Governance Initiative Review Paper

J-PAL Governance Initiative

October 2019

Governance Initiative Review Paper

Contents

1. What do we mean by governance?

Governments around the world spend billions of dollars annually to provide basic services and development programs aimed at improving the lives of those living in poverty. At the same time a large number of foundations and international aid organizations channel their development dollars through government-run programs. But the effectiveness of such public spending is often compromised by a number of connected factors: policies that do not reflect the needs or wishes of the people, leakages due to corruption, lack of community participation, and poor oversight of public spending.

Until recently, there was little quantitative evidence about the extent of corruption and its consequences, especially on program effectiveness, even though it is widely accepted that effective public service delivery requires "clean" governments (Acemoglu 2010). Several factors make empirical evidence on the causes and consequences of poor governance in low-income countries difficult to obtain. First, the illegal nature of many corrupt activities means that data on the extent of poor governance is often hard to gather. Second, obtaining causal evidence on why poor governance persists requires the researcher to disentangle the channels of influence, which is often difficult to determine. Similarly, it is challenging to measure the extent or intensity of participation and its true impact. For example, how do we know if observed voter choices reflect people's true electoral preferences, or choices limited by poor information about politician performance?

The feasibility and relevance of policy-oriented research on governance, therefore, depends on researchers' ability to address these challenges. Recent advances in data collection, which include improved methods of quantifying corruption, now enable researchers to collect sophisticated measures of misgovernance (Fisman and Wei 2004, Olken 2007 and Olken 2009). Furthermore, the use of field experiments allows researchers to directly provide causal evidence on the drivers of good governance.

Despite the crucial importance of good governance for development, and the recent advances in the literature, including studies supported by J-PAL's Governance Initiative (GI), many questions about how to effectively improve governance remain unanswered. There is also a need to build a better bridge between field experiments and underlying economic theories, using trials to test theoretical predictions and thereby strengthening our theoretical understanding of the issues underlying good governance and service delivery.

As a response to this need, this updated Governance Initiative (GI) review paper summarizes empirically rigorous evidence on governance issues in low-income countries and identifies new directions for research. The objective is to provide information about how to improve participation in the political and policy process, reduce leakages in public programs, and improve state capacity, and to identify gaps and specific open questions in the governance literature which researchers within the Governance Initiative will seek to fill.

The literature review began by the Co-Chairs of GI identifying research areas that have been influential in the fields of governance in the economic and political science literature. Within these areas, the research team conducted an extensive keyword search in the top economics and political science journals and reviewed the papers. Sources for the keyword search included journals in economics and political science and unpublished working papers presented at conferences or available through universities and top research centers.

From this range of publications, we selected the subset of papers that presented empirical evidence on the topic and addressed the issue of causal relationships in what we judged to be a reasonably effective way given modern standards of empirical work. The main criteria for inclusion in the GI review paper were that the papers should present studies that provide empirical evidence on the causes and

consequences of poor governance in low- and middle-income countries based on a 'rigorous empirical identification strategy.' We included both randomized and nonrandomized studies and made particular effort to include studies that provide causal evidence on why poor governance exists and persists, and what can be done about it. This reflects the ultimate focus of GI on supporting empirical research. That said, papers that provide a theoretical rationale for empirical tests were prioritized.

We recognize that the term "governance" means different things to citizens, researchers, and policymakers. Indeed, there is great diversity in the research and policy domains with respect to the definition of governance, and no strong consensus for a single definition of governance has been reached yet. For example, the World Bank has defined governance differently over time. An early definition described governance as "the manner in which power is exercised in the management of a country's economic and social resources for development" (World Bank 1992). Later, a broader definition described governance as "rules, enforcement mechanisms, and organizations" (World Bank 2002). Along this line, a more recent definition described it as "the manner in which public officials and institutions acquire and exercise the authority to shape public policy and provide public goods and services" (World Bank 2007).

In this GI review paper, we will not attempt to capture all such areas, nor attempt to build links where no natural links exist between areas. Further, we will not follow a particular operational definition of governance. Rather, throughout most of this review paper, we will cast the problem of good governance as a two-part principal-agent problem. In the first part of the problem, the principals are the citizens, who vote to elect as their agents politicians, whose job is to enact policies that are in line with the voters' interests. In the second part of the problem, politicians, in turn, become principals themselves, who then must work with their agents, civil servants, and other service providers to actually implement those policies.

In low and middle-income countries (LMICs), both of these links—the link between the electorate and the politicians they select, and the link between stated government policy and the civil servants that enforce it—are frequently weak. When citizen control over politicians is weak, the results can range from politicians directing state resources to connected firms (e.g., Khwaja and Mian 2005) to politicians implementing programs that in their original design are cost-effective, but in the presence of corruption are not (e.g., Olken 2006). When control over civil servants is weak, the results can range from public works officials stealing funds (e.g., Olken 2007) to teachers and nurses who rarely show up to work (e.g., Kremer el al, 2005 and Banerjee et all 2008).

Given this framework, this GI review paper tackles the subset of governance issues we believe are most theoretically and practically fertile, focusing on two key links: improving voter control over politicians and policy; and improving the incentives for civil servants to do their jobs properly and without corruption.

In this updated version of the review paper, we also introduce a new framework to approach corruption. This broadened theoretical framework considers bribes as part of the equilibrium of a bargaining game between officials and citizens.

As the literature in governance grows, we will continue to update our theoretical frameworks for understanding governance issues to reflect the latest research.

2. Participation in the political and policy process

The period since the mid-1970s has witnessed a significant democratization of governance structures across the globe, a fact that is often described as the "third wave of democratization" (Huntington 1991). Using subjective political freedom indicators, electoral archives, and historical resources in 174 countries from 1960-2005, Papaioannou and Siourounis. (2008) identify the third wave of democratization as being associated with 63 incidents of permanent democratic transitions, three reverse transitions from relatively stable democracy to autocracy, and six episodes of small improvements in representative norms (borderline democratizations). By the end of the analyzed period, democracy was the most common de jure institution of governance. More recently, some have identified democratization episodes in the Middle East (the 2011 Arab Spring) as the fourth wave of democratization; however, the robustness of this phase of democratization remains unclear.

Alongside country-level democratization of governance structures there has also been growing policy decentralization to local elected government (Bardhan and Mookherjee 2006) and significant interest, by governments and NGOs, in requiring direct community participation in public good provision and monitoring of service delivery. Finally, some have argued that coordinated citizen participation may have been aided by the rise of social media and swift spread of cellphone usage in LMICs.

Informed by these observations, this GI review paper focuses on citizen participation along four dimensions: as voters in democracies, as leaders, as community monitors, and as participants to a decentralization process. Greater citizen participation across all four dimensions is widely considered a key form of citizen empowerment. Here, we examine the evidence on whether, and how, such empowerment influences policy outcomes and citizen satisfaction. Throughout, we highlight research priorities.

We divide our review into two sections. We start with the impacts of participation, where we consider four types of participation—voting, leadership, community monitoring, and decentralization. Next, we examine the determinants of participation, where we consider the influences of five sets of factors: institutions, demographics, information, government programs, and technology.

2.1. Impacts of participation

The presumption of democratic governance is that citizen participation in decision-making processes can influence, and possibly improve, public good provision and redistribution. Beyond its direct impacts on citizen well-being, the form of resource redistribution may also influence economic growth.

Here, we focus on the impacts of citizen participation on three sets of outcomes: policy outcomes (public policies and service delivery), economic growth, and citizen satisfaction. Under each outcome we discuss the likely impact of different forms of participation, either through voting, leadership, community monitoring, or decentralization.

2.1.1. Impact on policy outcomes

2.1.1.1. Participation as voters

A first set of papers examines whether variations in the demographic composition of the electorate and, therefore, the mix of redistributive preferences among voters influences policymaking in democracies. Here, much of the evidence comes from studies of suffrage extension in rich countries

(Miller 2008 and Lott and Kenny 1999). Exploiting cross-state variation in the timing of female suffrage in the United States, Miller (2008) shows that within a year of female suffrage, local public health spending increased by 35 percent and child mortality fell by 8 to 15 percent. One notable study in a LMIC context comes from Indonesia (Martinez-Bravo 2014). This study considers the first election post-democratization, where there is variation across villages in the timing of the first election (until which time the leader is appointed). The study finds that village electoral results match those at the district level more often when officials are appointed. An explanation is that appointed officials had to signal their alignment with upper levels of government by influencing voters at the local level to support a specific party in district elections, often through clientelism or vote-buying. Additionally, Beath et al. (2013c) conduct an innovative randomized evaluation in Afghanistan, and find that democratically elected councils achieve better targeting of a food aid distribution program compared to traditional (non-elected) councils. However, this improvement only occurs when the responsibility for the program is explicitly given to the council. In the absence of a clear mandate, food aid leakage increases.

Micro-evidence on how the congruence of local preferences strengthens the impact of citizen participation comes from Martinez-Bravo et al. (2011). They examine the impact of local elections on leaders' accountability in rural China. The introduction of elections at the local level, even in the context of an authoritarian regime, decreased the enforcement of unpopular policies like the one-child policy and land expropriation, and increased the provision of appropriate public goods at the village level. Expanding on this research, Martinez-Bravo et al. (2012) find that the introduction of local elections in Chinese villages also has redistributive impacts on public goods provision. Elections significantly increase total public goods expenditure at the village level by 27.2 percent in accordance with village-specific demand. For example, in villages with more farmland, expenditure on irrigation increased whereas in villages with more school-going children, expenditure on schooling increased. Elections also led to income redistribution, as the income of the poorest households before the elections increased by 28 percent, and income of the richest households before the elections decreased by 29 percent. The authors provide evidence suggesting the causal mechanism is not improved leader selection by the voters but rather increased leader incentives to align actions with voter preferences.

Would similar results be obtained in ethnically diverse environments, such as those in sub-Saharan Africa? There are theoretical reasons to believe they may not. Padró I Miquel (2007) develop a framework to analyze political accountability when societies are ethnically divided and institutionally weak. Corrupt leaders may be elected if their supporters fear even greater extraction from the opposition groups. Thus, the fear of an equally inefficient leader that favors another ethnic group is sufficient to reduce electoral accountability. Banerjee and Pande (2009) suggest a different channel: voters with strong ethnic preferences may choose to trade off politician corruption against the ethnic identity of the politician. However, the severity of such trade-offs and whether they can be reduced by designing institutions and providing information needs more research. Bandiera and Levy (2011) consider the possibility that ethnic diversity allows elites to form coalitions with minority groups to rule against the preferences of the majority poor. Indeed, taking advantage of Indonesia's unique village governance structure in which some villages are elite controlled whereas others are ruled democratically, they find that the difference in public service outcomes between elite controlled and democratically ruled villages is decreasing in the level of ethnic diversity in the village. Their results suggest that ethnic diversity can undermine democratic reforms and meaningful participation, resulting in outcomes similar to those found in elite controlled autocracies.

In addition, politicians' reelection incentives can also improve electoral accountability to voters' preferences, as shown by Janvry et al. (2012) in their study of a conditional cash transfer program designed to reduce school dropout rates in Brazil. Exploiting variation in the implementation of the program across different municipalities, they find that the reduction in school dropout rates was 36 percent larger in municipalities with first term mayors in comparison to municipalities in which mayors

could not be reelected. First term mayors on average adopted more transparent and effective implementation strategies. Indicative of the electoral rewards to good performance when voters are responsive and informed, they also find that the probability of reelection was 28 percent higher for mayors who were in the top quartile of program impacts.

Open research questions:

- *How do voters trade off preferences for competent or honest politicians with redistributive preferences? How can institutional design and information about available policies alter these trade-offs?*
- *Does greater voter participation lead to more emphasis on populist policies?*
- *What is the incidence of electoral malpractices, and what policy solutions are effective?*

2.1.1.2.   Participation as leaders

Voters care about at least two dimensions of a politician's quality: honesty or competence, and policy preferences. A median voter model of politics predicts that elected leaders implement the policies favored by a majority of voters, rendering legislator identity irrelevant to policy outcomes. A large empirical literature, however, finds that, to the contrary, there is a clear link between a politician's identity, policy preferences, and implemented policies.

An important source of evidence is the body of literature related to use of electoral quotas in order to provide representation for historically disadvantaged groups in representative democracies. Interestingly, the findings for ethnic minority representation differ significantly from those for female representation—which points to the importance of theorizing about the nature of initial discrimination. Much of this research comes from India, where policies have enabled natural experiments to identify causal impacts of quotas.

 In the case of ethnic minority representation, Pande (2003) exploits rule-induced variation in quota incidence to show that greater presence of ethnic minorities in Indian state legislature increases targeted transfers to these groups. At the village level, Besley et al. (2005) report similar evidence. Chin and Prakash (2011) find that increasing the share of seats for one minority group in India significantly reduces rural and aggregate poverty in the state, but increasing the share of seats for another minority group does not impact poverty, making it difficult to generalize poverty impacts of increased minority representation. Consistent with the idea that the targeting of quota policies does not necessarily translate into welfare increases for the group, Jensenius (2015) uses a matching technique to study quotas for Scheduled Castes (SC) in India and finds no effect of quotas on development indicators or redistribution to SCs. In a similar spirit, Bhavnani (2016) finds that SC quotas do not have a lasting impact after they are removed on the chance that SC politicians win elections. Chauchard (2014) finds that SC quotas do not affect stereotypes, but do improve perceived norms of inter-caste interaction.

In the case of female representation, Chattopadhyay and Duflo (2004) exploit random assignment of gender quotas for council head position in India to show the reservation policy increased the provision of goods aligned with female preferences, such as drinking water and roads. In Beaman et al. (2011) the authors show that female leadership increases participation by female villagers in village meetings which (consistent with arguments in the previous section) suggests that leadership effects on policy may also reflect greater participation by certain groups. Iyer et al. (2012) estimate the impact of female representation on crime by taking advantage of state-level variation in the timing of implementation of quotas which set aside one-third of all local government council positions in India for women. They find

that reservation caused documented crimes against women to increase by 46 percent, with an increase of 23 percent in reported rapes and 13 percent of reported kidnapping. Their findings reflect an increase in the reporting of crimes against women rather than an increase in the incidence of such crimes, and they further find heightened police responsiveness, with a 30 percent increase in arrests for crimes against women. Pathak and Macours (2013) show that female leadership at the local level results in better nutritional and educational outcomes for local children.

Another set of papers looks at how the gender of a politician impacts policy outcomes in the absence of gender quotas by using the fraction of elections won by a female in a close election against a male as an instrument for the fraction of seats won by a female politician in a given jurisdiction. Using this technique, Clots-Figueras (2011) finds that female state legislators in India spend more than their male counterparts on schools, female teachers, primary education, and hospital beds. However, the socioeconomic background of the female legislator matters in spending decisions. Women that hold seats reserved for historically disadvantaged groups invest more in health and lower levels of education and favor laws made to benefit women or redistribute wealth. Women not from these reserved categories do not favor such laws, invest in higher education, and actually reduce social expenditure, providing further evidence that identity of a politician matters in policy outcomes.

Using a similar identification strategy to estimate how politician gender influences education outcomes at the district level in India, Clots-Figueras (2012) finds that a 10 percentage point increase in female representation increases the probability that an individual in that district attains a primary education by 7.3 percentage points. However, these results are not found for rural areas or the sample as a whole, suggesting that female legislators favor urban over rural areas for investments in education. In another district level analysis, Bhalotra and Clots-Figueras (2014) evaluate how health outcomes are influenced by female legislators in India. They find that a one standard deviation increase in women's political representation in the birth year or two years prior to birth decreases neonatal mortality by 1.5 percentage points. An increase in women's political representation also increases the probability that a village has required public health facilities and that women receive antenatal care, institutional delivery, and child vaccinations.

Looking at close mayoral elections between men and women in Brazil, Brollo and Troiano (2016) find that female winners are less likely to engage in corruption, measured by random government audits. Female mayors also have lower probability of reelection than male mayors.

At the macro level, Jones and Olken (2005) provide further evidence that leaders influence policy outcomes, mainly though monetary policy. In their study, an exogenous change on leadership, measured by the death of a leader while in office, is associated with changes in the level of inflation. There is, however, less convincing evidence on the effect of leaders on fiscal and trade policy.

Whether leader preferences also influence the efficiency of policymaking (e.g., by reducing corruption) is less clear. One reason is that it is difficult to know ex ante what individual traits predict performance. One characteristic that the literature has examined is education level of politicians. In a cross-country study, Besley and Reynal-Querol (2011) find that democratic elections deliver leaders who are 25 percent more likely to be highly educated. In terms of impact, Besley et al. (2011) show more educated leaders generate more growth. The transition from a college-educated to an uneducated leader reduces growth by 1.7 percentage points per year over a 5-year post-transition period. There is also some evidence that voters care about leader education. Banerjee et al. (2010) conducted a field experiment in which voters were given information about politician characteristics. They found that this information caused voters to favor relatively more educated candidates (on this also see sections 2.2.3.). However, more evidence on the channels of influence is needed.

Beath et al. (2016) provide experimental evidence on how different electoral systems cause voters to trade off leader competence with leader preferences. They organized an evaluation around local council elections in 250 villages in Afghanistan. In half of the villages representatives were elected at-large, and in the remaining cases each village was divided into several districts and voters in a given district could only select one district representative. The authors find that the leaders elected in at-large elections were more educated, and leader policy preferences were more extreme in single-member elections. This is consistent with a theoretical model where villagers in single-member elections anticipate bargaining that will take place in the village council between district representatives, and thus strategically select their leader.

Finally, evaluating the efficiency trade-offs (if any) in selecting leaders depends on the supply of politicians. Specifically, do voters find that politicians who share their redistributive preferences are less qualified? If so, does this reflect strategic placement of candidates by parties? Tracing the pathways from these questions to final policy outcomes is an area where more research would be welcome.

Open research questions:

- *Which qualifications or characteristics of leaders influence the efficiency of policymaking?*
- *How do selection procedures influence leader choice and therefore policy? What are the implications for disclosure laws?*
- *Do more competent or honest politicians care more or less about redistributive preferences?*
- *What types of policies are most affected by leader qualifications?*
- *What is the role of compensation, selection, and other incentives in influencing who becomes a politician and how politicians behave?*

2.1.1.3.   *Participation as community monitors*

Government programs have long incorporated mechanisms for local oversight, but there is little evidence on the effectiveness of these programs. On one hand, community participation is attractive because it offers solutions to many problems commonly associated with centrally administered services and projects. Relative to outsiders, local communities may have better information on needed goods and services and may also face stronger incentives to recognize and respond to inefficiency or corruption in implementation. As a result, locals may do a better job of monitoring and taking action when services are poor or corruption occurs. And if services are run or projects implemented by local workers and officials, then their peers may have a better chance of ensuring honesty through social sanctions. On the other hand, there is a risk that local communities may be prone to elite capture. Some decision-making may also require skills that are not be available at the local level. The extent to which devolving decision-making to the community level is effective may depend on the decisions being devolved, local capacity, and the risk of capture by local elites.

To investigate these issues, Alatas et al. (2012) study whether community involvement in the targeting of beneficiaries for Indonesia's national cash transfer program can improve targeting accuracy and community satisfaction. In a randomly selected subset of 640 villages, community members held a meeting to rank order the households in the village from poorest to richest, and this list alone was used to determine eligibility for the transfer. In another set of villages, the government used proxy means testing to identify eligible beneficiaries by collecting data on household assets and other demographics. Using the central government's consumption-based poverty measure as the metric of comparison, the authors find that the community method of targeting performed worse than the proxy means test in accurately identifying households as below poverty line, but the difference in accuracy between the two methods

was not large enough to have a significant impact on poverty. However, the community method performed much better in terms of community satisfaction and perceived program legitimacy. Furthermore, evidence suggests that the community's selection of beneficiaries was based on their conception of poverty that goes beyond the consumption per capita definition and takes into account their greater knowledge of the earning potential and vulnerability of various households. Notably, an individual's self-assessed poverty status was better reflected in community driven targeting outcomes than in proxy means test outcomes.

A more recent body of experimental evidence has begun examining the impacts of community oversight on education and health outcomes. Banerjee et al. (2010) find that 92 percent of villagers in rural Uttar Pradesh, India were not aware of existing Village Education Committees (VEC), which supposedly monitored teachers and administrators. Working with the community on monitoring tools that revealed just how little children were learning at school and informing the community of their rights to push for change, however, prompted no increased activity by the community, no increased teacher effort and no improvement in education outcomes.

In contrast, Pradhan et al. (2014) conducted a randomized impact evaluation in Indonesia and find that certain approaches to strengthen school committees can actually improve learning outcomes. Conditional on receiving a grant, facilitating linkage between the school committee and the village council to improve the status of the school committee increases Indonesian scores by .17 standard deviations and girls' math scores by .11 standard deviations. The combined intervention of this linkage with having committee members democratically elected to allow representation of previously excluded groups has the largest impact, leading to a .22 standard deviation increase in test scores. Taken together, this research suggests that community participation can be influential in public good outcomes, but in the case of school committees, the effectiveness is greatly enhanced by reaching out to stakeholders outside the committee through elections and linkage.

