitor broad expenditures.

Moe (1987) points to another trade-off in countries with organized politicians, but which is less salient when politicians are not organized. Legislatures delegate the supervision of bureaucracies to legislative committees. However, when median legislators do not trust that these committees will represent their interests, they are more reluctant to delegate authority to the bureaucracy. Bawn (1997) finds support in the US Congress for the argument that legislators prefer more statutory (*ex ante*) controls on agency actions if they are not on standing oversight committees; committee members prefer *ex post* supervision (which they themselves undertake). Epstein and O'Halloryn (1999) expand on these arguments and show, also with US data, that principal-agent problems within the legislature shape legislative delegation to the bureaucracy.

Our analysis suggests that the problems of intra-legislative credibility are likely to be much deeper in countries lacking programmatic political parties. As Cox and McCubbins (1994) argue, partisan organization inside a legislature can resolve transaction costs, particularly the danger of opportunistic behavior by committee members. In the absence of programmatic parties, politicians can rely neither on statutory (*ex ante*) controls on the bureaucracy – they cannot act collectively to enforce those controls; nor can they rely on *ex post* supervision (they lack the ability to act collectively to ensure that oversight committee members will look after their interests in supervising the actions of the public administration).

Other research (e.g., Huber and McCarthy 2004) predicts that dysfunctional public administration, particularly agencies staffed with less capable employees, distorts politicians' policy choices. The analysis here suggests that underlying political circumstances distort both political decisions regarding the bureaucracy and public policies more generally. Where politicians are not organized into programmatic political parties, they prefer clientelist public policies (Keefer 2011) and rely on clientelist interactions with the public administration to ensure that these policies are carried out.

The analysis here also extends the existing literature on the political economy of public administration to non-democracies. Autocrats vary in the degree to which they choose to fend off potential challengers by providing benefits to large groups of citizens. Those that choose to provide broadly distributed benefits to citizens must make credible commitments to reward members of the ruling party, who implement these policies. As Gehlbach and Keefer (2010) argue, the credibility of ruler commitments to party members depends on whether party members can act collectively to limit the scope for ruler shirking. However, party members able to act collectively are also more likely to demand such administrative reforms to facilitate oversight of the ruler's pursuit of party goals. Public sector reforms are therefore also more likely to succeed in non-democracies with programmatic parties.

Finally, the results here make a contribution to our understanding of state-building. Tilly (1990) argues that the threat of war, and the growing costs of war, prompted rulers to expand the state and build up state capacity. Expensive wars required higher taxes; the ruler's ability to extract higher taxes depended in part on benefits that the state could offer citizens. Both activities demanded greater administrative capacity than rulers had previously needed. Keefer (2012) explains why strengthened bureaucracies may not emerge even in countries beset by war: where politicians are not collectively organized, an organized public administration may constitute a greater threat to the regime than external actors. The argument here identifies other, less dramatic circumstances in which the political demand for improved state capacity might emerge even in the absence of war.

Conclusion

This paper develops a new argument that relates public sector organization to the presence of programmatic parties. First, a large literature argues that relatively high levels of institutionalization are required for political parties to make credible programmatic commitments to citizens. We argue that the more organized are politicians, the more likely are voters to hold them accountable for public sector failures and the better able are politicians to enforce rules governing public sector performance. Second, politicians from parties with programmatic agendas are more likely to value public policy outcomes that depend on the efficiency of public sector organizations.

New evidence on World Bank public sector loans allows us to conduct an exact test of the proposition that programmatic parties promote more efficient public sector organization. These data provide evidence on the success of World Bank programs aimed at improving two sets of rules that are universally associated with more efficient public sectors: transparent, regularized personnel and financial management decisions. Using data on programmatic parties from the Database of Political Institutions, we find strong and robust evidence for the proposition.

Our results have significant implications for public sector reform. In discussing the timing of reform and of donor collaboration with client countries, practitioners and donors emphasize that one should take advantage of windows of opportunity. Those windows are usually considered open when influential politicians demonstrate enthusiasm for reform. The analysis here presumably holds enthusiasm constant: all loans generated enough enthusiasm to persuade a few top ministers, including the one in charge of the relevant sector, to approve the loan. The individual commitment of a few turns out to be insufficient, however. In addition, and critically, reform succeeds when a larger group of politicians benefits collectively from it – that is, when politicians are gathered under the umbrella of a programmatic political party.