Björkman and Svensson (2010) find that informing Ugandan citizens of the dismal state of local health service delivery and holding meetings between citizens and health workers to agree on "action plans" significantly reduce provider absenteeism, increase utilization, and improve health outcomes. One possible reason for the striking difference between these results and findings of Banerjee et al. (2010) in India is that in Uganda specific action plans were agreed upon, while in India the community was encouraged to develop their own approach to addressing the problem. Indeed, when the program in India gave a subset of communities training in how to hold remedial tutoring sessions for local children, many volunteered and reading scores in the communities rose. Not only were concerned individuals in the community given something specific they could do to make a difference, the solution (remedial reading camps) did not rely on cooperation from the government teachers. Nevertheless, a recent study in Uganda from Raffler et al. (2018) find that a similar intervention as the one conducted by Björkman and Svensson (2010) had no effect on health outcomes, implying that information-based community monitoring become less effective as baseline health conditions and service provision improve.

In one study in Kenya, communities were given money to hire additional teachers on short-term contracts (Duflo et al. 2012a). In some ways these local teachers looked similar to the para-teachers for which VECs in India are nominally responsible. But in the Kenya program, power over the contract and money for the teachers clearly rested with the school committees and the NGO behind the program. These additional teachers performed much better than regular teachers—showing up more and achieving higher test scores. Training of the school committees improved results further.

Similarly, Gertler et al. (2012) study the impact of Mexico's AGE program, a government program which finances parent associations and allows parents to participate in the management of primary school grants. By exploiting the variation in the geographic expansion of this program over time, the authors find that it leads to a reduction in grade failure by 7.4 percent and grade repetition by 5.5 percent for children

in grades 1-3. The authors attribute these results to greater parental involvement in school matters and their increased ability to communicate with and monitor teachers. At the same time, the program had no effect in extremely poor communities where parents may lack the ability or stature to have influence. These findings suggest that while the program proved beneficial for some students, it could have also increased educational inequality between the poor and the extremely poor.

Turning to the private sector, Malesky at al. (2019) studied whether the participation of firms in the law-making process in Vietnam would influence their compliance with regulations. They find that those that participated in the drafting of new regulation had more positive views of the government's regulatory authority, were more likely to allow auditors to examine facilities, and ultimately improved compliance with factory floor regulations.

A related issue is whether a community can act as a single actor to improve policy outcomes, such as in response to community-level incentives. Olken et al. (2014) studied the Generasi program in Indonesia, which offered block grants worth approximately USD$10,000 to villages to be used to improve child health and education. Village management teams decided how the grants are used. The study randomly selected sub-districts where 20 percent of the grant allocated in the subsequent year depended on measured performance on a set of indicators. The authors find that grants have a positive significant effect relative to a pure control group, and incentives offer a small improvement for health outcomes in the short run (18 months) but no difference for education outcomes. In the longer term (after 30 months) this difference disappears, as the non-incentivized villages seem to make up the difference. A recent long-run evaluation of Generasi (Olken et al. 2018) found that in general the program effects became more muted with time, suggesting the need for long-run impact evaluation as conditions change.

Although some of these results show community monitoring can have a positive impact, it is still possible that alternative mechanisms are more effective at monitoring and in turn improving outcomes. Olken (2007)—also discussed in more detail in Section 4.3.2.1.1. below—compares additional community participation with the alternative of strengthened centralized oversight and monitoring. To address the possibility of elite capture and corruption in local road projects, in Indonesia some communities were told that their project would be externally audited while in others community monitoring was enhanced. The threat of central audit is more effective in reducing corruption—although it is also more expensive—than community monitoring.

Another possibility for the differing success of community monitoring programs, and one which is linked to our earlier discussions on the impacts of voter participation, is that differences in demographic make-up influence the relative success of community monitoring. Björkman and Svensson (2010) find that the treatment effects of community monitoring were significantly larger in communities that were more homogenous in terms of ethnicity and income. However, Fearon et al. (2015) offer an example in which homogeneous groups benefit less from a community intervention due to higher baseline ability to solve collective action problems. The authors study the impact of a Community-Driven Reconstruction (CDR) program in Liberia, measuring outcomes using an individual-level matching funds experiment. CDR increased contributions in communities where the matching funds experiment was run in mixed gender groups, while no effect was observed in communities with women-only groups. Glennerster et al. (2013) offers relevant evidence against a universal negative effect of ethnic diversity. Using earlier census data to instrument for present-day ethnic make-up, they find that in post-war Sierra Leone, ethnic diversity does not lead to lower public goods provision.

Overall, we have limited evidence on the absolute and relative importance of social capital and preference diversity for the possibilities of community monitoring.

Other studies examine community-based programs which seek to directly strengthen economic development and political representation. Gugerty and Kremer (2008) study whether outside funding strengthens community organizations or simply affects the composition of these groups. They exploit a random phase-in introduction of a program that provided funding to women's community associations in rural Kenya and found that outside funding had only little impact on agricultural outcomes. Instead, funding changed the membership and leadership in the communities, weakening the role of the disadvantaged in favor of younger and more educated women. Beath et al. (2013a) conducted a randomized field experiment in 500 Afghan villages to determine whether development programs that mandate women's participation can improve the status of women in historically repressive and discriminatory societies. They find that such programs increased women's participation in village governance, community affairs, and income-generating activities and increased support for female input in village level decision-making. At the same time, treatment did not change women's role in intra-household decision-making or their general societal role.

Recent experimental evidence examined whether the effectiveness of community monitoring can be improved through external intervention. In evaluating ways to improve the exercise of bottom-up accountability in Peru, Sexton (2017) finds that public training workshops can backfire. Providing training to citizens on participatory budgeting, recall of elected officials, and political protests led to mixed results on local government quality and increased the likelihood of initiating recall against poor-performing officials, and of civil unrest. The information provided may have increased participants' disillusionment with formal political institutions and led them to disengage from personally costly processes like participatory budgeting; the intervention led to participation dropping by five percentage points.

These studies provide further evidence that institutional design and outside intervention can influence how community-based programs work and their impact on economic and social outcomes.

Open research questions:

- *How does the institutional structure of community monitoring determine its impact on policy? What are the channels of influence?*
- *Are some programs and policies more amenable to being monitored by the community compared to others?*

2.1.1.4.   *Participation through decentralization*

Increased citizen participation may also be achieved by decentralizing government to allow more local decision-making. The public finance theory of decentralization, dating in economics back to Tiebout (and even debates surrounding the U.S. Constitution in the late 18th century), suggests that decentralization is optimal if there is substantial preference heterogeneity, limited spillovers across jurisdictions, and limited economies of scale. Beyond this, there is an argument that having government closer to the people results in better citizen monitoring of politicians.

Decentralization literature focuses on the redistributive consequences of greater citizen participation, and attributes the change to differences in the set of citizen preferences being represented. Yet the connection between group characteristics and policy preferences is often inferential and we lack direct causal evidence on how the economic status of voters influences policy outcomes. The evidence base on the links between the distribution of citizen preferences and the quality of policymaking is also mixed. Existing evidence suggests that greater citizen participation in policymaking processes can create winners

and losers. For instance, Beaman et al. (2011) used data from village meetings in India to show that men and women differ in their preferences over public goods. Thus, the welfare implications of changing the extent of citizen participation are ambiguous.

While few studies compare policy outcomes pre- and post-decentralization, those that do typically report improved representation of local voter preferences (see Gadenne and Singhal (2013) for a review). Faguet (2004) compared expenditure in public investment projects before and after the decentralization reform in Bolivia and found that decentralization increased public investment in eight of the ten analyzed sectors, especially in areas (e.g. education), where needs (e.g. illiteracy) were high. Similarly, Foster and Rosenzweig (2002) evaluated the consequences of local democratization and fiscal decentralization in India and found that a shift toward democracy made villages 15 percent more likely to have a paved road, 80 percent more likely to have public irrigation facilities, and 22 percent more likely to have a secondary school.

Moreover, the nature of change in public goods provision reflected underlying demographics. Decentralization in villages with more landless households increased investments in paved roads, which favored labor and reduced investment in public irrigation, augmenting land rents. In contrast, Malesky et al. (2014) studied a rollback of decentralization in Vietnam using a difference in difference approach and found that this change led to improvements in outcomes that were considered priorities for the central government, such as transportation.

We should expect the net welfare impact to be positive if the population is relatively homogenous in terms of economic and ethnic characteristics. In such settings, the main influence of increasing citizen participation should be greater electoral accountability of the representatives. Some suggestive evidence on the net effect comes from Fisman and Gatti (2002). Using cross-country data, they find that countries with greater decentralization have less corruption. The clearly endogenous nature of the data—less corrupt countries may be more likely to decentralize—implies that we cannot draw firm conclusions. However, it does raise the possibility that increased electoral accountability improves average governance practices.

One form of decentralization, community driven development (CDD), allocates control over the selection and implementation of public goods and services directly to communities. Recent research considers the ability of CDD projects to solve collective action problems and how community leadership and development aid programs interact with and affect local development outcomes.

Many randomized evaluations of CDD programs show that they do not change long-term participation and governance outcomes. Reviewing evidence from randomized evaluations of CDD programs conducted before 2017 in Afghanistan, the Democratic Republic of Congo (DRC), Indonesia, Liberia, the Philippines, Sierra Leone, and Sudan, Casey (2018) finds that they can improve citizen satisfaction and effectively deliver public services in difficult environments. Participation also entails a time-cost for citizens, with little evidence on whether CDD programs fundamentally change local institutions, empower marginalized groups, or improve other channels for collective action. Casey et al. (2012) find that a community-driven development program in post-war Sierra Leone has a positive impact on provision of local public goods. However, there is no evidence that the program improves the voice of underrepresented groups (youth and women), nor that treatment communities are better able to take advantage of subsequent experimental development opportunities. Similarly, Humphreys et al. (2015) report null results from a large randomized study of CDD in eastern DRC: despite good implementation and an absence of spillovers or anticipation effects in comparison group, they fail to find program effects on various measures of participation, accountability, information, and leakage. Fearon et al. (2014) shows that a community-driven reconstruction program in Liberia led mixed gender groups in treatment villages to contribute more to a matching funds experiment.

In principle, community leaders may have substantial influence on local projects, either positive, such as by rallying the community and improving public good contributions, or negative, such as by taking credit for external programs and reducing their own effort. Kelsey and Recalde (2015) show that when democratically elected local leaders are selected to make contributions towards a local project in public, this increases contributions from other members in the community. Guiteras and Mobarak (2015) show that when constituents do not have clear information on who is responsible for implementing a sanitation program, local leaders seek and partly succeed in taking credit for the project. However, this effect disappears when constituents receive clear information about the project, either from the beginning or at a later stage.

Open research questions:

- *Does the decentralization of policymaking change (i) the nature of citizen participation, (ii) policy outcomes, and (iii) electoral accountability?*
- *What is the optimal role of community participation and community ownership of shared resources to improve citizen control over policy?*

2.1.2. Impact on economic growth

2.1.2.1.   Voting

The macro evidence on whether democracy, measured by voter participation or institutional form, causes economic growth is mixed.

However, turning to the reduced form impact of institutional quality, a large institutions literature (largely associated with Daron Acemoglu and James Robinson)—that has made use of historical variation in the nature of colonization—demonstrates that the historical choice of institutions casts a long shadow over current policies and limits the role of citizen participation and voting. For instance, Acemoglu et al. (2001) and Acemoglu and Johnson (2005) show that the colonialism-related differences between countries that cause variation in expropriation risk (e.g., Nigeria versus Chile for high and low expropriation, respectively) can explain sevenfold difference on income. Under this view the introduction of de jure institutions of democracy per se is unlikely to influence economic growth. In contrast, Papaioannou and Siourounis (2008) exploit democratic transitions to examine the economic impact of countries that move from autocracy to more consolidated representative institutions. Their estimates suggest that on average democratizations are associated with a 1 percent increase in annual per capita growth. However, they are unable to provide evidence on the channels of influence. More generally, Pande and Udry (2005) argue that the very limited sources of variation in colonial experience at the macro level constrain our ability to use cross-country regressions to identify channels of influence.

An older macro and even more reduced form literature, has sought to directly quantify the relationship between inequality and growth (Persson and Tabellini 1994). One possible channel through which democratization in unequal societies can reduce growth is by excessive redistribution (if the median income is far below the mean). This, in turn, reduces the work incentives faced by the more educated and richer voters (Forbes 2000). However, countering this is the possibility that redistribution helps reduce the adverse impact of credit market imperfections on growth outcomes. Overall, the macro literature that has sought to estimate the reduced form relationship between inequality and growth has yielded very mixed results (Banerjee and Duflo 2003).

Open research question:

- *Does greater democratization of decision-making processes influence economic growth? What are the channels of influence?*

2.1.2.2. Leaders

Changes in national leadership are sometimes mentioned as a factor that can explain sharp variations on economic growth across countries. Jones and Olken (2005) study whether changes in national leadership which are associated with leader deaths (and are, therefore, more exogenous) influence growth. They use variation in national leaders due to natural or accidental cause as an exogenous change on leadership to estimate the effect of leader quality on growth rates. Their results show that a one standard deviation increase in leader quality increases growth rates by 2.1 percentage points per year in autocratic regimes but produce no effect on democratic regimes.

At some level this is puzzling, since micro-evidence reviewed in section 2.1.1.2 shows that a change in the group identity of leaders alters redistributive outcomes. One possibility is that the average transition at the time of leader death leaves the group characteristics and policy preferences of leaders unaffected. It is an area where we would like to see more research.

Certainly, the evidence suggests that voters believe that leader choice within democracies matters for efficiency. In section 2.2.3 we discuss results from field experiments in Brazil and India where voters were provided information about leader corruption and performance, respectively. In both cases, we observe significant improvement in voters' ability to select leaders.

Open research question:

- *How do leaders influence economic growth outcomes?*

2.1.2.3. Community participation

We did not find any research on the implications of community participation (as monitors or otherwise) for economic growth outcomes. We encourage research on this question.

2.1.3. Impact on citizen satisfaction

2.1.3.1. Voting

Recent experimental evidence that examines citizens' preference for direct democracy in Indonesia and Afghanistan suggests that direct democratic representation increases citizen satisfaction. Olken (2010) provides similar evidence from Indonesia, where villages were randomly assigned to choose development projects either by representative-based meetings or by direct plebiscites. Plebiscites substantially increased citizen satisfaction and perceived legitimacy of policy outcomes: plebiscites increased by 21 percentage points the likelihood that citizens would state that the project choice was either very much or somewhat in accordance with their wishes; by 18 percentage points citizens' belief that they would benefit either very much or somewhat from the project; and by 10 percentage points citizens' belief that they would make use of the project personally. The use of plebiscites, however, had very limited to no impact on project choice relative to representative-based meetings. Taken together the evidence strongly suggests that the increased ability to participate explains most of the reported satisfaction.

In Afghanistan, Beath et al. (2013b) finds that direct democracy both improves citizen satisfaction and changes allocation outcomes. They conducted a study across 250 villages in which half were randomly assigned to select local development projects through a secret ballot referendum and the other half were assigned to select these projects through meetings convened by local village councils. In villages that used the secret ballot, there was less elite capture over resource allocation decisions, and the level of satisfaction was significantly higher among men and women than in villages that used the village meetings to select development projects.

Open research questions:

- *Is citizen satisfaction with democratic process independent of actual policy choices?*
- *How does information about policy outcomes influence citizen satisfaction with the process?*

2.1.3.2. Leaders

A large body of survey evidence suggests that voters prefer competent leaders. In India, Beaman et al. (2009) exploit random assignment of gender quotas for leadership positions on a village council to first show that women leaders perform at least as well as male leaders. They then use the random variation in mandated exposure to female leaders across village councils to examine whether exposure to female politicians influences voters' future behavior. Electoral data show that, compared to councils with no previous reservation policy, women in councils no longer reserved were twice as likely to stand for and win positions if the positions had been reserved for women in the previous two elections.

To determine whether the electoral gains were related to voters updating their beliefs about the effectiveness of female leaders, the paper presents two types of evidence—vignette and speech evidence—in which the gender of the leader described (or heard) was randomly varied and evidence from implicit association tests (IAT, which investigated villager stereotypes regarding the association of women with domestic jobs and men with leadership jobs) was examined. The evaluation gap in vignettes disappears in currently or previously reserved villages. Similarly, IAT-based measures of gender-occupation stereotypes show that exposure to female leaders through reservation increased the likelihood that male villagers associated women with leadership activities (as opposed to domestic activities).

Corroborative evidence comes from Bhavnani (2009), who examined similar questions in an urban setting: municipal councilor elections in Mumbai. He finds that the probability of a woman winning office conditional on the constituency being reserved for women in the previous election is approximately 5 times the probability of a woman winning office if the constituency had not been reserved for women. He finds suggestive evidence that an important channel was party officials learning that women can make effective leaders and win elections. However, returning to the West Bengal setting, the authors also found that explicit attitudes towards female leaders remain unchanged and significantly more negative than those towards male leaders.

These results suggest that voters are willing to update their beliefs about the likely performance of politicians drawn from different groups. However, they also suggest that typical perception-based survey measures of satisfaction may not be correlated with either the performance of the group being evaluated and, indeed, voters' own willingness to select leaders from that group. In effect, Duflo and Topalova (2004) compare objective measures of the quantity and quality of public goods with information on how villagers evaluate the performance of male and female leaders in India. They show residents of villages headed by women were on average less satisfied in spite that the quantity and quality of public goods provided were as high as in non-reserved villages. Duflo and Topalova suggest two likely explanations for this finding. First, new leaders are often judged less favorably than established leaders (women leaders

are generally poorer, less experienced, less educated and less literate). Second, villagers generally expect women to be less effective and these priors adjust slowly. In both cases, the results suggest women start with a disadvantage when running for political positions.

One important take away from this literature is the importance of measuring actual policy outcomes in order to interpret the policy implications of attitude-based measures of performance.

Open research questions:

- *Are attitudes towards leaders influenced by performance information?*
- *What are the links between leader performance and satisfaction outcomes?*

2.1.3.3.   Community participation

Some evidence suggests that community participation in the monitoring or implementation of social service programs can improve citizen satisfaction in some contexts.

Alatas et al. (2012) (also discussed in 2.1.1.3) find that community involvement in the targeting of beneficiaries for Indonesia's national cash transfer program improved community satisfaction. Compared to a proxy means test, the community targeting method performed much better in terms of community satisfaction and perceived program legitimacy. The community targeting method also better reflected an individuals' self-assessed poverty status.

There is also evidence that direct involvement in the political process increases citizens' satisfaction with the process. Specifically, Olken (2010) and Beath et al. (2016) find that villagers have higher satisfaction and better perceptions of the local government when they get a direct say in the choice of local projects through a referendum, compared to consultative meetings, which are in theory open to everyone but in practice dominated by elites. We regard this area as important for future research.

2.2. The determinants of citizen participation

Existing evidence suggests that increased democratization of governance structures typically improves the representation afforded to citizens. However, several factors might in practice limit citizens' willingness or ability to participate. Much of the skepticism surrounding the use of citizen participation as a way of improving governance in these settings reflects a general concern that weak institutions, electoral malpractice, elite capture, and limited information flows constrain the type and impacts of participation such that the political elite largely determines policies.

Against this background, we ask what forces constrain the type of participation (and therefore its ability to influence policy) in low-income settings and the implications for public policy. We organize the section in terms of five characteristics of low-income settings that are often considered as key determinants of participation: institutions, demographics and identity, information flows, experience with government programs, and technology. For each characteristic we first ask whether each may influence participation and then we review the evidence (and open areas of research) on how policy innovations can either change the fundamentals, enhance or mitigate their impacts.

2.2.1. Institutions

2.2.1.1.   Weak institutions: electoral malpractice and entry barriers

Weak institutions can undermine voter participation by enabling electoral malpractice such as vote-buying, intimidation, or ballot stuffing. This, in turn, can worsen politician selection and reduce accountability or incentives for politicians to perform. Direct evidence on the importance of these channels in linking weak electoral institutions or electoral malpractice with poor governance is scant. There is, however, indirect evidence that electoral malpractice reduces the quality of elected officials and reduces the representation afforded to the relatively poor.

Golden and Tiwari (2009) use Indian data to show that parties are more likely to list criminal candidates if they are on the cusp of winning or losing, hence face a great deal of uncertainty ex ante. It is possible that a benefit of fielding criminal candidates is that they are more likely to engage in electoral malpractice. However, it is also possible that more competitive elections reflect, for instance, significant ethnic divides in the population and the latter is correlated with candidate criminality. More structured evidence on this issue comes from Acemoglu et al. (2013). They argue that one reason why the Colombian state may allow violent non-state groups, such as the paramilitary, to exist is because these groups can influence electoral outcomes, a benefit for certain central state actors. For example, they show that in areas with high paramilitary presence, the candidate favored by the paramilitary had, on average, an 11 percent higher vote share than the competing candidate. A third piece of indirect evidence comes from Baland and Robinson (2008). They hypothesize that the introduction of the secret ballot in Chile should reduce the electoral power of the landed elite, who earlier controlled the workers' vote. Consistent with this, they show that the introduction of a secret ballot reduces right-wing bias in electoral outcomes in areas with greater land inequality. However, much of this evidence remains indirect; there is significant need for more research on whether and how weak institutions increase electoral malpractices.