Table 2. Logistic Regression: Programmatic Parties and Project Ratings for Public Sector Reform

	Inst. Dev. Impact Rating		Overall Outcome Rating	
	Base Model	Full Model	Base Model	Full Model
Programmatic Parties	4.019**	5.893**	2.915**	3.504*
	(0.002)	(0.003)	(0.006)	(0.029)
Log GDP per capita	1.196	1.108	1.088	1.079
	(0.303)	(0.531)	(0.483)	(0.585)
Log Population	1.623+	1.763	1.249	1.386
	(0.086)	(0.196)	(0.216)	(0.211)
Log Land Area	0.840	0.804	0.915	0.879
	(0.397)	(0.429)	(0.552)	(0.565)
Progression to Secondary School (%)	0.976	0.974	0.993	0.996
	(0.144)	(0.156)	(0.610)	(0.816)
Funding Amount	0.998+	0.998*	0.998*	0.998*
	(0.083)	(0.049)	(0.032)	(0.011)
Project Duration	0.981	1.059	0.887+	1.018
	(0.780)	(0.440)	(0.086)	(0.799)
Loan % Allocated to PSR	1.006	1.008	1.009	1.014
	(0.367)	(0.334)	(0.199)	(0.129)
IEG Quality at Entry		1.925***		3.603***
		(0.000)		(0.000)
Initial ICRG Bureaucratic Quality		0.790		0.612*
		(0.336)		(0.050)
Region Dummies	Yes	Yes	Yes	Yes
Observations	255	214	259	218
Pseudo R^2	0.174	0.289	0.083	0.300

Odds ratios reported; p-values in parentheses + $p < 0.10$, * $p < 0.05$, ** $p < .01$, *** $p < .001$
GDP per capita and population are averaged over the project period; progression to secondary school is a country average; programmatic parties and ICRG bureaucratic quality are measured at the project start year

Table 3. Logistic Regression: Programmatic Parties and Project Ratings for Public Sector Reform--Additional Controls

	Inst. Dev. Impact Rating		Overall Outcome Rating	
	Media and Resources	Political Controls	Media and Resources	Political Controls
Programmatic Parties	7.237***	2.632*	3.486**	2.862*
	(0.000)	(0.038)	(0.005)	(0.017)
Newspaper circulation	0.996		1.001	
	(0.475)		(0.845)	
Natural Resource Exports (% GDP)	1.056		0.946	
	(0.227)		(0.368)	
PR		1.143		1.687
		(0.767)		(0.244)
Political System		1.384		1.191
		(0.217)		(0.392)
Mean District Magnitude		0.994		1.005
		(0.482)		(0.480)
Democracy (ACLP)		1.552		0.856
		(0.450)		(0.662)
Baseline controls	Yes	Yes	Yes	Yes
Region Dummies	Yes	Yes	Yes	Yes
Observations	185	230	188	234
Pseudo R^2	0.246	0.182	0.129	0.100

Odds ratios reported; p-values in parentheses + p < 0.10, * p < 0.05, ** p < .01, *** p < .001
Programmatic parties and ICRG bureaucratic quality are measured at the project start year; all other variables are averaged over the project period

Table 4: Logistic Regression: Programmatic Parties and Project Ratings, Democracy Subsample

	Inst. Dev. Impact Rating		Overall Outcome Rating	
Programmatic Parties	7.362*	10.32*	6.334**	11.69*
	(0.012)	(0.016)	(0.003)	(0.012)
IEG Quality at Entry		2.181***		3.397**
		(0.000)		(0.003)
Initial ICRG Bureaucratic Quality		0.981		0.418**
		(0.935)		(0.009)
Baseline controls	Yes	Yes	Yes	Yes
Region Dummies	Yes	Yes	Yes	Yes
Observations	145	131	148	134
Pseudo R^2	0.266	0.337	0.142	0.337