Another example of weak institutions is elite capture of the entry process, such as through dynastic politics. However, the evidence base is thin. Rossi (2010) uses a natural experiment to test the hypothesis of self-perpetuation on power in Argentina. With the return of democracy and the formation of the Congress in 1983, the duration of elected legislators' terms was randomly allocated, with half of the legislators elected for two year terms and the remaining for four year terms. The results show that having a longer tenure in Congress increases the probability of having a relative enter Congress in the future by approximately 8 percent.

None of these papers examine whether political dynasties are worse for governance outcomes. Ex ante, the effect could go either way. Olson (1993), for instance, argues that in environments with weak institutions, dynastic politicians may be more willing to undertake policies that have limited short-term benefits but significant long-term benefits. Thus, long-term investment and growth outcomes may improve. It is clear that there is need for more research here.

Elite control can also undermine the role of participatory processes in LMICs. In Sao Tome and Principe, Humphreys et al. (2006) examined whether a participatory approach to decision-making was vulnerable to manipulation by political elites. In their study, the leaders who moderated a public consultation process to discuss policy priorities were randomly assigned to groups throughout the country. They show a strong influence of the leaders in the outcome of the deliberation. Groups were more likely to prioritize investment that reflected the preferences of their leader. This result suggests that the way these participatory processes are structured matters. However, the study does not necessarily provide evidence against participatory processes as it is not able to compare the final outcomes with and without them.

Some of the most prominent evidence on how the poor and less educated are more susceptible to vote-buying comes from the era of urban machines in the United States. Drawing a comparison between the political process in the (at the time) new nations such as Malaysia and the Philippines and the urban political machines in U.S. cities, such as the Tweed machine in 19th-century New York or the Dawson

machine in 20th-century Chicago, Scott (1969) argues that the most fundamental quality shared by the mass clientele of the machines was poverty. Stokes (2005) shows that machine politics in Argentina target the poorest and Blaydes (2006) reports similar evidence for Egypt. Further, Blaydes (2006) shows that in Egypt, illiterates turned out at more than twice and sometimes at three times the rate of literates in the 2000 and 2005 parliamentary and 2005 presidential elections. She argues that this turnout difference reflects greater vote-buying among the poor in Egypt.

A first hypothesis is that the immediate cash benefits of vote-selling is high for the poor—within this view, democratic governance in low-income settings will always be susceptible to corruption of voters (Stokes 2005). Yet it is also the case that the poor are often the group that suffers the most from low-quality service delivery. This would suggest that it is relatively expensive for the poor to sell votes. A second possibility is that the poor are more easily intimidated.

In a survey of Nigerian voters, Bratton (2008) documents that almost 20 percent of Nigerians are personally exposed to vote-buying and almost 10 percent experience electoral violence. He finds that political intimidation was not correlated with income, gender or other demographic characteristics, suggesting violence is a general condition that affects individuals equally during election time. On the other hand, he finds that vote-buying is correlated with income and education. Poor voters were more likely to be approached by a politician who offered to buy their vote, while this is less likely to happen for educated voters.

This could either reflect the pure income effect discussed above or the possibility that these practices are sustained by reciprocal beliefs. Thus, if low-income settings are characterized more by informal institutions and norms of reciprocity, then vote-buying may persist. Combining individual survey information on vote-buying experienced in a 2006 municipal election in Paraguay with measures of reciprocity obtained from lab experiments carried out in 2002, Finan and Schechter (2012) find that reciprocal individuals are more likely to experience vote-buying. This evidence suggests that social preferences may help sustain vote-buying practices. To the extent that reciprocal relationships are more important in low-income settings we may infer that it is easier to sustain vote-buying among the poor.

Recent experimental evidence examines whether targeted interventions have the potential to reduce vote-buying, with mixed evidence to date. Larreguy et al. (2018) study how a large-scale campaign against vote-buying affects not only citizens' willingness to sell their votes but also politician and party behavior in the 2016 election in Uganda. They do not find evidence that the campaign significantly reduced the extent to which voters were offered (and accepted) cash or gifts in exchange for their vote. Nevertheless, the campaign reinforced opposition candidates as researchers found evidence that voters took the reward offered by politicians but nonetheless voted for their preferred candidate. On the other hand, Green et al. (2016) evaluate a non-partisan anti-vote-buying radio campaign during the 2014 Indian general election. Using electoral data, they find that exposure to the radio campaign significantly decreases the vote share of putative vote-buying parties, with estimates ranging from 4 to 7 percentage points.

2.2.1.2.   *Can policy reduce the impact of weak institutions?*

Even when elections look clean and fair while in progress, there are other types of preelection irregularities—such as voter registration fraud—that can negatively impact political participation. Ichino and Schundeln (2010) conduct a randomized field experiment deploying domestic observers during a 13-day voter registration period in the 2008 general elections in Ghana. They find that observers can help reduce registration programs through a direct effect on the electoral area where they were deployed. However, this effect may be reduced by spillovers in other areas. The increase in voter registration from

2004 to 2008 was, on average, 3.5 percentage points smaller in electoral areas with observers. However, some of this reduction was due to displacement rather than absolute effects: voter registration increased by 2.8 percentage points in electoral areas near treatment areas. This evidence suggests that policies aimed at controlling fraud during elections can fail if they push misconduct to less visible stages of the election process. Whether a more comprehensive strategy for monitoring before, during, and after elections can be more successful at reducing irregularities is an interesting area for future research.

2.2.1.3.   *Political parties and participation*

Recent research indicates that the composition and ideology of political parties along with intra-party politics can exert significant influence over the nature and outcomes of participation.

An important question is how parties promote politicians among their ranks and whether this process is efficient. Jia et al. (2013) show that provincial leaders in China are more likely to be promoted when economic growth is higher, and this relationship is stronger for leaders who are already connected with top politicians. The authors then argue that this pattern may not be distortionary from the point of view of talent allocation, because connected leaders are more likely to be loyal.

Dunning and Nilekani (2013) take advantage of a natural experiment and employ a regression discontinuity design and find that ethnic quotas for village council presidents in three Indian states do not induce greater targeting of jobs, benefits or government welfare schemes based on caste or tribe. Instead, they find that political parties, which are multi-caste, are more likely to target benefits along party lines rather than ethnic lines, and voters in turn seem to value party affiliation more than ethnicity when stating their voting intentions.

Aside from ethnic minorities, another group that often lacks political representation is women, especially in socially conservative societies. Meyersson (2014) explores whether rule by conservative political parties, namely Islamic parties, impedes or bolsters women's empowerment. He utilizes a regression discontinuity design to compare outcomes in Turkish municipalities in which a prominent Islamic party, the Refah Party, barely won or barely lost the mayoral elections of 1994. He finds that having a mayor from the Refah Party significantly increased female high school completion and decreased adolescent marriage, and the effects are more pronounced in municipalities that are poorer and more religious. In these environments, where parents are more conservative, Islamic leadership was able to reduce barriers to education faced by females by setting up schools sponsored by religious charities and ignoring the nationwide headscarf ban, making parents more willing to send their daughters to school. In the long run, Meyersson finds that female political participation increased in these municipalities relative to the control, with more women elected to the municipal council.

Internal decisions and electoral processes of political parties can also influence participation decisions and voting outcomes, as highlighted by research done in Ghana by Ichino and Nathan (2013). Before the 2004 and 2008 parliamentary elections, parties held primaries in some constituencies and not in others. Exploiting this variation, the authors find that the electoral impacts of holding primaries is different for the two major Ghanaian parties. In the ruling party, primaries reduce the vote share and the probability of winning the election, whereas for the opposition party, primaries increase vote share and probability of winning the election. These results are in line with the authors' theory of primaries based on patronage relationships in which the ruling party is at greater risk of facing defections and internal disputes after the result of a primary election. Voters respond to such disputes and defections by switching their vote for the opposition party who faces less internal setbacks from the primaries and is able to present a more unified platform.

Open research questions:

- *How can elite capture of electoral processes be constrained in low-income settings?*
- *What are the strategic responses of political actors to transparency initiatives during elections?*
- *What is the impact of different institutional arrangements in influencing the political voice afforded to different citizen groups, and what are their implications for policy outcomes?*
- *What is the incidence of electoral malpractices, and what policy solutions are effective?*

2.2.2. Demographics

2.2.2.1. The role of demographics

2.2.2.1.1. Ethnicity

In the governance literature, demographic concerns in low-income countries typically relate to ethnic diversity. Using cross-country data, Easterly and Levine (1997) found a strong negative relationship between ethnic diversity and public good provision, political stability, and long-run economic growth. According to them, Africa's high ethnic fragmentation explains a significant part of the low levels of schooling, political instability, underdeveloped financial systems, government deficits and insufficient infrastructure among other determinants of long-run growth. On the other hand—and contrary to existing assumptions about ethnic preferences—Berge et al. (2015) conducted multiple lab studies in Kenya and find no evidence of ethnic bias in decision-making related to sharing of economic benefits and political alignment.

It is likely that in low-income settings an important reason why ethnic diversity limits economic development and political participation is an inability to overcome public good free rider problems due to monitoring and enforcement limitations across ethnic groups. Miguel and Gugerty (2005) find that ethnic diversity is negatively related to local funding of primary schools and community water wells in western Kenya, because ethnically diverse communities have lower contributions at public fundraising events. Their estimates show that a change from complete ethnic homogeneity to the average school ethnic diversity leads to a drop in school funding equal to 20 percent of mean local school funding per pupil, and this relationship is robust to socioeconomic, geographic, and demographic controls. In addition, ethnic diversity is associated with poor maintenance of wells—areas with average levels of ethnic diversity are 6 percent less likely to have functioning water wells compared to homogenous areas—which suggests that ethnic diversity may have implications for collective action beyond school funding.

Habyarimana et al. (2006) conduct a field experiment in Uganda to understand whether ethnic fragmentation undermines cooperation in the provision of public goods. The researchers used a number of experimental games to compare patterns of play among co-ethnic and non-co-ethnic groups. The results of the study suggest that in similar settings, co-ethnics are more likely to play cooperatively, plausibly due to the threat of social sanction.

At the same time, there is some evidence that ethnicity is less politically salient when cross-cutting cleavages or geographic considerations offset the influence of ethnicity in driving vote choice. Dunning and Harrison (2010) run an experiment in Mali to understand the importance of historical alliances called "cousinage" relative to ethnicity in voting behavior. They exposed experimental subjects to video-taped political speeches where the content was held constant but the last name of the politician, which in Mali conveys both ethnicity and cousinage, was changed. They find that both cousinage ties and ethnic ties are politically salient, and the causal impact of cousinage on preferences for candidate works to offset the

effect of ethnicity. In addition, individuals reported greater trust and expectations of politician responsiveness from cousins than non-counsins. Similarly, Ichino and Nathan (2013) use polling station data from Ghana and find that individuals living in rural areas are less likely to vote for a party that identifies with their own ethnic group when another ethnic group makes up a larger proportion of the population in that geographic area. In this case, voters support the party identified with the more dominant ethnic group in anticipation that when elected, the politician will bestow areas in which his ethnic group has a majority with non-excludable public goods.

One continuing concern with this literature is researchers' limited ability to control for unobserved regional characteristics, which vary alongside ethnic diversity. In settings where recent migration has been high, results look different. As discussed in Section 2.1.1.3, Glennerster et al. (2010) provide some counter-evidence from Sierra Leone, one of the world's poorest and most ethnically diverse countries; they find no correlation between local ethnic diversity and local public goods or collective action outcomes, such as road maintenance, community group membership, trust, and school funding and staffing.

2.2.2.1.2.   *Gender*

As evidenced throughout the literature discussed in above sections on participation and governance, many lab and field experiments have identified gender as a salient factor in political participation. Evidence suggests that gender directly influences political participation and selection and that gender bias limits women's access to public office. For example, in Uganda, Braton (1999) finds evidence that gender has a greater influence on political participation than socioeconomic status. In a study in India, Besley et al. (2005) find that women are less likely to participate in village meetings and that higher participation influences policy outcomes including beneficiary selection for welfare programs. Evidence also suggests that women are disadvantaged when running for political positions (also discussed in 2.1.3.2). Duflo and Topalova (2004) show that residents of villages headed by women in India were on average less satisfied, despite the fact that the quantity and quality of public goods provided were as high as in non-reserved villages. The findings suggest that villagers generally expect women to be less effective leaders and that these priors adjust slowly.

Gender quotas are one policy tool that can increase women's representation in politics. Evidence suggests that gender quotas can indeed improve women's representation and subsequently lead to changes in the provision of public goods. In the long run, gender quotas can also change beliefs and aspirations around female leaders. Increasing women's participation in community-based programs can also improve the status of women in historically repressive and discriminatory societies. For example, Beath et al. (2013a) find that development programs in Afghanistan that mandate women's participation can increase women's participation in village governance, community affairs, and income-generating activities and increase support for female input in village level decision-making (also discussed in 2.1.1.4).

There is a rich literature that examines the impact of gender quotas on political participation and service delivery. Chattopadhyay and Duflo (2004) show that gender quotas in India can increase the provision of goods aligned with female preferences, such as drinking water and roads. Beaman et al. (2011) find that female leadership increases the participation of female villagers in village meetings, suggesting that leadership effects may reflect greater participation by certain groups. Iyer et al. (2012) find that female representation leads to an increase in the reporting of crimes against women as well as heightened police responsiveness to such crimes. Pathak and Macours (2013) show that gender quotas can improve a child's early life nutrition and educational outcomes. Ghani et al. (2014) show an increase in new businesses owned by women after the introduction of quotas in India, consistent with a change in aspirations of women entrepreneurs.

Existing evidence also examines how a politician's gender impacts policy outcomes in the absence of gender quotas (also discussed in 2.1.1.2). Clots-Figueras (2011) find that female state legislators in India spend more than their male counterparts on schools, female teachers, primary education, and hospital beds, although the socioeconomic background of the female legislator matters in spending decisions. Brollo and Troiano (2016) examine close mayoral elections between men and women in Brazil and find that female winners are less likely to engage in corruption, as measured by random government audits.

Increased female political participation, such as through gender quotas, can also lead voters to update their beliefs about the likely performance of female leaders (also discussed in section 2.1.3.2.). Beaman et al. (2009) study gender quotas in India and find that women leaders perform at least as well as male leaders. Furthermore, they find that women in councils no longer reserved were more likely to stand for and win positions if the positions had been reserved for women in the previous two elections, compared to village councils with no previous reservation policy. Similarly, Bhavnani (2009) finds that the probability of a woman winning office is higher if the constituency was reserved for women in previous elections. Evidence suggests that an important channel was party officials learning that women can make effective leaders and win elections. Bhalotra et al. (2016) find that electoral gains for women were related to voters updating their beliefs about the effectiveness of female leaders. Similarly, they find that exposure to female leaders through reservation increased the likelihood that male villagers associated women with leadership activities. However, they also find that women were less likely to enter politics in areas with existing female leaders and entrenched gender bias, suggesting a backlash effect.

We encourage future research that examines the drivers of gender exclusion in political and policy processes, and policies that work (or do not work) to reduce gender bias in political participation.

2.2.2.1.3. *Education and socio-economic status*

Other demographic characteristics which have been found to directly influence political participation and politician selection include education and socio-economic status. Using household and village survey data from South India, Besley et al. (2005) study who participates in village meetings called by elected local governments, and whether these meetings have an influence on beneficiary selection for welfare programs. The results show that an individual is more likely to attend a Gram Sabha meeting if that person is literate, male and landless[1]. Higher participation influenced the allocation of resources. Landless individuals in a village that held a Gram Sabha were 7 percent more likely to received targeted benefits. The results suggest that this type of opportunity for political participation can help improve the targeting of resources towards the neediest groups.

Along these lines, the political science literature has also examined how socioeconomic factors influence political participation, with mixed results. In a cross-sectional study for democratic countries, Blais and Dobrzynska (1998) find that election turnout increases with the level of a country's income and literacy rate. Using survey data from Uganda, Bratton (1999) suggests gender is a greater influence on political participation than socioeconomic status. In Latin America, Fornos et al. (2004) find that, unlike institutional and political variables, socioeconomic variables are unrelated to turnout.

2.2.2.2. *Clientelism*

---

[1] The Gram Sabha is a meeting of all adults who live in a common area. Anyone who is 18 years old or more and who has the right to vote is a member of the Gram Sabha.

A related growing literature is that on clientelism. Clientelist relationships between political leaders and the population often times take the form of ethnic patronage, which can influence the nature of citizen participation, public good provision, and development outcomes.

Burgess et al. (2015) provide evidence for the role of clientelism along ethnic lines in hindering development outcomes, with evidence from historical road building in Kenya. Districts in Kenya are proxies for ethnicity, and the authors construct a district-level panel dataset of road construction over time since Kenyan independence. Using data from both expenditures and maps, the authors compare actual roads to a counterfactual situation of equivalent-length road construction that would provide the greatest connectivity by population. They find that road building occurs disproportionately in areas that share ethnicity with a sitting president during less democratic periods, but that these effects are mitigated during more democratic periods.

Anderson, Francois, and Kotwal (2015) study local government in rural Maharashtra, India, and identify clientelist vote-trading structures operating in a context of free, highly participatory elections. The researchers use data collected from nearly 10,000 households to show that villages dominated in land-holding by the Marathas—a caste of both landholders and peasants—have lower provision of public goods and services but more informal insurance. The authors interpret this as evidence of clientelism and argue that the Maratha landlords artificially depress provision of public goods to make the poor more dependent on them. Landlords compensate by providing insurance in times of distress, which allows them to rely on the votes of the poor.

Another strand of research seeks to understand how clientelist networks function, and how they interact with and are affected by voting and competition. The degree of competition between political actors, and the relative sizes of their support groups, may have an influence on development outcomes. Acemoglu et al. (2014) study paramount chiefs in Sierra Leone, and find that areas with fewer ruling families, a proxy for a low degree of competition for prospective chiefs, have worse outcomes in education, income and health today. Nevertheless, in the same areas, chiefs are more respected and social capital is higher, which the researchers attribute to the chiefs' control over civil society.

How do national parties incentivize brokers to control local clientelist networks? Larreguy (2013) studies this principal-agent problem empirically by exploiting variation in the amount of information that the national PRI party in Mexico can extract about brokers operating in communal land areas. Specifically, the researcher compares communal lands that overlap better or worse with electoral areas. In a simple model, areas with better overlap give the PRI a more informative measure of broker effort, allowing the party to incentivize the broker more strongly. The paper indeed finds evidence that the vote share of the PRI is higher in such areas, when the party has access to resources to incentivize brokers.

Larreguy, Marshall, and Querubãn (2016) explore turnout buying by parties in Mexican elections. Their model shows that the extent of turnout buying depends on the party's ability to monitor local brokers, as well as the availability of favorable voters who can be induced to turn up to vote. The researchers then exploit variation in the number of polling stations in an electoral precinct, as mandated by sharp cutoffs in the number of registered voters in the precinct. They find that an increase in the number of polling stations increases turnout as well as the number of votes for the main parties (PRI and PAN).

Fafchamps and Labonne (2016) use a rich dataset of politicians, individuals, social networks based on blood and marriage ties and a bouquet of public services to understand the nature of clientelism in Philippines. Households that are deemed to be more central (using the "betweenness centrality" measure defined in network theory) receive more services from their municipal governments. This result throws up some interesting heterogeneous findings: one, central relatives of either the winner or the

runner-up are less likely to get public services; two, betweenness centrality is a major factor where elections are close or in more populous municipalities; finally, in municipalities where politicians target more goods to their relatives, they target less of these to those with high betweenness centrality.

2.2.2.3. Can policy reduce the impact of demographics?

While ethnicity can influence political participation and public good provisions, there is also evidence that salience of ethnic identities increases during election time, reinforcing the initial relation. Eifert et al. (2010) collect data from ten public opinion surveys in Africa and find evidence that ethnic identities in Africa are strengthened during periods of political competition. On the one hand, these results suggest that playing the ethnic card could be an efficient instrument for politicians to mobilize voters. On the other hand, the higher level of ethnic identity could be explained by voters reacting to the belief that public good provision is concentrated in the ethnic group of the ruling party. In both cases, elections might be strengthening the negative effect of ethnicity on political accountability (Padró I Miquel 2007), and corruption (Banerjee and Pande 2009).

A plausible interpretation is that electoral competition creates fluctuations in ethnic saliences that are correlated with the timings of elections. It is possible that this reflects strategic behavior by political parties. As we discuss in section 2.2.3, limited information about candidate qualifications and performance may be an important reason why ethnicity or other group characteristics become dominating forces in elections in low-income and low-information settings.

Ethnic tensions at the time of elections may interact in a negative way with non-partisan civic campaigns. Marx et al. (2016) study an initiative by Kenya's newly formed Electoral Commission to stimulate voting participation by sending text messages to registered voters. Turnout was indeed higher in the randomly selected polling stations where voters received text messages. However, contrary to the original hope that the campaign would also improve trust in the new electoral institutions, there was a significant fall in trust, driven by voters from the Luo and allied tribes, who supported the candidate who eventually lost. The researchers argue that the text message campaign raised expectations for the election outcome, which was then disappointed, in part due to a poor performance in how the ballot was organized.

An example of how institutions can be designed in specific response to challenges of diversity and low social capital is provided by Khwaja (2009), discussed in Section 2.1.1.3. The result shows the improvement of project performance in communities that lack social capital can be addressed by designing projects that face fewer appropriation risks, lower complexity, and more equitable distribution of project return and by investing in simpler and existing projects.

2.2.2.4. Social networks

Social networks may inform citizens' decisions to participate in the political or policy process. Network analysis can help shed light on how information or behaviors spread through groups of people. Understanding who the key individuals are, and the nature and magnitude of spillovers within networks, has implications for the effectiveness of policy interventions.

Recent research leverages social network analysis to further our understanding of governance issues. Gulzar, Khan, and Sonnet (2017) are analyzing whether the level of social connectedness of the organizer of women's political participation meetings affects the impact of the meetings on participation. Using a natural experiment in rural India, Prillaman (2017) finds that women's access to networks of other women can reduce the gender gap by enhancing their coordination and political participation. Lowe (2018)

evaluates how interventions to foster integration affects discrimination and the social networks of participants.

Social network analysis research going ahead will help us understand both how networks influence political participation, and how participation in turn affects networks. We see network analysis as a promising area for future research.

Open research questions:

- *How can institutional design reduce the impact of demographic identities?*
- *When and how does ethnic diversity worsen collective action and policymaking?*
- *How does institutional design of elections and community programs influence entry?*
- *How can more committed or qualified candidates be encouraged to stand for elections?*
- *How can people be encouraged to be more involved in community participation?*
- *What are the behavioral underpinnings of citizens' decision to participate?*
- *Do policy and public good preferences differ by gender?*
- *Through what channels do beliefs, institutions, or policies drive gender exclusion in the political and policy process?*
- *How does gender bias limit women's access to public office? What types of interventions impact bias?*
- *How does improving women's representation in politics affect service provision, citizen outcomes, and beliefs and aspirations regarding women's participation?*
- *How can social and other networks be best utilized to encourage participation in the political process?*

2.2.3. Information

2.2.3.1. The role of information

A direct implication of principal-agent models of politics is that limited information about how elections work, the roles and responsibilities of elected representatives, and politician characteristics may pose severe constraints on voter participation in low-income settings.

Poor citizens may be particularly susceptible to such information deficits. This, in turn, may lead them to place weight on more easily observed markers of a politician's policy preferences (e.g., their gender or ethnic identity). Parties and other political players, in turn, may exploit these informational asymmetries to underperform. This would suggest that policy interventions, such as audits and information campaigns, can increase electoral accountability. As reviewed in section 2.2.1, this view predicts that improvements in the information available to voters can directly reduce electoral malpractices. An emerging body of evidence shows that information can affect political outcomes. Voters are sensitive to the provision of information on candidates and policies relevant to them, and such information can affect voting behavior. Information is more likely to improve the selection of politicians when it is credible and directly links politician behavior to policy outcomes.

Providing information to citizens might be at the core of state-building. In a recent paper, Acemoglu et al. (2018) used two lab-in-the-field games to identify the source of the lack of trust in state institutions in rural Punjab, Pakistan, where reliance on non-state actors is common. They find that providing citizens with information about reduced delays in state courts leads them to report a higher willingness to use state

courts and to greater fund allocations in a dictator fund game. The information also had the indirect effect of reducing the stated confidence in non-state actors and hence reversing the feedback loop between state ineffectiveness and the legitimacy of non-state actors.

2.2.3.2.   *The role of media*

A first set of papers have examined the role of mass media in enabling citizens to monitor incumbent behavior and use this information in voting decisions. Brunetti and Weder (2003) examine whether cross-country variation in corruption is influenced by freedom of the press, as measured by an index provided by Freedom House. They find that an improvement of 46 points in the press freedom indicator (the move of the average country to full press freedom) is associated with a 1 point decrease in corruption (measured on a 0-6 scale with mean of 3.4).

Media studies from low-income countries also support this hypothesis. Using variation in government responsiveness across states and time—as measured by public food distribution and calamity relief to weather-related shocks and the degree of development of mass media—Besley and Burgess (2002) find that a 10 percent weather-related drop in food production leads to a 1 percent increase in public food distribution in states that are at the median newspaper circulation per capita, as well as a 2.28 percent increase in states that are in the 75th percentile. However, this literature is largely unable to distinguish the channels of influence.

A recent body of evidence shows the importance of media in mobilizing citizens. Aker, Collier, and Vicente (2016) randomize three voter information interventions during an election in Mozambique. One treatment sent information using SMS, the second allowed voters to report electoral misconduct via SMS, and the third consisted of regularly distributing a free newspaper. They find that all interventions boost turnout; however, only the free newspaper intervention statistically significantly increased the probability that voters took a costly action (sending an SMS) to signal their policy preferences and decreased independently observed electoral problems. Similarly, Larreguy, Marshall, and Snyder Jr. (2014) conduct a study that highlights the importance of local media—in their case local radio stations—for the dissemination of information. The authors analyze whether the release of municipal audit reports that document leakage in infrastructure projects for the poor reduce the incumbent's vote share. They find a positive answer when comparing areas covered by local radio stations to those without or with fewer radio stations.

Ravallion et al. (2015) offers a case where conveying information about rights to program beneficiaries in poor rural India changed their knowledge and perception about the scheme, yet did not affect outcomes. The intervention used an engaging movie that illustrated the rights under the antipoverty scheme.

A recent body of evidence based on randomized and natural experiments has made significant advance in disentangling channels of influence. These studies are aimed at answering two types of questions. First, how do voters react to information about politician performance? And second, does information about policy-making and the importance of voting influence voters' behavior?

2.2.3.3.   *Learning about politicians*

2.2.3.3.1.   *Learning through electoral campaigns*

TV campaign advertising is often used by politicians to reach and influence voters prior to an election. However, it is unclear whether and to what extent this advertising influences voting behavior and shapes election outcomes, due to the difficulty of capturing this effect while avoiding biases of reverse causality and omitted variables. Silveira and De Mello (2011) exploit a quasi-natural experiment provided by Brazilian electoral legislation to measure the impact of TV campaign advertising on election outcomes. Gubernatorial elections in Brazil work in a two-round system. If no candidate gets at least 50 percent of the vote in the first round, there is a runoff between the top two candidates in the second round. They find that changes in the TV advertising share between round 1 and round 2, which are mandated by the law, are responsible for closing 70 percent of the vote gap between the winner and the runner up in the second election in comparison to the first election. They find stronger effects in cities with poor education and high TV penetration, which adds credibility to the findings.

In places with low media penetration and low levels of education, in-person political campaigns become an important vehicle through which voters learn about politicians. Wantchekon (2003) randomized the form of political campaigns in eight non-competitive districts of Benin during Benin's national election, and examined whether voters favor clientelism per se, and whether clientelism increases ethnic voting. In each district, two villages were randomly chosen for treatment. In treatment villages, the incumbent and challengers were assigned different campaign messages, which could be either "public good" or "clientelist." While both campaigns focused on the same issues, the public good campaign framed the issue as part of a national program, while the clientelist campaign framed the issues as a specific project to transfer government resources to the region. The clientelist message also highlighted the candidate's ethnic affiliation. In control villages, candidates continued to use their typical campaign, which combined elements of the clientelist and public goods campaign. The results show a positive and significant clientelist treatment effect: a 15 percentage point average increase in vote share of candidates using clientelist speeches in the Southern region; up to a 17 percentage point increase for opposition candidates; and a negative and significant public policy effect on incumbent candidate's vote share—up to 24 percentage point lower vote share for public policy speeches in the Northern region of Benin. While the paper's main finding is that voters respond positively to clientelist campaign messages, it cannot say much about the type of signal related to politician quality that voters expected to extract from the treatment.

In a follow-up paper, Fujiwara and Wantchekon (2013) aim to answer whether informed public deliberation can overcome clientelism. Before the 2006 presidential election in Benin, candidates were randomly assigned to implement different campaign strategies in different villages. In the treatment villages, candidates held town hall meetings to discuss policy platforms and then opened the floor to voter deliberation. In control villages, candidates pursued their standard clientelist strategies. Surveys taken in the villages immediately after the election reveal that the treatment reduced self-reported measures of clientelism with no effect on voter turnout, suggesting that information campaigns and public deliberation can be just as effective in mobilizing turnout as clientelism but likely more cost effective. Interestingly, when candidates used the treatment campaign in their stronghold villages, their vote shares went down by 13.3 percentage points on average, but when candidates administered the treatment outside of their strongholds, their vote shares increased by 16.8 percentage points on average. Thus, the authors find that it is always in the best interest for at least one candidate to deviate from clientelism in favor of an information-based campaign strategy in a given voting area. This implies that optimal behavior by candidates should lead to a reduction in overall clientelism and an increase in information-based campaigns. Continuing this line of research, Wantchekon et al. (2015) ran an experiment around congressional elections in the Philippines. They find that town-hall style campaigns have a positive impact on small issue-based party's vote share, and that this effect is larger for women and poor voters (target demographics for the issue-based parties), whose attitudes are changed more by the town-hall meetings.

Political debates are often also considered an integral part of campaigns, but what is the impact of debates on voters and candidates? Bidwell, Casey, and Glennerster (2016) conducted a randomized evaluation to study the effect of publicly-screened debates during the run-up to Sierra Leone's 2012 parliamentary elections. The authors find that debates impacted voters' political knowledge and voting decisions. Exposure to the public debate screening improved alignment between the voters' reported policy positions and those of the candidates they voted for. The debate viewings also encouraged voters to support quality candidates; vote shares for the debate winner (as judged by the debate audience) rose by 5 percentage points at polling centers that screened the debate relative to centers that did not. Debates also caused candidates to invest more in their constituencies, both during the campaign and one year later. In the year after their election, members of Parliament who participated in the debates conducted an average of 4.2 community visits (compared to 2.9 in the comparison group) and held about twice as many public meetings. Similarly, Bowles and Larreguy (2018) evaluated a nationwide debate initiative in Liberia designed to solicit and rebroadcast policy promises from legislative candidates. By randomly encouraging debate participation more strongly in some districts than in others, they find that when leading candidates participated in debates, citizens were more engaged and informed and the turnout on election day was on average 1.5 percentage points higher relative to the comparison group.

These papers clearly demonstrate the malleability of voter behavior. However, they also underline real difficulties with interpreting the impact of treatments which seek to inform voters prospectively about leader behavior. In environments with limited policy commitment the credibility of campaign promises is likely low. Hence, while there is evidence that differences in political campaigns can influence voter behavior, it is unclear whether differences in political campaigns will have impacts on policy outcomes.

2.2.3.3.1. *Learning about past incumbent performance*

A second set of papers examining how voters use information on incumbent performance afford tighter tests of the information hypothesis. Such tests build on a key prediction of the political agency model—that voters should use retrospective voting to ensure electoral accountability. Several recent papers provide experimental evidence which supports or expands this prediction.

Ferraz and Finan (2008) use a natural experiment from Brazil in which the federal government began to select municipalities at random to audit their expenditures of federally-transferred funds, and compare the electoral outcomes of municipalities audited before (versus after) the 2004 elections, with the same levels of reported corruption. Audit findings were made publicly available and disseminated to media sources. Their results show that the re-election probability for incumbents who committed two corruption violations in municipalities with preelection audit was 7 percentage points (17 percent) lower than that of incumbents where audit findings were released after elections. The effects increase to almost 14 percentage points in municipalities with 3 violations associated with corruption. In all cases the comparison groups were municipalities where the audit reports were released after the election. These effects were more pronounced in municipalities where local radio was present to broadcast the information. Compared to municipalities audited after the elections, the audit policy decreased the likelihood of re-election by 11 percentage points among municipalities with one radio station and where two violations were reported.

Banerjee et al. (2010) also look at the role of disclosing information about incumbent performance in influencing a politician's re-election probability in the run-up to the Delhi municipal elections. They conduct a field experiment in conjunction with an NGO that used India's Right to Information Act to obtain disclosures on legislator spending, committee attendance, and participation in the legislature. The experiment provided slum dwellers in treatment polling stations with preelection report cards on incumbent performance and candidate qualifications. Dwellers in control polling stations did not receive any preelection cards. The intervention was associated with a 3.5 percentage point increase in turnout in

treatment slums. For the median performing incumbent, the campaign had no impact on the vote share. However, for the best performing incumbent the treatment increased vote share by 7 percentage points. Voters exhibited sophistication in interpreting qualification information. If the incumbent moves from being the best-qualified candidate (in terms of education, assets and criminality) to the second-best candidate, his vote share was reduced by 6.2 percentage points. This provides evidence for the effect of the information campaign in influencing voter behavior.

Another important question is whether information about performance interacts with ethnicity ties. Adida et al. (2016) show in the context of Benin that offering voters positive information leads them to update their prior beliefs only if the candidate is of the same ethnicity with the voter. They find symmetric results for negative information; that is, only non-co-ethnics are "punished" by updated negative beliefs.

At the same time, however, increasing information does not necessarily improve political accountability if voters react to the information by withdrawing from the political process altogether. Chong et al. (2013) conducted a randomized field experiment in Mexico in which a non-partisan corruption and public expenditure information campaign was carried out one week before the 2009 municipal elections. Their findings show that information about rampant corruption led to a decrease in voter turnout by 2 percent, a decrease in incumbent and challenger support, a decrease in voter identification with the corrupt incumbent's party by 40 percent, and a decrease in voter identification with any political party by 30 percent. They show that the main mechanism driving the results is one of voter disengagement rather than voters' strict party identification causing them to abstain from voting for a corrupt member of their own party rather than giving a vote to the opposite party. As the authors point out, the unanticipated findings of this research suggest that information may be necessary but is not sufficient for improving accountability, and more work must be done to ensure information leads to the desired outcome of enhanced democratic processes.

Humphreys and Weinstein (2012) conducted a field experiment in Uganda to test whether greater transparency influences voter response to information as well as politicians' behavior. They use random variation in exposure to scorecards produced by a local NGO. The scorecards provide information on initiatives undertaken by members of Uganda's parliament, and positions advocated by those members in plenary and committee sessions. Voters react to a politician's engagement in national politics and seem less influenced by ethnic ties or patronage politics. However, there is no evidence that politicians change their behavior as a response to exposure to scorecard information.

Another interesting dimension of voter responsiveness to politician performance is how voters react to events beyond politicians' control, such as weather-related catastrophes. Cole et al. (2012) study this phenomenon by using the quality of monsoon rains in India as an exogenous shock to welfare and looking at the decisions of voters in state elections. They find that, on average, voters punish ruling political parties but not all incumbent politicians for bad weather. If the incumbent politician belongs to the ruling party, his or her average vote share decreases by 3 percent for each standard deviation that district-level rainfall deviates from its optimum level. The results further indicate that incumbents do better when they respond to a crisis with emergency relief, but this only makes up for 1/7 of the loss in vote share for ruling during the crisis in the first place. More importantly, the evidence shows that voters only reward relief activities in the period leading up to an election. Governments respond to this recency bias and voter forgetfulness by concentrating relief activities during election years, resulting in suboptimal welfare outcomes for voters, especially when crises occur in non-election years.

2.2.3.3.2.   *Learning which individuals are good leaders: The role of quotas*

Voters may use performance of the incumbent to correct mistaken priors about the leadership quality of other individuals who share incumbent attributes. Changing stereotypes about groups is most likely to

occur when the leader belongs to a group which is underrepresented in politics, one such group being women. We have discussed Beaman et al. (2009) in Section 2.1.3.2, which shows how quota-induced representation influencing voter beliefs about the effectiveness of women leaders and, therefore, their future electoral behavior.

Chauchard (2014) studies how Schedule Caste (SC) quotas impact beliefs of dominant groups. The author finds that while quotas to not affect stereotypes, they do improve perceived norms of inter-caste interaction. As mentioned in section 2.1.1.2, Bhavnani (2016) finds that Scheduled Caste (SC) quotas do not have a lasting impact after they are removed on the chance that SC politicians win elections.

Bhalotra et al (2016) use a regression discontinuity design to study the impact of women narrowly winning against a man in state legislator elections in India. They find that a woman's victory makes her more likely to run again after her mandate ends. However, they also find a discouragement effect, in that other women are less likely to enter. This effect is concentrated in states with entrenched gender bias, suggesting a backlash effect against women leaders.

2.2.3.4.   *Policies and participation*

A second set of papers in the experimental literature on voter behavior has, in common with the US "Get Out the Vote" (GOTV) literature, focused on non-partisan information campaigns typically delivered by civil society organizations. Here again, a striking finding is the malleability of voter behavior. However, a key challenge faced by these papers is disentangling the role of motivation (or mobilization efforts) from learning in influencing voter behavior. We present these papers and how their experimental design seeks to disentangle channels of influence.

2.2.3.4.1.   *Lack of information as a barrier to political participation by women*

What constrains female political participation in low-income countries? One view is that low levels of education and limited female labor force participation make women disinterested in the political process. Another view is that patriarchal systems constrain the political freedom women can enjoy. Both suggest a limited role for informational campaigns in increasing female participation in politics.

Evidence from a field experiment in rural Pakistan that informed female voters about the importance of politics provides evidence to the contrary (Gine and Mansuri 2011). Randomization occurred at the level of geographic clusters within a village and within treated clusters a random sample of households were randomly assigned to receive the awareness campaign via a door to door campaign. There were two treatment arms: one which focused on the relationship between electoral process and policy and the second which included an additional module on the significance of a secret ballot. Turnout increased by 12 percent for women in treatment clusters with similar-sized effects in the two treatments. The ballot treatment, however, influenced women's ability to keep their voting decision secret—husbands in treatment areas were less able to predict their wives' voting decision.

2.2.3.4.2.   *Electoral malpractice as a barrier to voter participation*

One of the more glaring examples of electoral malpractice is electoral violence. Can information about the importance of elections and enabling voters to coordinate actions (which may have been hindered by limited communications between voters) influence voter behavior and the incidence of electoral violence? Collier and Vicente (2014) examine these issues in the context of the 2007 Nigerian election which was marred by significant violence—during the two days of the Nigerian election, over 300 citizens were killed. The field experiment featured a campaign against political violence conducted

by a major international NGO that specializes in community participatory development. The campaign caused voters to report a 10 percent improvement in political freedom and an 11 percent decline in conflict within the local community. These improvements in citizen perception were accompanied by large but noisily estimated improvements in citizen willingness to directly take costly actions in the form of visiting the post office to return postcards demanding more attention to countering voter intimidation.

The intervention also increased turnout by 7 percent, and in the presidential election this was associated with a decline in the vote share of the opposition candidate. The fact that the campaign had the effect of decreasing the vote share of the candidate associated with violence while the vote shares of the other two main presidential candidates were not significantly changed suggests that the campaign succeeded in transmitting information to voters about candidate characteristics.

2.2.3.4.3. *Changing voter preferences about politician quality through non-partisan campaigns*

Banerjee et al. (2010) worked with an NGO in the North Indian state of Uttar Pradesh to evaluate the impact of two non-partisan campaigns. Treatment villages for the two campaigns were drawn from the same set of jurisdictions. The first campaign encouraged villagers not to vote along caste lines and instead to vote for the candidate who is best able to deliver development. The second campaign encouraged villagers not to vote for corrupt campaigns and again exhorted villagers to vote for the candidate best able to deliver development. The campaign message was delivered in treatment villages by a NGO team using multiple methods, a prominent one being a puppet show. The caste campaign increased turnout by 7.6 percent and the survey data shows a 9 percentage point decline in voting for the caste-preferred party. The paper also finds that the caste treatment reduced the incumbent vote share by 35 percentage points when the incumbent faced criminal charges, showing that declines in caste-based voting are related to the quality of politicians, since they cause voters to re-evaluate the trade-off between ethnic voting and criminality.

In contrast, the corruption campaign largely left voter behavior unaffected. One interpretation of the differing results of the two campaigns is voters' ability to interpret the campaign message. In an environment where corruption is the norm, voters are unable to translate broad messages about fighting corruption into concrete action regarding whom to vote for. This interpretation is consistent with the earlier findings that voters react to specific information about incumbent's corruption records.

Finally, providing voters with information about incumbent performance and candidate qualifications may also lead to a reduction in electoral malpractices. As discussed in section 2.2.3.3.2. Banerjee et al. (2011) provided slum dwellers in treatment polling stations with preelection report cards on incumbent performance and candidate qualifications. As a result, slums in the treatment area witnessed a 19 percentage point decline in vote buying. However, identifying whether information disclosure led to a decline in vote-buying is difficult, since voters may have responded to better information by increasing the price at which they sell their vote. Alternatively, parties may be more wary of vote-buying in areas that had an NGO campaign.