Odds ratios reported; p-values in parentheses + p < 0.10, * p < 0.05, ** p < .01, *** p < .001
GDP per capita and population are averaged over the project period; progression to secondary school is a country average; programmatic parties and ICRG bureaucratic quality are measured at the project start year

Table 5. Logistic Regression: Programmatic Parties and Bank Supervision and Performance for Public Sector Reform Projects

	Bank Supervision Rating		Bank Performance Rating	
	Base Model	Full Model	Base Model	Full Model
Programmatic Parties	0.760	0.631	0.758	0.572
	(0.542)	(0.458)	(0.632)	(0.575)
IEG Quality at Entry		2.018***		5.069***
		(0.001)		(0.000)
Initial ICRG Bureaucratic Quality		0.962		1.505
		(0.883)		(0.132)
Baseline controls	Yes	Yes	Yes	Yes
Region Dummies	Yes	Yes	Yes	Yes
Observations	252	211	245	206
Pseudo R^2	0.155	0.259	0.136	0.554

Odds ratios reported; p-values in parentheses + p < 0.10, * p < 0.05, ** p < .01, *** p < .001
GDP per capita and population are averaged over the project period; progression to secondary school is a country average; programmatic parties and ICRG bureaucratic quality are measured at the project start year

Table 6. Instrumental Variables Probit Regression: Programmatic Parties and Project Ratings for Public Sector Reform

	Inst. Dev. Impact Rating	Overall Outcome Rating
Programmatic Parties	8.725**	4.322+
	(0.001)	(0.072)
Baseline controls	Yes	Yes
Region Dummies	Yes	Yes
First Stage: Programmatic Parties		
Latitude	1.017***	1.018***
	(0.000)	(0.000)
Baseline controls	Yes	Yes
Region Dummies	Yes	Yes
Observations	250	254
R^2		
Pseudo R^2		

Odds ratios reported; p-values in parentheses + $p < 0.10$, * $p < 0.05$, ** $p < .01$, *** $p < .001$
GDP per capita and population are averaged over the project period; progression to secondary school is a country average; programmatic parties and ICRG bureaucratic quality are measured at the project start year

References

Aldrich, John H. (1995) *Why Parties? The Origin and Transformation of Political Parties in America*. Edited by B. I. Page, *American Politics and Political Economy Series*. Chicago: University of Chicago Press.

Alvarez, Mike, José Antonio Cheibub, Fernando Limongi and Adam Przeworski. 1996. "Classifying Political Regimes." *Studies in Comparative International Development* 31(Summer):3-36.

Bawn, Kathleen (1997). "Choosing Strategies to Control the Bureaucracy: Statutory Constraints, Oversight and the Committee System." *Journal of Law, Economics and Organization* 13: 101-126.

Beck, Thorsten, George Clarke, Alberto Groff, Philip Keefer and Patrick Walsh (2001). "New Tools in Comparative Political Economy: The Database of Political Institutions." *World Bank Economic Review*, 15(1): 165-176.

Cain, Bruce John Ferejohn and Morris Fiorina (1987). *The Personal Vote: Constituency Service and Electoral Independence* Cambridge: Harvard University Press.

Calvert, Randall L., Mathew D. McCubbins and Barry R. Weingast (1989). "A Theory of Political Control and Agency Discretion." *American Journal of Political Science* 33:3, 588-611.

Chandra, Kanchan (2004). *Why Ethnic Parties Succeed* . Cambridge: Cambridge University Press.

Cheibub, Jose Antonio, Jennifer Gandhi, and James Raymond Vreeland. 2009. Democracy and Dictatorship Revisited. Public Choice 143(1-2): 67-101.

Cox, Gary W. and Mathew D. McCubbins (1994). "Bonding, Structure, and the Stability of Political Parties: Party Government in the House." *Legislative Studies Quarterly* 19:2, 215-231.

Eaton, Kent (2003). "Can Politicians Control Bureaucrats? Applying Theories of Political Control to Argentina's Democracy." *Latin American Politics and Society* 45:4, 33-62.

Epstein, David and Sharyn O'Halloran (1994). "Administrative Procedures, Information, and Agency Discretion." *American Journal of Political Science* 38:3, 697-722 (August).