2.2.3.4.4. *Monetary incentives and voter participation*

Leon (2013) combines a field experiment with a change in Peruvian voting laws reducing the fine for abstention to study how monetary incentives affect voting and the composition of the electorate. Until 2006, the fine for not voting in Peru was around USD$50, but after 2006 the fine was reduced to either USD$25, USD$12.5, or USD$6 depending on how relatively poor a district was. Despite the substantial reduction in the fine, however, no news outlets covered the reduction nor were there information campaigns about it, meaning that very few people knew about the new law at the time the 2010 municipal

election. Noting this lack of information, Leon conducted an experiment in which voters in a randomly selected treatment group were given information about the new fine levels just before the 2010 election, whereas in the control group voters were reminded about the fine but not explicitly told the new levels. He finds that the elasticity of voting with respect to cost is -0.21. Extrapolating, he suggests that if the fine was reduced to zero, turnout would decrease from 94.7 percent to 74 percent. Further analysis shows that the reduction in turnout is driven by centrist voters, those who are less politically informed, and those with less interest in politics. Therefore, the change in turnout did not lead to an average change in the policy preferences of the electorate.

2.2.3.4.5. *Vote-buying and participation*

Information can also be used to discourage voters from engaging in activities such as vote-buying by politicians. Vicente (2013) conducted a randomized experiment in Sao Tome and Principe just before the presidential election of July 2006 to test whether vote-buying actually affects voting decisions. The treatment was a leaflet door-to-door campaign against vote-buying sponsored by the country's Electoral Commission. He found that vote buying increases turnout of the electorate, but there was no change in turnout due to the campaign. However, the campaign not only diminished the frequency of vote buying but also its effectiveness. In the treatment group, the probability of an individual voting for the incumbent went up by 40 percent in comparison to preelection intentions. Therefore, the challenger in the election, who engages in vote-buying later (closer to the election) than the incumbent, was more affected by the intervention. Overall, his results show that vote buying can change voting outcomes by changing the preferences of the electorate in favor of the challenger who buys votes later in the game. When informed not to engage in this exchange, voters instead voted for the incumbent.

Hicken et al. (2015) study whether inviting voters to make a pledge or promise not to sell their vote affects actual vote-buying. They find that this treatment indeed reduces vote-selling, more so in areas where vote buying payments are smaller. These results are consistent with a model where vote-selling is a temptation good, and partly sophisticated voters can reduce this behavior through promises that impose a utility cost when they are broken.

2.2.3.4.6. *Property rights and participation*

A large institutions literature shows significantly greater expropriation risk in low-income countries, importantly driven in part by colonial histories. The basic argument is that colonial powers set up institutions that help them extract resources and these institutions form the backbone of current institutions in these countries (Sokoloff and Engerman 2002; Acemoglu et al. 2001; and Acemoglu and Johnson 2005).

Using a natural experiment, Janvry et al. (2014) study how strengthening property rights over agricultural land in Mexico impacted voting behavior. The 1992 Mexican land reform was rolled out across the country over a period of 14 years, providing certificates of ownership to individuals who previously had access to land but faced limitations in official property rights. The authors match affected communities before and after the change in property rights with voting outcomes in corresponding electoral sections over six federal election cycles. They find that after the receiving complete property rights, voters exhibit a conservative shift, increasing their support for the pro-market party by 6.8 percent of its average vote share. This shift to the right is stronger in regions with more valuable land. These results support the investor class theory, which argues that individuals with productive assets, such as land, prefer pro-market politicians who will support low taxes on capital and labor income, allowing for greater benefit to be derived from the assets. Interestingly, voters did not respond to the increase in property rights through reciprocal support for the pro-state party that implemented the reforms. Thus, the authors point out that redistributive land policies, which are often championed by pro-state parties,

actually may end up being politically costly for such parties.

Open research questions:

- *What are the best ways to inform voters to elicit a change in their behavior or attitudes towards corruption, gender, ethnicity, or politician underperformance?*
- *Does the form of information provision influence voter responsiveness?*
- *How do political parties and other political actors respond to greater information disclosure?*
- *How should disclosure laws be structured? Is there an anti-incumbency bias built into such laws?*

2.2.4. Government programs

Some theories assert that voters are retrospective, voting according to whether or not they are satisfied with the incumbent's performance in office. If this is the case, government programs may increase support for the implementing administration among the program's beneficiaries. This may be especially true for cash transfers, which can deliver a substantial increase in income to poor families.

Manacorda et al. (2011) test this theory in Uruguay, where PANES (*Plan de Atención Nacional a la Emergencia*) delivered monthly cash transfers and food cards to qualifying households between 2005 and 2007. Eligibility for the program was determined through a predicted income score using household characteristics, and researchers used this discontinuous rule for program assignment to compare households just above and just below the cutoff. They found that program eligibility was associated with a 13 percentage point increase in support for the current government and that the positive effect continued after the program ended.

Labonne (2013) examines the effect of cash transfers on local elections. The study used the phased roll out of a conditional cash transfer program in 17 municipalities in the Philippines. Municipalities were randomly assigned to either have the program implemented in all villages before the 2010 election or have the program implemented in only half of the municipalities' villages. The incumbent mayor's vote share was 26 percentage points higher in municipalities where all villages received the program. In this case, the municipal mayor had no influence on where the program was implemented or who could receive benefits, but the transfer program still improved the electoral outcomes of the incumbent.

Open research questions:

- *Do governments strategically target programs to increase their chances of reelection?*

2.2.5. Technology

2.2.5.1.   Technologies for monitoring votes

Many LMICs are somewhat democratic, electing their governments through regular elections. However, many of these elections are rife with manipulation and fraud: some people are prevented from voting, others are paid to vote multiple times, and officials reallocate votes in exchange for kickbacks. In countries with little government accountability, new technologies for monitoring votes may help to create freer and fairer elections.

In the 2010 election in Afghanistan, Callen and Long (2015) test the effectiveness of Photo Quick Count in Afghanistan, a new camera monitoring technology. Researchers took photographs of the

declaration of results forms posted at each polling center and compared them with photographs of the forms submitted to the national count. Although this form of monitoring only captures manipulation by elections officials that occurs after the initial ballot count, the program reduces the submission of fraudulent results forms by 11.8 percentage points.

Although monitoring technologies can be effective, corrupt agents often adapt to monitoring efforts, especially if they become predictable. Callen and Long find some evidence that when the photo monitoring made manipulation of results forms riskier, candidates moved the manipulation of forms to (supposedly) unmonitored polling stations or had officials reallocate cast votes from one candidate to another. There is also evidence that fraud levels influence voters' attitudes toward their government. Berman et al. (2014) find that in areas in Afghanistan with the Photo Quick Count technology described above, where electoral misconduct was reduced, voters reported more positive attitudes toward and support for the government.

Ichino and Schundeln (2010) conducted a randomized evaluation in Ghana where they sought to understand how monitoring spilled over to nearby areas. During the registration period before the 2008 election, they randomly assigned constituencies to a treatment or comparison group. Monitors then visited 25 percent of registration stations within the treatment constituencies. Within treatment constituencies, growth in voter registration declined by 3.5 percentage points. However, some of the deterred registrations appear to have been displaced to nearby electoral areas, where voter registration increased by 2.8 percentage points.

More connected politicians intuitively benefit more from electoral fraud. Moreover, it is possible that they can more easily evade monitoring efforts due to their connections. Callen and Long (2014) show that more connected local politicians benefit disproportionately from vote aggregation fraud. However, they also find that the above described Photo Quick Count technology reduced their vote counts by 25 percent.

A separate question relates to the distributional consequences of information technology, specifically if the poor will benefit more or less in relative terms. Grossman, Humphreys, and Sacramone-Lutz (2014) ran an experiment in Uganda where they offered voters the chance to send messages to their representatives. They find that marginalized populations are actually more likely to use this method compared to alternative communication channels, compared to other voters, which assuages concerns that ICT widens inequality in representation.

2.2.5.2. *Technologies for voting*

Another obstacle to fully democratic elections may be that the voting process is too complicated for individuals with little education. If individuals fill out their ballots incorrectly, their votes are not counted. As a result, uneducated (and often poor) individuals have little influence on policymaking.

Fujiwara (2015) tests the effect of electronic voting on participation and health spending using the phase-in of electronic voting machines in Brazil. The machines, which displayed photos of candidates and alerted voters when they were about to cast an invalid ballot, were used in the 1998 election in municipalities with more than 40,500 registered voters. Fujiwara used this discontinuity to show that electronic voting increased the share of valid votes by 12 percentage points, effectively enfranchising many poor and illiterate voters. Furthermore, this change had an influence on policymaking: spending on healthcare, an issue especially important for the poor, increased.

2.2.5.3.   Technologies for citizen participation

Leveraging technologies and especially information and communications technology to foster direct citizen participation at scale is a new avenue for research. In Pakistan, Golden et al. (2017) test how efficient an Interactive Voice Response (IVR) system is at achieving this objective. Of those called by the IVR recorded by a politician asking specific questions, 31 percent responded to at least one question. Further, they find that a responsive follow-up from the politician is effective at improving accountability and ultimately contributes to more effective engagement of citizens in the political process. Similarly, George et al. (2018) test the potential of using voice and text messages in India to inform citizens of the criminal charges held against candidates in their constituency. They find that votes for candidates with severe charges drop by 7.7 percent and votes for candidates with no charges increase by 6.7 percent while increasing the overall turnout by 1.6 percent.

Open research questions:

- *What strategies can help to prevent corrupt agents from adapting to monitoring technologies or tools?*
- *What is the influence of technology and social media on electoral outcomes?*

3. Impacts and determinants of leakages

Corruption is widely understood as "the misuse of public office for private gain." For our purposes, we use the word corruption as "incidents where a bureaucrat (or an elected official) breaks a rule for private gain." (Banerjee et al. 2012b). This broader definition would include the most obvious type of corruption—a bureaucrat taking a bribe in order to bend or break a rule—but also encompasses more nuanced forms of bureaucratic corruption, including nepotism to favor a family member rather than issue a competitive bid. We use the word "leakage" to include not only this definition of corruption, but also other acts that are sometimes not seen as corruption but nevertheless lead to a loss of public funds or create inefficiencies in the delivery of public goods and services. Examples include unauthorized absenteeism while still collecting a paycheck, needless delays in providing services, and lack of effort towards final outcomes.

From an economic perspective, there are two basic issues we would like to address about the impacts of leakages: magnitude and efficiency costs. The first, which we examine in Section 3.1, is simply put: how large are leakages? Although anecdotal evidence suggests that leakages are common in LMICs and more prevalent in LMICs than in rich ones, there are remarkably few reliable estimates of the actual magnitude of leakages and corruption, and the credible estimates we review in Section 3.1.7 reveal a surprisingly high level of heterogeneity.

However, just knowing the magnitude does not tell us much about how serious the problem is from an economic perspective. After all, it is at least theoretically possible that corruption could just represent a transfer from one party (say, the government) to another party (say, bureaucrats), with little efficiency cost. In fact, if bureaucrats' official salaries were less than their market wage in expectation of the corrupt rents they would obtain—and there is evidence that this is indeed exactly what happens—there could be no net costs of corruption at all. In practice, however, the evidence we review in Section 3.2 suggests that the efficiency costs of corruption can be quite severe, as corruption may raise the marginal tax rate of firms, decrease business activity, raise the marginal costs of public funds, make certain government

projects economically unviable, and undo the government's ability to correct externalities, leading to inefficient outcomes.

Finally, understanding the determinants of leakages is essential to devising policies to reduce corruption. Key determinants include civil servant incentives and the structure of the bureaucracy, external factors (technology, transparency, judicial corruption and demographics), and long-term anticorruption policy.

3.1. How large are leakages, really?

3.1.1. Perceptions

The early economic literature estimated corruption through surveys of perception. These perception surveys have the advantage of good coverage—it is much easier to ask someone's perceptions of corruption than to actually measure corruption directly. As such, they still form the basis of most cross-country corruption indices, such as Transparency International's Annual Corruption Perception Index (CPI) and the World Bank's Control of Corruption Index—although the latter incorporates very different aspects of corruption, ranging from the frequency with which firms make "additional payments to get things done," to the effects of corruption on the business environment, and finally to measuring "grand corruption" in the political arena, or the tendency for people to obtain positions of power based on patronage rather than level of ability. Perception-based measures were also used in some of the first empirical work in economics on corruption, such as Mauro's (1995) well known study of the relationship between corruption and growth, discussed in more detail below.

The challenge with perception-based measures is that they may not measure corruption accurately. Olken (2009) examines the reliability of corruption perceptions in a micro-environment: villagers' perceptions of the level of corruption in a local road building project. Olken asked villagers to assess the likelihood of corruption in the road project on a scale from 1 to 5. At the same time, he developed a much more detailed measure of the amount of corruption that was actually present in the road project by comparing the amount the village government spent on the road to the amount independent engineers estimated the road would actually cost to build (based on the amount of materials used, the local prevailing prices for materials, and so on). While villagers' perceptions do reflect actual corruption in the road project, the magnitude is quite weak: increasing the actual missing expenditures in the road project by 10 percent increases the probability a villager reports any corruption in the road project by just 0.8 percent.

Moreover, villagers' perceptions appear to be biased in two ways. First, villagers are much better at detecting marked up prices (i.e., overcharging for cement) than inflated quantities (i.e., billing for 1,000 m^3 of rocks but only delivering 800 m^3)—and given this, it is not surprising that most of the corruption occurs by inflating quantities. This may account for the relatively low correlation between perceptions and actual corruption, since people must make an inference about the aspects of corruption they cannot perceive—which end up being where the bulk of corruption is usually hidden. Second, Olken shows that individual characteristics, such as one's education, have much more predictive power than actual corruption itself. If a perception survey has different compositions of people evaluating different projects (or countries), this could create systematic biases in the use of perception.

These types of biases could create problems in macro-level perception indices as well. For example, after the fall of Suharto in 1998, many commentators perceived that corruption in Indonesia became worse. For example, Kuncoro (2006) argued that decentralization in the post-Suharto era replaced "centralized corruption—one-stop shopping […] with a more fragmented bribe collection system." Many

players at both the local and the national level started demanding bribes, and their failure to coordinate their bribe-taking behavior resulted in a higher total level of bribes. The worsening of perceptions of corruption was captured by the Transparency International Index—measured on a scale from 0 (highly corrupt) to 10 (highly clean)—which fell from a value of 2.0 in 1998 to 1.7 in 1999, and stayed at the same level in 2000. This may well have been the case, but another explanation is that the fall of Suharto's dictatorship resulted in a free press. Unlike the press under the Suharto regime, which was tightly restricted, the free press was newly able to report on allegations of corruption, which it did. It is therefore possible that perceptions of corruption rose even though actual corruption fell. For these types of reasons, economists have been moving to more direct measures of corruption whenever possible. Developing more direct measures of corruption not subject to these types of biases that are applicable across countries and contexts remains an important topic for future research.

3.1.2. Estimating bribes: Survey evidence

Perhaps the most direct way of measuring bribery is through the use of surveys of bribe-payers. In most contexts, there is relatively little stigma associated with paying bribes, and so in many cases bribery can be measured using surveys of firms or households. One early example of this is Svensson (2003), who surveyed firms in Uganda and examined how much they paid in bribes. On average, firms in the survey report bribe payments of about 88 USD$per worker, or about 8 percent of their total costs.

Since this type of survey-based measure of bribes is the most easily replicable, it is one of the only areas where consistent measurement is now being carried out across countries and over time. One key dataset is the International Crime Victim Surveys (ICVS) from 49 countries, in which individuals are asked whether any government official in that country has asked them or expected them to pay a bribe for his or her services during the previous year. Using this data, Mocan (2008) finds that income and education of the individual have positive impacts on the likelihood of being asked for a bribe in LMICs. The impact of these variables is not statistically significant for high-income nations. Hunt (2004) uses the 1999 and 2000 data on bribes paid by individuals to public officials in 34 countries from the same source to examine the implications of trust networks. The World Bank Enterprise Surveys (WBES)[2] have asked comparable questions about firms' informal gifts or payments in obtaining water, electricity, telephone connection, operating and import licenses, construction-related contracts, meetings with tax officials, government contracts, and more generally "getting things done." This data is available in the main Enterprise Survey, as well as in various types of surveys conducted under the umbrella of WBES such as the Indicator Survey and Informal Survey, and covers most LMICs.

As this type of data becomes more available we will be able to produce more reliable estimates of bribery over time, across countries and within countries. Bai et al (2017), for example, use this type of direct data on bribe-paying from repeated cross-sectional surveys of firms in Vietnam to estimate the impact of industry growth rates on bribe-paying behavior, finding that the share of revenues paid as bribes decreases as industries grow.

3.1.3. Estimates from direct observation

The best way to measure corruption is often to observe it directly. Needless to say, this is difficult, since corrupt officials rarely will allow corrupt behavior to be observed. Nevertheless, there are several notable examples of direct observation of corrupt activity. Perhaps the most famous example is the case of Montesinos in Peru, documented by McMillan and Zoido (2004). Montesinos, who was the secret-

---

[2] See https://www.enterprisesurveys.org for exact details on the number of countries and years available for each type of survey.

police chief under President Alberto Fujimori in Peru, bribed judges, politicians, and the news media to support the Fujimori regime. Remarkably, he kept detailed records, with signed contracts from those he bribed and videotapes of them accepting the bribes. After the fall of the Fujimori regime the contracts and videotapes became public, and McMillan and Zoido use them to estimate the cost of bribing various types of government officials. On average, they show that politicians received bribes ranging from 3,000 - 50,000 USD$per month, depending on whether the politician was in the opposition party (higher) or Fujimori's party (lower), with judges receiving bribes of the same order of magnitude. The bribes to control the media were orders of magnitude larger—as much as USD$1.5 million per month for one television station's support.

Olken and Barron (2009) provide direct data on actual bribes in a more prosaic setting: the bribes truck drivers pay to police on their routes to and from the Indonesian province of Aceh. Over a nine-month period, enumerators accompanied truck drivers on their regular routes, dressed as truck drivers' assistants, and simply noted the amounts that truck drivers paid each time they were stopped at a police checkpoint or weigh station. On over 300 trips, they observed more than 6,000 illegal payments. Usually the amount of each payment was small — averaging 0.50 to 1 USD$, sometimes in cash and sometimes in kind (such as a pack or two of cigarettes). In total, the illegal payments represented 13 percent of the marginal cost of the trip. By comparison, the salary of the truck driver was only 10 percent of the marginal cost of the trip.

Sequeira and Djankov (2014) use a similar methodology in Mozambique and South Africa, shadowing clearing agents who process customs for cargo as it passes through the ports. They find that, on average, bribes represent 14 percent of the total shipping costs for a standard container passing through the port of Maputo, Mozambique, and 4 percent of total shipping costs for a standard container passing through Durban, South Africa.

To study bed net distribution programs in Ghana, Kenya, and Uganda, Dizon-Ross, Dupas, and Robinson (2015), perform an audit study to estimate corruption by health workers when distributing bednets to the poor. They hired enumerators to present themselves as ineligible "mystery clients" to procure subsidized bednets. In sharp relief to many other studies on bribery, only 4.7 percent of those hired were successful, despite being incentivized to try to obtain a bednet. Overall, a survey of those eligible revealed that only 1.4 percent were asked to pay a bribe.

3.1.4. Estimating graft: Estimation by subtraction

The most common method for estimating graft is what we term estimation by subtraction. In this method, one obtains two measures of the same quantity, one measure before corruption takes place and one measure after corruption takes place. The estimate of corruption is the difference between the two measures.

One of the first estimates using this technique is the pioneering study by Reinikka and Svennson (2004). In their paper, they use what they term a Public Expenditure Tracking Survey (PETS) to track educational funds in Uganda. Specifically, they compare the amount of a special education block grant sent down from the central government and compare it to the amount of the block grant received by schools. They estimate a leakage rate of 87 percent. Once the results were publicized, a subsequent study found that the leakage rate fell to less than 20 percent. An important question in such an approach is the quality of recordkeeping: if schools have poor records, some of the money might not show up on the books even though it may have been received. Studying the importance of recordkeeping quality in PETS is an important issue for the replicability of this technique. Subsequent to this work, similar PETS studies have been carried out, largely by the World Bank, in a variety of contexts.

Another recent study in Indonesia estimates the extent of missing funds in an anti-poverty program called *Inpres Desa Tertinggal* (IDT, presidential aid for poor villages). Suryadarma and Yamauchi (2013) compare the amount of disbursed funds in a given district and year to the amount actually received by households of IDT villages in these districts to find that only 69 percent of disbursed funds on average reach the intended beneficiaries. They also find that districts which had a high initial human capital index and a greater number of female village heads received a greater proportion of the disbursed funds and targeted these funds more accurately to the poorest 20 percent of households in the villages. Therefore, their evidence suggests that local characteristics can influence the level of leakage that occurs.

Using a similar approach, Fisman and Wei (2004) measured tax evasion by comparing Hong Kong's reported exports and China's reported imports of the same products. The strategy allowed them to differentiate three different aspects of tax evasion: underreporting of unit value, underreporting of taxable quantities, and mislabeling of higher-taxed products as lower-taxed products. These calculations are then used to estimate the effect of tax rates in tax evasion. They found that higher-taxed products were associated with a forty percent higher median evasion rate.

Olken (2007) implements a related exercise in the case of rural road projects. Instead of using an expenditure tracking survey to estimate the amount received, he compares the official amount spent on the road to an independent engineering estimate of what the road actually costs to build, where engineers dug core samples of the roads to estimate materials quantities, conducted price surveys to estimate local prices, and interviewed villagers to estimate actual wages paid. Importantly, since some amount of materials naturally disappears during construction, Olken built several small "test roads" to calibrate the metric so that it would show zero corruption when, in fact, corruption was zero. Using this approach, Olken estimated that "missing expenditures"—the difference between what the village claimed the road cost and what the engineers estimated it actually costs—averaged about 24 percent of the total cost of the road.

Another way of implementing this technique is to compare administrative data to a generally administered household survey. Olken (2006) uses this approach to estimate theft of rice from a subsidized rice distribution program in Indonesia. Using this approach, he estimates that on average, at least 18 percent of the rice cannot be accounted for, with greater amounts in ethnically heterogeneous and sparsely populated areas.

Employing a similar approach, Duflo et al. (2013) carry out a field experiment involving industrial plants in Gujarat, India in order to measure the extent of underreporting of pollution by third party auditors. In their control group, plants continue to follow the status quo of hiring and paying their own auditors whereas in the treatment group plants are assigned auditors who are paid a fixed wage from a central pool and are subject to random performance checks. By comparing pollution levels reported by third party auditors to those reported by an independent technical agency, the authors find that 29 percent of audit reports from the control group falsely state that a plant's pollution level is just below the regulatory standard. Therefore, in the status quo, auditors maximize on their business incentive to underreport pollution in order to continue receiving the auditing contract and payment from the industrial plant.

Dizon-Ross, Dupas, and Robinson (2015) perform a similar exercise to estimate corruption at the lowest level in the distribution chain of bednets and find that leakages are small. In Kenya, while official estimates implied a maximum potential coverage of 98 percent of households, survey evidence seemed to suggest that about 91 percent of households received bednets. In Uganda, the difference was negative, implying that more bednets were distributed than allocated. The authors argue that this was because the districts they surveyed received a higher number of bednets per pregnant woman on average.

Niehaus and Sukhtankar (2013b) use the subtraction method to estimate the marginal rate of corruption in India's National Rural Employment Guarantee Scheme (NREGS). Comparing official records and survey responses, they find that the initial average leakage in wage payments was close to zero. However, a government mandated increase in wages under the scheme resulted in no increase in reported wages. In other words, the *marginal* rate of corruption with respect to an increase in wages was close to 100 percent. The paper thus highlights the difference between the average level of corruption and how corruption responds to policy changes.

An important component of corruption that is often hard to measure is personal, non-wage asset accumulation by politicians while in office, which is exactly what Fisman et al. (2014) set out to measure in their innovative study in India. Since November 2003, all political candidates in a number of Indian states have been required to report the value and composition of their assets when standing for office. Taking advantage of this policy, Fisman et al. gathered data on the assets of all winners and runner ups in an election in a given year who stood for office again in the following election. Therefore, they were able to compare the growth in assets between the two elections of those who were in office to those who served as a good counterfactual—the runner ups that almost made it to public office. They find that annual asset growth of the election winners is between 3 to 6 percent higher, on average, than the runner ups. However, they find evidence that the results are largely driven by those who hold high positions in the government, with the asset growth of Parliamentarians in the Council of Ministers being 13 to 29 percent greater than that of runner ups. Furthermore, the effect was minimal for first time politicians compared to an estimated incumbent premium of 12.6 percent in asset growth, suggesting that career politicians have low earning potential in their outside options.

When examining corruption through price manipulations, one can compare an official price to the market price and use the difference as a measure of price manipulation. Hsieh and Moretti (2006) do this for a very famous case: corruption under the Iraqi Oil-For-Food program administered by the United Nations. Specifically, they compare the price received by Iraq for its oil to the going price for comparable oil on the world spot market and use a model of the market for oil trading to infer what share of that under-pricing was likely received by Saddam Hussein's regime. While the total amount of corruption they estimate is enormous—approximately USD$1.3 billion—it amounts to only about 2 percent of the total volume of oil sold.

Of course, not all price markups are corruption—they could simply reflect incompetence in obtaining good prices for the government, or a lack of incentives in the government sector to obtain the best prices. In the Italian case, Bandiera et al. (2009) compare corruption with incompetence by examining the decision to switch to procurement through a centralized price agency rather than procuring by oneself. If price markups were due to corruption, those agencies paying the highest prices (the most corrupt) would be less likely to switch to the centralized procurement regime; if they are due to incompetence ("passive waste" in their terminology), those paying the high prices would be more likely to adopt centralized procurement. On average, they estimate that 83 percent of the above-average prices in the Italian data they examine are due to "passive waste," and only 17 percent due to corruption. Of course, it is possible that in many other countries, the share due to corruption is much higher.

3.1.5. *Estimating other types of leakage from surveys*

Shirking by government employees can also be considered an important type of leakage. To examine this, Chaudhury et al. (2006) conducted an innovative survey in 6 countries—Bangladesh, Ecuador, India, Indonesia, Peru, and Uganda—to examine the degree to which public sector teachers and health workers were not performing their job. To conduct the survey, surveyors showed up unannounced at government schools and health clinics during hours they were supposed to be open and simply counted attendance

rates. Averaging across the countries in their sample, about 19 percent of teachers and 35 percent of health workers were absent. For primary school teachers, this ranged from a low of 11 percent absence in Peru to 27 percent absence in Uganda; for primary health centers, it ranged from 25 percent in Peru to 40 percent absence in India and Indonesia. The authors find lower rates of absence in richer countries and, looking within India, in richer Indian states—OLS regressions of absence on log of per capita GDP in the sample of countries and in Indian states surveyed revealed that a doubling in income is associated with a 5.8 percentage point and 4.8 percentage- point drop in absence rates respectively.

Although many would not consider this a form of corruption, it can be thought of as a type of leakage, as these public servants were not working a substantial fraction of the time for which they were paid. Banerjee et al. (2004) conducted a similar exercise for health workers in the state of Rajasthan and found comparable absence rates of 45 percent for medical personnel in sub-centers and aid posts, and 36 percent in the (larger) primary health centers and community health centers. A larger, nationally representative survey of over 1400 public health centers in India by Chaudhary et al (2011) reinforces these findings: nearly 40 percent of doctors and medical service providers were absent from work on any given day. Doctors who had longer commutes, who worked in poorer and more remote areas, were more likely to be absent. The absence problem was quite widely distributed and was not concentrated amongst a few doctors. In a snapshot of teacher absences across Indian states, Kremer et al. (2005) find that 25 percent of teachers were absent from school, and only about half were teaching during unannounced visits to a nationally representative sample of government primary schools in India. Higher absence rates were concentrated in the poorer states. Muralidharan et al (2016) revisit the same schools seven years later and only find a modest improvement in teacher absence, despite significant improvements in schooling inputs. The panel nature of their dataset allows them to estimate the costs of teacher absence, which they estimate at approximately USD$1.5 billion per year.

Another way in which people have measured leakage is by looking at who receives a particular program, and measuring whether certain types of individuals are more likely to receive benefits than they should, based on official program rules. For example, Alatas et al (2013) study Indonesia's largest government programs and a separate high-stakes cash transfer scheme and test for elite capture and consequent leakages at two different stages—the designing of the list of beneficiaries and the actual receipt of the subsidies during implementation. The authors find that relatives of those holding official village positions (so-called "formal elites") are more likely to have actually received targeted benefits during implementation in certain cases. The estimated difference in magnitude is as high as 8 percentage points (19 percent) for the temporary cash assistance programs in 2005 and 2008 and the health insurance program.

3.1.6. Estimates from market inference

Finally, in some cases one can use the theory of market equilibrium, combined with data on market activity, to estimate the amount of corruption. In a pioneering study, Fisman (2001) applied this approach to estimate the value of political connections to Indonesian president Suharto. Specifically, he obtained an estimate from a Jakarta consulting firm of how much each publicly traded firm was "connected" to Suharto, on a scale of 0-4. He then estimated how much each firm's price moved when Suharto fell ill to estimate how much the stock market assessed the value of those political connections. If the efficient markets hypothesis holds, then the change in stock market value surrounding these events captures the value of the political connection to the firm. Since investment bankers in Jakarta estimated that the total market would fall by 20 percent if Suharto died, he can calibrate these estimates to estimate the total "value" of the connections to Suharto. On net, for the most connected firms he estimates that about 23 percent of their value was due to Suharto's connections.

Faccio (2006) pursues a similar approach using a large sample of countries—she examines political connections to 20,202 publicly traded firms in 47 countries. For each of these firms, she defines the firm as having a political connection if a board member or large shareholder is a politician (e.g., Member of Parliament or minister). She focuses on corporations where a previous board member and large shareholder becomes a politician. She finds that, on average, having a member of your board or large shareholder become a politician is associated with a 2.29 percent increase in the company's share value. Echoing the contrast between Suharto in Indonesia and Cheney in the United States, when she splits the sample into countries with below and above average corruption levels (as measured by the World Bank perceptions index), she finds that the impact comes entirely from high corruption countries: in above median corruption countries, having a board member or large shareholder become a politician increases stock market value by 4.32 percent, but in below median corruption countries, having a board member or large shareholder become a politician has no impact on stock value.[3]

Another approach to measuring corruption uses equilibrium conditions in the labor market. Specifically, one can use the hypothesis that people in the public sector must, on the margin, be indifferent between their public sector job and alternative jobs in the private sector (differences between public and private sector employees will be developed in section 4.1.1). If, controlling for their job opportunities, pay is lower in the public sector, the result could simply reflect a compensating wage differential. But if pay in the public sector is lower but consumption levels are the same, one could infer that the difference between pay and consumption in the public sector relative to the difference between pay and consumption in the private sector tells us something about how much those in the public sector are likely receiving in the form of bribes. Gorodnichenko and Peter (2007) perform this exercise using a household survey in Ukraine. They find that, controlling for education, hours of work, job security, fringe benefits, job satisfaction, and secondary employment, public sector workers received 24-32 percent less income than their private sector counterparts. Crucially, however, they have the same level of consumption and assets, suggesting that a large part of the gap must be made up in bribes. Aggregating across the economy they estimate that the total amount of the gap (and hence bribery) is between USD$460 million to USD$580 million, or about 1 percent of GDP.

3.1.7. So how much corruption is there, really?

The magnitudes of corruption raise several important questions. First, a striking correlation that comes up in a variety of datasets—from the perception indices to the absence study of Kremer et al (2005) to the Faccio (2006) and Fisman (2001) studies of the value of political connections to the Sequeira and Djankov (2014) comparison between ports in South Africa and Mozambique—is the strong negative relationship between income and corruption: as best we can measure it, richer countries in general appear less corrupt. The causality likely runs in both directions. It is easy to see how low corruption could cause countries to become rich if corruption hinders economic activity (Mauro 1995).

However, the relationship in the other direction—that richer countries become less corrupt—is less well understood. Treisman (2000) argues using geographic instruments for per-capita income that, at least in part, exogenously richer countries tend to be less corrupt, but understanding the mechanisms behind this relationship is an important open question.

---

[3] Not all studies find zero effect of connections in wealthy countries. Faccio (2006), for example, finds that stock price of firms headquartered in the home town of politicians who died suddenly fell on average by -1.7 percent around the death of the politician. Jayachandran (2006), who uses the surprise event of Senator Jeffords leaving the Republican Party and tipping control of the U.S. Senate to the Democrats to show that a firm lost 0.8 percent of market capitalization the week of Jeffords' switch for every $250,000 it gave to the Republicans in the previous election cycle.

On the one hand, certain types of income shocks, such as natural resource shocks, may lead to there being more rents to be expropriated and more corruption. For example, Caselli and Michaels (2009) present the case of oil revenues distributed to municipalities in Brazil, as a result of the large increase in the country's off-shore oil production[4]. On the other hand, more complex business relationships may lead to demand for better government, and higher incomes may mean that countries have more resources to invest in cleaning up corruption. Bai et al (2016) suggests, using data from Vietnam, that perhaps economic growth that increases competition between jurisdictions may reduce corruption. Nevertheless, this remains an open and important question. Yet another idea, suggested by Rogoff (2016), is that the need to pay bribes in cash puts a limit on how large bribes can be—a USD$100 billion is a much larger share of incomes or GDP in a poor country than a rich one, limiting the (real) size of bribes that can be paid in rich countries.

Second, One issue with virtually all of the "hard" estimates of corruption discussed here is that they may suffer from selection bias in both directions. To the extent that measures of corruption depend on voluntary disclosure, such as surveys of bribery or disclosing links to politicians sitting on corporate boards, one might expect that corruption would be understated, as places where corruption is most severe might be less likely to disclose it. To the extent that researchers explicitly choose cases to study, corruption may be overstated, as researchers interested in corruption are not going to invest the time and energy necessary to develop a careful metric of corruption if they have a prior belief that they are not going to find much. Developing careful, rigorous metrics of corruption that are not subject to these types of selection bias is an important area for future research.

Finally, although there are some cases where corruption is substantial (e.g., the 87 percent of missing funds in Uganda or the 24 percent of missing construction expenditures in Indonesia), in many cases one is left wondering why corruption is not greater.

Open research questions:
- *How much corruption would we find if we could construct an unbiased sample?*
- *What are the mechanisms through which higher incomes lead to less corruption?*
- *What prevents corrupt officials and political leaders from extracting even larger amounts?*
- *What is the relationship between payments in cash and corruption? Does transitioning to electronic payments from cash reduce corruption?*

3.2. Does corruption matter?

Although the previous section has shown that corruption is substantial in magnitude—whether in the form of bribes given to civil servants, graft from public expenditures, or public employees not showing up to work—this does not necessarily answer the question of whether corruption actually has a negative impact on economic activity. Instead, one must test directly for whether corruption has efficiency costs.

Whether corruption matters in general equilibrium is not obvious. For example, Gorodnichenko and Peter (2007) showed that, on average, public employees in Ukraine have the same consumption levels as their private sector counterparts, even though their salaries are 24-32 percent lower. Corruption in this case does not seem to be providing extra income to these public employees, as what the government pays

---

[4] Complementing the results of Caselli and Michaels (2009), Monteiro and Ferraz (2012) find that royalty payments create a large incumbency advantage in the two elections that followed the oil windfall boom that benefited some Brazilian municipalities, but that this effect disappears in the medium-run, while Tabellini et al. (2009) find that larger transfer to municipalities increased political corruption and reduced the quality of candidates for mayor.

them is reduced exactly to offset the amount they receive in bribes. Whether corruption "matters" for economic activity in this case does not depend on the amount of corruption: instead, the economic efficiency losses (or gains) from corruption depend on whether the deadweight loss imposed by the bribes they collect is greater than (or smaller than) the equivalent deadweight loss from taxation that would be needed to raise the revenue to pay the equivalent amount of money in salaries were corruption was eliminated.

More generally, corruption could have either efficiency costs or lead to efficiency gains. This section lays out the evidence thus far on four ways in which corruption may have aggregate efficiency costs: macro impacts, the costs imposed on firms, the costs imposed on government activity, the costs imposed through the government's lack of ability to correct externalities and costs on individual citizens.

3.2.1. Macroeconomic impacts

The macroeconomic literature on corruption has tried to estimate the effect of corruption on macro variables like economic growth, investment and foreign aid. Most of these studies are cross-country comparisons that rely on international survey and perception-based indices to provide some evidence on the negative effect of corruption at the aggregate level. For example, in one prominent study, Mauro (1995) examines the effect of corruption on economic growth and finds a negative correlation between corruption and investment and growth.

One challenge is that corruption is potentially endogenous, and finding credible instruments for corruption at the macro level is difficult. Microeconomic research can offer reliable estimates of the impacts of corruption that, when appropriately synthesized, offer important lessons for policymakers and researchers alike.

3.2.2. Impact on firms

To estimate the efficiency cost of corruption on firm behavior, ideally one must know several things. First, one would like to estimate how corruption changes the effective marginal tax rate faced by firms. To the extent that bribery is used to reduce tax liabilities (e.g., bribing tax officials to reduce tax payments), one would expect the marginal bribe rate to probably be lower than the official marginal tax rate, so corruption reduces effective tax rates. On the other hand, if bribes are charged for other types of government activities, this could add to the effective marginal tax rate faced by firms.

Svensson's (2003) study of firm bribe-paying behavior in Uganda provides some clues that while there is a positive relationship between bribes and firm profits, it is very flat. Specifically, he estimates that each USD$1.00 in firm profits per employee leads to about USD$0.004 in additional bribes paid, for a "marginal bribe rate" of 0.4 percent on profits. He also finds that each USD$1.00 in capital stock per employee leads to an additional USD$0.004 in additional bribes paid, representing an additional 0.4 percent "marginal bribe rate" on capital stock. Note that these are marginal rates: the average level of bribes is substantially higher, but bribes increase relatively weakly with profits and capital stock. If the only impact of corruption was to impose a tax of 0.4 percent on profits and 0.4 percent on capital, one might expect that the impact of corruption on firm activity would be quite modest. By way of comparison, the marginal tax rate on corporate profits for large corporations in the United States is 35 percent.

Note that the Svensson study does not answer the question of the impact of corruption on firms. Instead, it only establishes what their effective corruption tax rates are. Furthermore, there may be other factors ways in which corruption could affect firm behavior beyond the marginal tax rate. For example,

many have argued that the uncertainty surrounding corruption makes it costlier than an equivalently-sized tax. Wei (2000), for example, makes this argument looking at foreign direct investment and measuring uncertainty through perceptions-based metrics. Malesky and Samphantharak (2008) examine the impact of changes in governors in Cambodia. Using survey data, they show that these changes are associated with increases in uncertainty about corruption, but reductions in actual corruption levels and decreased firm-level investment.

One recent study that seeks to quantify the impact of corruption on firm behavior is Sequeira and Djankov (2014), who examine a different type of distortion: changes in the firm's production choices designed to avoid corruption. Specifically, they estimate the economic costs and distortions associated to corruption acts at two ports in Mozambique and South Africa by directly observing bribe payments to port and border post officials for a random sample of 1,300 shipments. The study calculates the difference in the amount of bribes paid between the two ports and between goods that experienced a tariff reduction and those that did not. The results show bribes are high, frequent, and vary significantly across ports. Bribes can represent up to a 14 percent increase in total shipping costs for a standard container while accounting for a 600 percent increase in the monthly salary of a port official. Further, the study shows bribes do not only increase shipping cost but also impose high distortions on firms. Estimations suggest that even when accounting for distance and the urgency of the shipment among other variables, corruption is the strongest predictor of the choice of port. About 46 percent of South African firms located in regions in which overland costs to the port of Maputo are 57 percent lower go the long way around to Durban to avoid higher bribe payments. This represents a real efficiency loss: firms are willing to pay higher (real) trucking costs to avoid having to pay bribes in Mozambique.

Given that corruption could have both direct effects (through changing the effective marginal tax rate) as well as indirect effects (through uncertainty or other channels), it is necessary to examine the net impact of corruption on firm decisions directly. Fisman and Svensson (2007) perform this exercise in Uganda, using the same dataset as in Svensson (2003). They calculate both bribes and tax payments as a function of total firm sales, with bribes averaging 1.3 percent of sales and taxes averaging 8.5 percent of sales. They regress firm growth over a two-year period on the bribe and tax rate, instrumenting for the bribe and tax rate with industry-by-location averages. They find that a 1 percentage point increase in bribes reduces annual firm growth by three percentage points. By comparison, a 1 percentage point increase in taxes reduces annual firm growth by 1 percentage point, so bribes have three times the negative impact of taxes on firm performance. They interpret the findings as showing that the negative impacts of bribes on firm activity are higher than the corresponding impacts of taxation—with substantially large magnitudes for both.[5]

While this study provides one data point on the relationship between bribes and firm growth, substantial work remains to be done. In these studies, bribe rates are not randomly assigned, and there is also no clear estimation of the marginal bribe rate and the impact of bribes on marginal tax rates. Ideally, one would have plausibly exogenous variation in corruption levels across jurisdictions, which could then be used to estimate the impact of different types of corruption on firm behavior. Finding such a change and linking it to firm behavior, seems an important next step in understanding the efficiency costs of corruption.

Corruption could also imply that firms lack even the opportunity to compete on contracts to provide public services because government officials set up their own companies to enjoy the lucrative contracts.

---

[5] Although the level of this effect seems enormous, it is worth recalling that the bribe and tax rates are expressed as fractions of sales, not profits. Since profits are much smaller than sales, the implied bribe and tax rates on profits are much higher than those on sales, so the estimated impact of a 1 percentage point increase in a tax on profits would be substantially smaller than what they estimate. Nevertheless, the large magnitudes they estimate remain something of a puzzle.