_____ (1999). *Delegating Powers: A Transaction-Cost Approach to Policymaking under Separate Powers*. Cambridge: Cambridge University Press.

Ferejohn, John and Charles Shipan (1990). "Congressional Influence on the Bureaucracy." *Journal of Law, Economics and Organization* 6:1-20

Huber, Gregg and Nolan McCarty (2004). "Bureaucratic Capacity, Delegation, and Political Reform." *American Political Science Review* 98: 481-94.

Johnson, Ronald and Gary D. Libecap (1994). *The Federal Civil Service System and the Problem of Bureaucracy*. Chicago: University of Chicago Press.

Keefer, Philip (2011). "Collective Action, Political Parties and Pro-Development Public Policy." *Asian Development Review* 28:1, 94-118.

_____ (2012). "Why Follow the Leader? Collective Action, Credible Commitment and

Conflict." Chapter 32, *Oxford Handbook of the Economics of Peace and Conflict*, Michelle Garfinkel and Stergios Skaperdas, eds., 816 – 839. Oxford: Oxford University Press.

_____ and Razvan Vlaicu (2008). "Democracy, Credibility and Clientelism." *Journal of Law, Economics and Organization* 24:2, 371-406 (October).

_____ and Stuti Khemani (2009). "When Do Legislators Pass on Pork? The Role of Political Parties in Determining Legislator Effort." *American Political Science Review* 103:199-112 (February).

Knack, Steven (2008). "Sovereign Rents and the Quality of Tax Policy and Administration." World Bank Policy Research Working Paper 4773 (November).

Knack, Stephen and Philip Keefer (1995). "Institutions and Economic Performance: Cross-Country Tests Using Alternative Institutional Measures." *Economics and Politics* 7(3): 207-228.

McCubbins, Mathew and Thomas Schwartz (1984). "Congressional Oversight Overlooked: Police Patrols versus Fire Alarms." *American Journal of Political Science* 28: 165-79

McCubbins, Mathew, Roger Noll and Barry Weingast (1987). "Administrative Procedures as Instruments of Political Control." *Journal of Law, Economics and Organization* 3:243 – 77.

McCubbins, Mathew and Frances Rosenbluth (1995). "Party Provision of Personal Politics: Dividing the Vote in Japan." *Structure and Policy in Japan and the U.S.*, Peter Cowhey and Mathew McCubbins, eds., 35-55. Cambridge: Cambridge University Press.

Moe, Terry M. (1987). "An Assessment of the Positive Theory of 'Congressional Dominance.'" *Legislative Studies Quarterly* 12:4, 475-520.

Rauch, James and Peter Evans (2000). "Bureaucratic structures and economic performance in less developed countries." *Journal of Public Economics* 75, 49–71.

Shefter, Martin (1994). *Political Parties and the State: The American Historical Experience*. Princeton, New Jersey: Princeton University Press.

Siavelis, Peter M. (2000). "Disconnected Fire Alarms and Ineffective Police Patrols: Legislative Oversight in Postauthoritarian Chile." *Journal of Interamerican Studies and World Affairs* 42:1, 71-98.

Snyder Jr., James M., and Michael M. Ting (2002). "An Informational Rationale for Political Parties." *American Journal of Political Science* 46 (1):90-110.

Thies, Michael F. (2001). "Keeping Tabs on Partners: The Logic of Delegation in Coalition Governments." *American Journal of Political Science* 45:3, 580-98.

Tilly, Charles (1990). *Coercion, Capital and European States, AD 990-1990*. Cambridge: Basil Blackwell.

van de Walle, Nicolas (2003). Presidentialism and Clientelism in Africa's Emerging Party Systems. *Journal of Modern African Studies* 41 (2):297-321.

Weingast, Barry R., and Moran, Mark J. (1983). "Bureaucratic Discretion or Congressional Control? Regulatory Policymaking by the Federal Trade Commission." Journal of

Political Economy 91 (October): 765-800.

World Bank (2008). *Public Sector Reform: What Works and Why? An IEG Evaluation of World Bank Support*. Washington DC: International Bank for Reconstruction and Development/The World Bank.