Understanding how this dynamic results in overpriced contracts and the under-provision of products and services is also an important part of the future research agenda.

3.2.3. Impact on government provision of goods and services

Another way that leakages could have efficiency consequences is through their impact on government provisions of goods and services. This could happen in two ways. First, if leakage simply increases the cost of government goods and services, this could have an effect similar to raising the price of these goods and services. The efficiency loss would arise if projects that would be cost effective at the true costs are no longer cost effective once the costs of corruption are included, and hence are not done. Second, leakages could create additional efficiency costs through distortions. Corrupt officials usually cannot steal cash directly, as that would be easily detected; instead, they need to go through a variety of more convoluted procedures to extract rents. These convoluted procedures themselves may induce inefficiencies, which could potentially be larger than the direct cost of corruption itself. We explore both of these issues in turn.

3.2.3.1.   Price effects

One way corruption may matter is if theft of government resources increases the cost of government activity, so that otherwise worthwhile government projects—such as redistribution schemes or public works projects—become non-cost effective. Olken (2006) examines this possibility in the context of a large Indonesian anti-poverty program that distributed subsidized rice to poor households. As described above, by comparing survey data to administrative data, Olken estimates that at least 18 percent of the rice was lost from the program.

To examine the costs of corruption here, Olken performs a welfare calculation of the benefits of the program, both as it was implemented and using a counterfactual with the same targeting of beneficiaries but without corruption. The estimates imply that the welfare losses from this "missing rice" may have been large enough to offset the potential welfare gains from the redistributive intent of the program, so that the program without corruption might have been cost effective but, in the presence of corruption, it likely was not. In this particular case, the government implemented the program anyway, so in a sense the efficiency costs from lost redistribution were not realized. However, the estimates imply that in many similar cases, governments may not find redistribution programs worthwhile given the costs of corruption.

The Olken study demonstrates that a maximizing government might not have wanted to pursue the rice program given the costs of corruption in the program. The open question, however, is whether governments do indeed endogenously adjust their composition of expenditures in response to the higher prices imposed by corruption. To demonstrate these types of inefficiencies, it would be useful to show that if there was a shock to corruption that affected some sectors but not others, the government changed its composition of activities so that the quantity of activities where corruption declines increased. Note that this would be challenging to measure since we would be interested in measuring net quantities of corruption, not expenditures which include corruption—since if corruption on an activity declined, total expenditures on the activity (including corruption) could decline even though quantities could increase. We regard this question—of whether governments indeed optimize taking the price effects of corruption into account—as an important one for future research.

3.2.3.2.   Distortions

Corruption may have another type of efficiency cost beyond the price effects discussed above. In particular, since corrupt officials need to hide their activity, they may introduce two types of distortions into the procurement of government activity. First, since corruption is secret, the government may not anticipate the amounts lost to corruption (in some ways, this is the countervailing force to the price effects discussed above). It may then effectively underfund some activities relative to its preferences, once the losses due to corruption are taken into account. Second, the need to keep corrupt activity secret could also introduce distortions, as procurement officials may substitute the types of goods that make hiding corruption easier. We discuss the evidence for both of these types of corruption in turn.

The first type of efficiency impact is the effective under-provision of government activities, since the government does not fully anticipate the impact of the losses due to corruption. As described above, Olken (2007) and Olken (2009) provide evidence for this type of efficiency loss in studies of perceptions vs. reality for rural roads in Indonesia. Since villagers are better able to detect corruption where prices are marked up (where there would only be a price effect), village officials instead hide their corruption by deflating quantities, i.e., they claim to procure enough rock, sand, and gravel to make a road that is 20cm thick but instead build a road that is only 15cm thick. Since the roads they build are thinner than official engineering guidelines, they will not last nearly as long, and will need to be replaced sooner. Although Olken was not able to directly detect this quicker rate of decay in the timeframe of his study, engineers estimate that the impact of the thinner-than-design roads on road lifespan is substantial enough to cause significant efficiency losses. Using a randomized design, Olken shows that reducing corruption in the program increases the thickness of the roads and brings them closer to engineering guidelines, reducing the efficiency losses.

On the other hand, Blum (2016) evaluated the impact of the introduction of a price cap on poultry vaccines in Tanzania, compared to public service delivery agents' discretion over prices. While the average prices charged dropped by 17 percent due to a reduction in leakage, it reduced coverage of more remote areas of the country by 4.5 percentage points, from a base of 37 percent of villages in the comparison group. Taken together, these findings suggest that the price cap led to a tradeoff between affordability and service coverage, and that introducing price caps may result in reduced welfare.

Ferraz, Finan, and Moreira (2012) provide more direct evidence of the efficiency costs of corruption. They show that students in municipalities where corruption was detected in education have test scores that are 0.35 standard deviations lower than those without corruption, as well as higher dropout and failure rates. They demonstrate that higher corruption translates into lower quantities received: teachers in corrupt municipalities are 10.7 percentage points less likely to receive pedagogical training and less likely to have a computer or science lab. The study does not discuss the composition of school budgets, so it is hard to know if what the authors are picking up is price effects (there is less spending on schools because the government anticipates corruption) or distortions from corruption. One challenge in the study is that the level of corruption may be endogenous: while the authors control for other municipal characteristics, as well as corruption in other sectors and some indicators for school management practices, the level of corruption could be correlated with unobservable variables related to the quality of the school.

Examining leakages more broadly, Duflo, Hanna, and Ryan (2012b) examine the impact of teachers' lack of attendance at schools. As described above, teacher attendance is a problem because it reduces a potentially key input of the learning process. In their study, their change incentives—coupled with monitoring—obtained a 21 percentage point lower rate of absenteeism, which led to a 0.17 standard deviation increase in student test scores. The results show that the lack of attendance was indeed translating into real education losses.

Another direct estimate of the efficiency costs due to distortion is the allocation of capital from state banks. Khwaja and Mian (2005) investigate whether politically connected firms, defined as those with a

politician on their boards, obtain preferential lending. The results suggest politically connected firms receive 45 percent larger loans from government banks in spite of having a 50 percent higher default rates on these loans. Privately owned banks, on the other hand, show no such political bias. According to estimates, and assuming the default rates are equivalent to transfers from taxpayers, the deadweight loss due to corrupt lending is between 0.15 percent and 0.30 percent of GDP. When the effect of inefficient investment of politically connected firms is considered, an additional 1.6 percent of GDP is estimated to be lost each year due to preferential lending.

Sukhtankar (2012) examines distortions in prices paid to farmers in sugar mills of Maharashtra, India during election years. He finds that in politically controlled mills, those that have a chairman who sought state or national office at least once between 1993 and 2005, farmers face lower prices by approximately Rs. 20 per ton during election years in comparison to prices paid to farmers in uncontrolled mills. In the year following the election, however, farmers in politically controlled mills receive higher prices than usual by about Rs. 80 per ton if their chairman wills the election. He finds that lower prices in election years reflect revenues which are embezzled for campaign expenses, but supporters are later rewarded. Sukhtankar's rough estimates indicate that a farmer can expect to receive about Rs. 27 per ton on their principal of Rs. 20 that is 'stolen' during an election year. As farmers are likely to be aware of this pattern of decrease and subsequent increase in sugar prices during and after an election, the author notes the siphoning off of revenues for campaigning is better thought of as rents paid to politicians rather than outright theft.

Corruption can also lead to misallocation if it leads to less efficient firms producing. Khwaja and Mian (2005) argue that political influence over state banks in Pakistan channels capital to less productive firms, increasing misallocation and reducing the overall productivity of the economy. Using a list of six-years' worth of banking transactions allegedly leaked to the Russian public in 2005, Mironov and Zhuravskaya (2012) identify tunneling of money by established firms to politicians via fly-by-night firms. They find that firms with government procurement contracts intensify tunneling around the election cycle; firms without contracts do not. Furthermore, the authors report a higher correlation between tunneling around elections and obtaining government contracts in regions deemed to be corrupt. After rejecting several other potential mechanisms to explain these findings, the authors conclude that corruption is the main reason for such tunneling. The authors argue that tunneling is inefficient, as less productive firms are allocated procurement contracts and efficient firms lose out.

3.2.4. Impact on correcting externalities

A third way in which corruption may lead to inefficiency is if it lessens the government's ability to correct an externality. For example, if someone can bribe a police officer or judge instead of paying an official fine, the marginal cost of breaking the law is reduced from the official fine to the amount of the bribe. Even worse, if the police officer extracts the same bribe regardless of whether the person has broken the law, the marginal cost of breaking the law falls to zero and the law ceases to have a disincentive effect altogether. A number of studies now confirm this type of effect.

Olken and Barron (2009) examined this possibility in their study of trucking in Aceh. Specifically, they examined what happens when these trucks stop at weigh stations. Overweight trucks are a classic example of an externality: the benefits to a trucker from loading on additional weight are concave, whereas the damage the truck does to the road rises to the 4th power with the truck's weight. For this reason, governments around the world weigh trucks and impose fines on trucks that are overweight.

Olken and Barron found that virtually all the trucks in their sample were substantially over the weight limits—and in fact, 42 percent of trucks were more than 50 percent over the legal weight limit. The data

suggest that corruption at weigh stations is the likely culprit. Whereas according to the law all trucks more than 5 percent over the legal weight limit are supposed to receive a ticket, immediately unload their excess cargo, and appear in court to face a fine, in fact virtually none of the trucks received an official ticket. Instead, virtually all paid a bribe. While more overweight trucks did pay higher bribes, this relationship was very flat, and even those trucks that were not overweight at all still had to pay a bribe. Corruption thus dramatically reduced the marginal cost of driving overweight, leading to more overweight trucks.

Bertrand et al. (2007) examined a similar question in the context of drivers' licenses in India. Specifically, they randomly allocated applicants for driving licenses into three groups. The first group received a bonus if they obtained a driver's license quickly, the second group received free driving lessons, and the third group served as the comparison group. The findings confirm an efficiency loss from corruption: many people who were completely unable to drive were able to obtain licenses by paying a fee to an agent—and, in fact, the fee charged by the agent was unrelated to one's ability to drive. This efficiency loss effect was greater among the group that received the bonus for quickly obtaining a driving license since they faced a higher incentive to bypass the official procedures. On the other hand, those who were randomly allocated to the driver's license training class and who were better drivers were able to obtain their license with lower payments on average, mostly because they avoided using agents and instead used the official channel.

Oliva (2015) uses a more structural econometric approach to examine corruption in vehicle emission tests in Mexico City. The author uses indirect evidence to measure cheating by detecting serial correlation patterns in consecutive emissions, generated by car owners bribing center technicians to use emissions from a clean car to substitute for those of a cheater. Her results predict that 79 percent of centers have engaged in this behavior. Her model suggest that 9.6 percent of car owners paid bribes of approximately USD$20 to circumvent the regulation. These results suggest that the incentives for car maintenance are very low given the relatively low price of the bribe—again, showing that corruption has undermined the government's ability to counteract the externality.

The Bertrand et al. (2007), Olken and Barron (2009), and Oliva (2015) studies have very similar findings: in all three cases, those who are doing the activity the government wishes to discourage (getting a license if you cannot drive or having a truck that is overweight) do pay a higher cost that those who obey the laws. However, the marginal cost of breaking the law is much lower with corruption than it would be without corruption, so the net impact of corruption is to decrease the marginal cost of breaking the law and, thus, to decrease the effectiveness of the law.

These papers all argue that corruption has an efficiency cost because it reduces the marginal cost of the negative behavior.

Duflo et al. (2013), however, go a step further, and show that innovations that lead to decreased corruption actually do reduce the desired negative behavior. In their randomized field experiment of industrial plants in Gujarat, India, half of the plants continue to follow the status quo of hiring and paying their own third-party auditors, and the other half are assigned to auditors with fixed pay from a central source whose work is subject to random verification checks by an independent technical agency. In comparison to the control group, they find that auditors in the treatment group are 80 percent less likely to falsely report a plant's compliance with pollution regulations. Furthermore, knowledge of this increased accuracy in reporting actually compelled plants in the treatment group to pollute a significant .21 standard deviations less on average than those in the control group. Pollution reductions were entirely concentrated among the highest polluting plants, or those who face the largest risk of penalty. The key point is that reducing corruption actually leads to a reduction in the desired externality.

These studies raise an important question: given that corruption exists, how should the government structure the official laws so that the net of corruption marginal cost faced by citizens matches the government's true objective function? Put another way, if a sophisticated government knew there would be corruption at weigh stations, could they re-set the underlying fine schedule—which, after all, is the off-equilibrium punishment that determines the bribes that weigh station officials can extract— so that the net-of-corruption bribe schedule yields exactly the marginal cost curve they want? Studying this question—which would entail watching how corruption changes as the government changes the official cost schedule—is an important topic for future research in this area. It would also be informative to have evidence on this topic in sectors where the government is trying to correct externalities other than the transportation sector, such as criminal justice, tax enforcement, and environmental regulation.

3.2.5. Impact on individuals

The final question is how corruption and leakages affect individuals. In part, the impact on individuals comes from all of the channels described above: through the macro impacts, the impacts on firms' behavior, the efficacy of government service provision, and through the government's ability to correct externalities. But corruption and leakages can also affect individuals with the impact disproportionally distributed across society.

One study by Fisman and Wang (2015) shows this directly, by comparing worker fatality rates between politically connected and unconnected publicly traded Chinese companies, using firm level data on worker fatalities from 2008-2011. They define a company to be politically connected if its CEO or another C-suite executive previously held a high-level government position. They find that the rate of workplace deaths at connected firms is 5 times higher than at unconnected firms and that this relationship is best explained by firms using political connections to bypass safety regulations. Interestingly, however, when officials' promotion is contingent on meeting safety standards, then officials of connected firms are less prone to circumvent regulations, providing evidence that stricter penalties can lead to less abuse of power. Regardless, the study reveals that corruption can significantly impact individuals; in this case it leads to the disproportionate loss of life of workers whose firms ignore safety procedures due to their political connections.

3.2.6. Concluding thoughts on efficiency

The evidence in this section has demonstrated that the impacts of corruption and leakages on efficiency can be substantial. But it has also raised a number of important questions for new research.

One common theme that has emerged is that we know little about how governments respond endogenously to the presence of corruption. For example, if there are higher or lower rates of corruption in certain types of government spending, does it re-optimize spending as theory would predict, and does this re-optimization mitigate the efficiency costs of corruption? Or given that government rules to correct externalities are partially (but not completely) undone by corruption, does the government set official fines higher than they really want, knowing the official fines will not be implemented exactly? Understanding how the government responds to corruption is crucial to helping governments mitigate the efficiency losses from corruption in cases where it cannot be eliminated.

An issue on the flip side of this is the degree to which governments create regulations to maximize opportunities for corruption. An old hypothesis about corruption is that it actually increases efficiency, in the sense that it allows citizens to cut through red tape and allows bureaucracies to be more responsive to economic needs. The Bertrand et al (2007) paper provided suggestive evidence of that view, in that

people who had a high need to get a license quickly were able to do so by using an agent. But the red tape itself may be endogenously created as a way to maximize the corrupt rents captured by bureaucrats, as suggested by Banerjee (1997). Understanding whether the red tape itself is an endogenous response is an important, unanswered question, and another form of inefficiency that merits further study.

Similar issues apply to the costs of corruption for firms. While the Fisman and Svenson (2007) study suggested that bribes were costlier for firms than equivalent amounts of taxes, the tax rate could also be endogenous to the level of corruption. Gordon and Li (2009), for example, suggest that the tax code of LMICs is endogenously shaped by the presence of tax evasion, as governments reallocate tax systems towards those areas that are less prone to corruption. However, whether marginal tax rates on firms are higher or lower in corrupt countries and therefore whether the net distortions taxes imply for firms is higher or lower in corrupt countries, is an open question for future research.

Open research questions:

- *What is the impact of corruption on a firm's performance, and why is the impact of corruption on firms higher than the impact of taxes?*
- *Does corruption exacerbate misallocation across firms in the economy? How can this be reduced?*
- *How does corruption impact individuals? Does it have distributional consequences?*
- *Does corruption cause governments to re-optimize expenditures? If so, do governments shift into lower corruption sectors (to avoid distortions) or higher corruption sectors (to maximize rents)?*
- *How can the government design the official law to best achieve its desired ends in the presence of corruption?*

3.3. What determines corruption?

3.3.1. The incentives that bureaucrats, citizens, and firms face

The previous sections have discussed the impact of corruption on economic outcomes. This section examines what we know about why corruption exists, and related to this, what can be done with it.

To organize ideas, we first examine a simple framework that models the perspective of an individual bureaucrat, following the ideas of Becker and Stigler (1974). This framework treats the gains from corruption (the bribe) as fixed and asks when corruption will be preferable to honesty. We then examine what happens when the optimal bribe is determined by the bureaucrat taking into account market forces, following the ideas of Shleifer and Vishny (1993). The subsequent sections discuss the empirical evidence to date along all of the dimensions suggested by the simple theoretical framework.

Suppose that the bureaucrat receives a wage w from the government and, if fired, can receive an outside option v. The bureaucrat can decide to be corrupt or honest. If he is corrupt, he is detected with probability p, is fired, and receives outside option v. If he is undetected, he receives his wage w plus the bribe b, less a dishonesty cost d. In equilibrium, he will be corrupt if and only if $w - v < \frac{1-p}{p}(b - d)$.

This simple framework suggests several avenues for reducing corruption. One could increase the returns to staying on the job (w), or, equivalently in this context, one could decrease the outside option (v) by increasing punishments. One could also increase the probability of detection (p).

One implication of this simple framework is that if there is heterogeneity in *d* among potential bureaucrats, there can be selection where those who are most likely to be corrupt (those who have the lowest dishonesty costs *d*) will self-select to be more likely to become bureaucrats. Suppose that *d* in the population is distributed uniformly from 0 to $\bar{d}$. If $w\text{-}v > \frac{1\text{-}p}{p}$ (b), then nobody will be corrupt, regardless of their level *d*, and there is no reason that the distribution of *d* among bureaucrats will be different than the distribution of *d* in the population.

If, however, the above inequality does not hold, then people with low *d* will have a higher utility from becoming bureaucrats than those with high *d*, since they will be relatively more efficient at corruption, so depending on how the government allocates jobs we might expect to have more low *d* people among bureaucrats than in the population. This implies that corruption may be harder to combat than one might expect, since a corrupt system may attract those types of bureaucrats who are more prone to corruption. It also implies that the effect of a given anticorruption policy (i.e., a vector (*w,p*)) will depend on past levels of anticorruption policies, since those past policies will influence the selection of bureaucrats.

The simple framework thus far has treated the amount of the bribe, *b* as exogenous. In practice, however, the bribe *b* may be set by the bureaucrat to maximize his profits. Specifically, conditional on deciding to be corrupt, the bureaucrat will set his bribes to maximize his profits, which are the number of bribes he receives multiplied by the price, i.e. $q_i b_i$.[6] The key insight of Shleifer and Vishny (1993) is that the optimal solution depends on what other bureaucrats are doing and how they set prices. If a person needs permits from two different bureaucrats to complete a transaction, and both set prices independently, then each bureaucrat solves $\max_{b_i} Q(b_i + b_{-i})b_i$, taking the other bureaucrat's bribe b_{-i} as given. In such a case, the total amount of the bribes ($b_1 + b_2$) will be higher than if there had only been a single bureaucrat, and the total quantity will be lower. Conversely, if a consumer needs a single permit which can be obtained from either bureaucrat, they will compete against each other and reduce the bribe beyond what a single, monopolistic bureaucrat would charge. The key insight is that the bribes themselves may be a function of the structure of the bureaucracy, and that changing the nature of the organization may have important implications for the level of corruption.

This simple framework is, of course, quite stylized, but it highlights the important role that both the incentive structure faced by individual bureaucrats (be it compensation, monitoring, selection, or other incentives) as well as the bureaucratic organization may play in influencing the amount of corrupt behavior. This section discusses the evidence to date on each of these factors in turn.

3.3.1.1. *The bureaucrat's decision problem: The role of compensation, selection, and other incentives for bureaucrats*

Research in many LMICs has attributed poor governance to the individuals performing government functions. However, rather than their personality traits, it may be that the weak institutional structure within which public employees work in these countries is the dominant cause. The two may also interact: weak institutions may cause corruptible individuals to select into public service. Cooper (2018) provides some insight on this reverse causality issue through an experiment in Papua New Guinea where applicants are randomly appointed to be police officers and tested on their prosociality eight months later. In running behavioral games, he finds that the policing power granted to applicants did not corrupt them: civilians hired as community police cheat against themselves to award a windfall to the community member

---

[6] To simplify the analysis, we will assume that conditional on being caught, the probability of being caught does not depend on the amount of the bribes or the quantity of bribes, though one could easily generalize the model to include these effects.

against whom they are playing. The dependence of community police on support from their community to coerce offenders might explain this finding.

Section 4 focuses on the importance of the "personnel economics" of the government sector—in particular, compensation, selection and recruitment of employees, incentives, and monitoring. The potential of field experiments in advancing our understanding and identifying civil services reforms of these three arenas is highlighted. We refer the reader to that section for an in-depth treatment of these issues.

3.3.1.2.   The market for bribes: Changing the structure of the bureaucracy to harness the forces of competition

The approach taken in analyzing the bureaucrat's decision problem in Section 4 describes an approach to corruption and leakages based on a principal-agent framework, focusing on how a principal—the government—can best monitor its agents—civil servants. The key focus of analysis is therefore the strategic relationship between principal and agent. In other settings, however, strategic interactions between corrupt agents themselves become important. In this view, articulated by Shleifer and Vishny (1993), corrupt agents behave like profit-maximizing firms, and the level of corruption is determined not just by monitoring, but also by the structure of the "market" for bribes, the elasticity of demand for the officials' services, and the degree to which corrupt officials can coordinate with one another in setting prices. As outlined in the theoretical section above, depending on how the market is structured, these kinds of strategic interactions can either raise or lower the bribe amounts.

If a person needs to bribe multiple corrupt officials to perform a given task, Shleifer and Vishny argue that that the "double-marginalization" problem can arise. Specifically, if each agent does not fully internalize the effect of their bribes on other officials' bribe revenues, the total amount of bribes one would need to pay could be higher than if the agents had acted independently.

Olken and Barron (2009) use the data they collected on the bribes truck drivers pay to empirically test idea that market forces partially determine the level of corruption and specifically to test for this type of double-marginalization. They exploited the fact that, during the period studied, the number of checkpoints along one of the roads was reduced in accordance with a peace agreement signed earlier in the year. They used this change in market structure to estimate the elasticity of the average bribe paid with respect to the expected number of checkpoints. They show that the average price paid at checkpoints increases when the number of checkpoints declines, consistent with double-marginalization along a chain of vertical monopolies. The results provide evidence for the Shleifer-Vishny view that the market structure has an impact on the total amount of bribes charged, and more specifically, that price setting in this particular context is decentralized rather than centralized. These findings highlight the need to consider strategic interactions between corrupt agents themselves, in addition to interactions between principals and agents, in designing effective anticorruption policy.

An implication of this view is that a policy reform that moves from having a large number of independent agents to a single agent may reduce corruption and increase economic efficiency. Bruhn (2008) uses the sequential implementation of a reform that simplified business entry regulations across municipalities in Mexico to estimate the economic effects of such reforms more convincingly than in cross-country data. Although the paper does not look at the effects of the reform on corruption directly, the results do show that simplified regulation leads to increased efficiency. She finds that the reform increased the number of registered businesses by 5 percent, which was accounted for by former wage earners opening businesses. Wage employment also increased by 2.2 percent as a result of the reform, while competition from new entrants decreased the income of incumbent businesses by 3 percent.

Conducting a similar study of the same reform of business entry regulation in Mexico, Kaplan et al. (2011) compare eligible industries to ineligible industries in municipalities which implemented the reforms. They find that the program generated 5 percent more new formal firms per month in eligible industries compared to ineligible industries. At the same time, however, they find these results to be temporary, concentrated in the first 15 months of implementation. Still, their back of the envelope calculations indicate that the benefit of the reforms are six times greater than the costs, so there are certainly efficiency gains with the simplification of the previous business entry regulation.

The flip side of strategic interactions between bureaucrats is that if bureaucrats are competing against one another, this could reduce the bribes paid and lead to lower bribes and more output.

One recent study that examines this is Burgess et al. (2011), which explores this issue in the context of deforestation in Indonesia. In particular, the study explores a setting in which local district forestry officials can allow logging beyond the legal logging quota in exchange for bribes. The study shows that as the number of political jurisdictions increases, so that there are more bureaucracies with the potential to facilitate illegal logging in a province, logging rates increase and prices for wood fall, consistent with a model of Cournot competition between bureaucrats.

Despite the potential for competition between bureaucrats to reduce bribes, other than the Burgess et al. (2011) study we know of no other evidence that examines how competition between bureaucrats works in practice. In the Burgess et al. study, competition occurs only through the product market—each district chooses how much wood to extract and market forces—a common demand curve—determine how much they receive in rents. In many other settings, however, individual agents would be able to choose which bureaucrat to work with to obtain a service, and the bureaucrats might compete on price. This type of Bertrand competition could result in even larger impacts of competition than the Cournot competition studied by Burgess et al. (2011). We regard studies examining Bertrand competition between agents as a first-order question for future work.

It is important to note that, in this context, competition leading to lower bribes is not necessarily socially optimal. In particular, it depends on what the government is trying to accomplish and whether the bribes are on top of, or instead of, official government fees. For example, in the case of deforestation studied by Burgess et al., bribes were to allow more logging than the government had deemed optimal (for example, for reasons of watershed protection or biodiversity protection). Competition meant lower bribes and greater quantities, which in this context meant more illegal logging, and hence greater social losses, than had there been less competition. On the other hand, in the case of the road checkpoints studied by Olken and Barron, traveling the road should have been free, so lower bribes would have meant greater road travel and greater efficiency. Understanding the welfare implications of these types of strategic efficiency depends therefore on whether higher or lower bribes would increase or decrease social efficiency, and we do not yet know of an empirical example demonstrating how competition between bureaucrats could lead to greater social efficiency.

Open research questions:

- *Can changing the structure of the bureaucracy to encourage competition between government officials lower bribes?*

3.3.1.3. Incentives for citizens and firms

An alternative approach is to focus on the citizen, who is contemplating paying a bribe or being honest. Can this approach reduce corruption?

One view comes from Kaushik Basu, then the chief economic advisor for the Indian government, who made a provocative argument that the act of paying bribes should be legal (Basu 2011). His argument is that if the bribe payment itself is legal, then after the bribe is paid, the bribe payer will then have an incentive to cooperate with law enforcement against the bribe taker. Anticipating this, bribe takers will not want to take bribes. This is a very innovative idea, but one for which—to the best of our knowledge—there is no evidence on to date.

Of course, the converse could also be true: increasing enforcement actions against citizens could also reduce corruption. On this side, the best evidence comes from work on tax evasion. Carrillo, Pomeranz, and Singhal (2016) point out some of the limits of this approach. Working with the Ecuadorian government, they sent letters to randomly selected firms informing them of detected revenue discrepancies. They found that most firms were able to substitute to alternate forms of tax avoidance behavior and in response paid almost more taxes. Castro and Scartascini (2015) worked with the Argentine government and sent letters to various types to taxpayers. They found that sending letters informing taxpayers of the penalties of non-compliance increased tax compliance by about 4 percentage points, but other letters—such as those emphasizing the importance of equity or fairness in the tax system—had no impact.

Recent research has, however, highlighted another very important role for citizens: that they themselves can improve enforcement and reduce illicit behavior. Several studies point to this, in different ways. Pomeranz (2015) conducts a randomized trial to study the self-enforcing properties of the Value Added Tax (VAT) system in Chile, where VAT evasion was estimated to range between 12 and 27 percent in the period preceding the study. Pomeranz finds that the VAT system, which generates a paper trail on transactions between firms, can facilitate tax enforcement by spreading the impact of enforcement measures up the production chain—that is, by creating incentives for firms to keep each other honest. In a randomized evaluation with over 400,000 firms in Chile, Pomeranz finds that firms that received a deterrence letter from the tax authority increased their median monthly VAT declarations by USD$2.65 on average, a 7.6 percent increase from USD$35.04 in the comparison group. Furthermore, announcing plans to audit a firm increased tax payments by its suppliers, but not by its clients, illustrating spillover effects of audits through a firm's trading network.

Naritomi (2016) finds very similar effects in her study of a policy in Brazil that encouraged consumers to make sure that firms paid tax. She studies a program that provided monetary rewards for citizens based on whether firms had paid value added tax on each transaction. She finds that the introduction of this program led to increases in tax payment for firms that were more exposed to this (in virtue of having more final consumers). Kumler, Verhoogen, and Frias (2015) have a related finding in Mexico, where they argue that a tax reform which made workers' pensions a function of how much firms paid in payroll taxes resulted in firms paying more payroll taxes, which they also interpret through the lens of citizens-as-monitors. Dunning at al. (2017) test the limits of incentivizing citizens for good behaviors. The experience of a lottery-based one-year tax holiday in Paraguay to reward good taxpayers led to persistent negative effects: lottery winners were 3 percentage points less likely to pay their taxes 3 years after the tax holiday, suggesting that taxpaying is habitual and disruptions may backfire.

Researching the citizen side of corruption, Reid et al. (2007) randomly offered financial incentives for motorcycle taxi drivers in the Democratic Republic of Congo to provide official toll receipts and receive reimbursement (rather than engaging in unofficial, non-reimbursed bribes for lesser toll amounts). They find that, even with full reimbursement, less than 14% of the drivers provided receipts, representing a 4 percentage point increase relative to the control group that received no incentives. The researchers suggest that these results support the view that corruption is preferred by citizens when it enables faster

access to public services and avoids navigating "red tape," particularly in settings with low-quality institutions.

The role of citizen behavior in corrupt relationships is an important area for further research.

Open research questions:

- *Does making it easier to report corrupt activities (for example, by making paying bribes legal, or otherwise) discourage corrupt activity in the first place?*
- *Can citizens be leveraged as anticorruption agents outside of the tax setting? If so, how?*
- *How should government best allocate limited audit resources against the population to minimize tax evasion and bribery?*

3.3.1.4. Corruption as a bargaining outcome

Having considered incentives for corrupt officials and incentives for citizens, we can now consider the two together. In particular, a relatively new empirical approach to corruption thinks of bribes as part of the equilibrium of a bargaining game between both sides. In that approach, the amount of the bribe is determined by each side's outside option (i.e. the payoff to both the corrupt official and the citizen from being honest and not engaging in corruption) and their relative bargaining weights. While this approach may seem evident, it can yield surprising and counterintuitive predictions.

Khan, Khwaja, and Olken (2016) study this issue in the context of tax collectors in Pakistan. They examine, using a randomized evaluation, the impact of providing incentives to tax collectors, where tax collectors are paid a fraction of all taxes they collect. The results show that this increases tax revenue collected quite substantially. But, they also show evidence of the perverse consequences of bargaining: the incentive payment increases the tax inspector's outside options for being honest, which (for those taxpayers that continue to still pay bribes) means that they need to pay the inspector a *higher* bribe, to compensate for the foregone incentive payment.

An implication of this bargaining approach is that changing the underlying threat point—for example, the true tariff should be paid in an honest equilibrium—can reduce the bribe paid, since it improves the outside option of the tax collector. Sequeira (2016) exploits exogenous changes in the tariffs between South Africa and Mozambique to examine how changing regulation can affect corruption. Using data on trade gaps, as well as direct bribe evidence collected from mandatory clearing agents at the ports, Sequeira finds that reductions in tariffs lowered bribes on both the extensive margin and the intensive margin. She finds that a 10 percent decline in tariff rate corresponds to a 22 percent decline in probability of paying a bribe, and that a 1 percent decline in tariff rate is associated with a 20 percent decline in the amount of bribe paid. There is some evidence of displacement where, in the absence of tariff evasion bribes, officials threaten to delay the imports using inspections or other coercive techniques. However, these effects dampen the corruption reduction by only approximately 10 percent. Because lower tariffs reduce bribing power as well as tariffs, there are minimal net effects on firm import decisions.

Open research questions:

- *Would providing incentives to citizens to behave honestly decrease bribes?*
- *How can the government reduce the bargaining power of potentially corrupt officials?*
- *Can incentives for government officials be deployed effectively without increasing bribe rates?*

- *Are there multiple equilibria in corruption? If so, what causes this? And if so, can temporary corruption crackdowns have permanent effects?*

3.3.2. External factors

The levels of corruption may be influenced by factors that intermediate the relationship beyond the control of either the bribe payer or the bribe taker. This section reviews the evidence on four important factors: technology, transparency, the judiciary, and demographics.

3.3.2.1.   Technology

Technology is a very broad term, and there are many ways technology could affect the level of leakages. Broadly speaking, we consider two ways in which technology could matter: by providing tools that are hard for humans to tamper with and by enhancing communication. It is important to point out, however, that there are other mechanisms through which technology could matter, many of which likely have yet to be discovered.

3.3.2.1.1.   Providing tools that are difficult to tamper with

For many (though not all) corrupt activities, the corrupt agent needs to somehow evade the rules or procedures that the official government bureaucracy has set up. Technology can help address this problem by ensuring mechanically that certain procedures are followed. Two landmark studies are reviewed here, and a longer discussion on the move towards smart governance is provided in section 4.4.1.

The first example is the cameras study conducted by Duflo et al. (2012b) in India. In this study, teachers need to verify their attendance at school by having a student take a picture of them with the entire class at the start and close of each school day. Technology—the camera—means that teachers could not simply fudge their attendance records. The use of this monitoring device coupled with monetary incentives that linked salaries with attendance reduce absenteeism from 42 percent to 21 percent and raise children's test scores by 0.17 standard deviations.

Of course, technology is only as effective as the system it feeds into. Banerjee et al. (2008) evaluate the impact of an incentive program on nurse attendance in India showing that when the incentives were effectively in place, as they were during the first 6 months, they produced a substantial improvement in attendance, according to some measures a 100 percent increase. However, over time, the local health administration deliberately undermined the incentive system by allowing more and more excused absences, so that by the end of the sixteenth month the program produced no effect on absence rates.

3.3.2.1.2.   Increasing communication

Technology can also have a substantial impact on corruption by facilitating communication. Banerjee et al. (2016) provide a useful example from India. In their study, they randomized blocks in India into a new, real-time back-end payment structure. Specifically, in the new system, payments for a workfare program were only made after local administrators had entered all the names of those who had worked into the system, rather than being made in advance. This technology-based advance, which allowed up-front payments, reduced leakages by about 15 percent.

Electronic procurement is another area where technology may be able to play an important role. Lewis-Faupel et al (2016) study the introduction of e-procurement systems for public works contracting in both India and Indonesia, using a differences-in-differences design. They find that the key effect of e-procurement seems to have been enhancing access to bidding information—perhaps because it was being intentionally withheld from non-favored bidders, but also perhaps because technology allowed firms located physically further away from the job site to bid. While e-procurement did not reduce the prices paid by governments, it did seem to improve both timeliness and quality.

3.3.2.1.3. *Other roles for technology*

A major innovation in LMICs over the past five years has been the rise of electronic payments. Muralidharan, Niehaus, and Sukhtankar (2016) show using a randomized evaluation how introducing smart-cards for payments reduced leakage from the NREGA program in India. Other countries around the world are beginning to introduce smart-cards, but these card systems may raise a whole other set of issues, such as the power of the agents who redeem the money and the ability of low-literacy populations to use these systems effectively, which may need to be addressed.

Technology can also be used to detect fraud. In many countries, technology has played an important role in the design and administration of the tax system. Yet this has been underexplored in a LMIC context where bribery and tax evasion are more prevalent. One key idea of tax enforcement is double-reporting, where the tax department compares two independent reports about tax performance and investigates discrepancies. Here, the best evidence is from high-income countries: Kleven et al. (2011) analyzes a randomized tax enforcement experiment in Denmark and find that the tax evasion rate is very small (0.3 percent) for income subject to double-reporting and much higher (37 percent) for self-reported income. Technology plays a key role here: in a manual system actually doing the matching from all the double-reported information would be very challenging, but once the system is automated it is much easier. Machine learning techniques to detect fraud are also an exciting new area for research. Given the large number of countries in the process of modernizing and computerizing their tax infrastructure, it should be possible to study the impacts of this type of technology in the context of a LMIC where tax evasion is usually higher.

Open research questions:

- *How can technology be leveraged to reduce tax evasion and decrease leakages from tax?*
- *Does e-procurement prevent corruption or prevent collusion?*
- *What is the role for electronic payments in reducing leakages? How can the challenges of electronic systems be mitigated?*
- *How can technology be used to effectively detect and deter fraudulent behavior, reducing leakages to the government?*

3.3.2.2. *Transparency*

One of the key themes of the international anticorruption movement is the role of transparency—so much so that the largest worldwide anticorruption NGO is called "Transparency International." But does transparency matter? And if so, what are the mechanisms through which transparency could matter?

The basic idea about transparency is that by enabling information about government actions, citizens and civil society can do a better job monitoring what the government is doing and demanding better performance. However, the effect of making information about politicians publicly available is a priori

unclear. While disclosure of information can improve political accountability, it can also undermine politicians' privacy and thus may disincentive more qualified candidates from participating in politics.

Several pieces of evidence suggest that, indeed, there is a relationship between providing access to information about politicians' performance and both the political accountability and the quality of government. In a cross-sectional, cross-country study, Djankov et al. (2010) study the relationship between disclosure rules for information about parliament members and a numbers of measures of quality of government and corruption. Their main conclusion is that public disclosure, but not internal disclosure to parliament, is associated with lower perceived corruption and better government. They further find that information about politicians' assets, liabilities, income sources, and potential conflicts, as opposed to simply income and wealth levels, are more consistently associated with better government. Since the study is cross-sectional in nature, they cannot rule out reverse causality (i.e., higher quality governments adopt better transparency laws).

In a more micro example, Banerjee et al. (2010) study how public disclosures about politicians' performance and qualifications can influence electoral accountability in settings characterized by weak institutions and a less educated population by conducting a randomized experiment in Delhi, India. Using the Indian Right to Information Act and candidates' affidavits, they created report cards for ten assembly jurisdictions during the run-up to the 2008 election in Delhi. They then randomly provided slum dwellers with pamphlets and free newspapers containing information on candidate qualifications and legislator performance. The information increased voter turnout by 3.5 percentages points and reduced the incidence of vote buying by 19 percentage points. The information campaign seems to increase the quality of government: the vote share of the best performing incumbent increased by 7 percentage points in the treatment group relative to the controls. An interesting next step would be to see if, over the long term, increased transparency in certain assembly districts encourages better quality candidates to run and/or encourages politicians to improve their behavior while in office.

A related example is Ferraz and Finan's (2008) study in Brazil on how information about incumbent performance increases accountability and how the presence of the media can influence this result. The authors take advantage of the implementation of a federal government anticorruption program that randomly selected the municipalities where the use of federal funds was audited and the results publicly disseminated. The study showed public dissemination of corruption scandals in local governments had a negative effect on incumbents' electoral performance, especially in those areas where local radio stations were present to broadcast the results of the audit reports. The probability of reelection for an incumbent who committed two corruption violations in municipalities with preelection audit was 7 percentage points lower than one who had zero violations and 11 percentage points lower if radio stations were present in the municipality. One interpretation for the larger effect in municipalities where local radios were present to divulge the information is that radios are more efficient in transmitting information about local politics to smaller municipalities.

In a follow-up paper, Avis, Ferraz and Finan (2016) also find long run effects. Examining Brazilian municipalities receiving audits, they find municipalities receiving a subsequent audit exhibit significantly lower levels of corruption than newly audited municipalities. This is consistent with audits having a disciplining effect due to electoral, legal, or reputational costs.

Given the importance of the media in facilitating transparency surrounding government actions, a biased media can be a major roadblock to accountability. Research carried out in Argentina by Di Tella and Franceschelli (2011) finds evidence of the media being influenced by government transfers in the form of advertisement bought in newspapers. Documenting the front page coverage of government corruption in the four main national newspapers from the period between 1998-2007, they find that a one standard deviation increase in government advertising in a given month for a given newspaper is

associated with a reduction in coverage of corruption scandals in that newspaper by .23 of a front page. This relationship is statistically significant at the 1 percent level even when controlling for month, newspaper, and a newspaper and president interaction dummy. Their results suggest that even when the media is not out rightly controlled by the government through censorship or intimidation, the government can use financial incentives to influence the media in a manner that limits its usefulness as a mechanism for holding government officials accountable and providing valuable information to voters.

A second way that transparency may matter—and the way that many suggest it does—is by providing citizens with information on what they are entitled to. Banerjee et al. (2018) show using a randomized trial in Indonesia that providing identification cards to beneficiaries of a subsidized rice program with information as to their rights under the program increased the amount of subsidy they received by 25 percent. Additional information, such as printing the official price on the card, or announcing the cards publicly, further increased the amount received. The authors show that this represented a decrease in leakages, not a transfer from some households to another.

A third way in which transparency could matter is by allowing citizens to signal interest in a particular outcome. Peisakhin and Pinto (2010) examine this by conducting a randomized experiment in India to test whether freedom-of-information laws can be used to obtain greater access to basic public goods that are otherwise attainable only through bribery. The experiment randomly assigned individual applicants in India to one of three mechanisms used when requesting public benefits and then tested the effect of these mechanisms on the time that elapsed before the applicant received the benefit. In the first treatment group, applicants submitted an information request under the Right to Information Act shortly after their applications. The second group of applicants presented a letter of support from a local NGO with their application. Finally, the third group of applicants paid a bribe to a local to obtain the benefits. According to the results 94 percent of those who pay bribes or sent an information request received benefits over the course of one year, as opposed to 21 percent in the NGO and control groups. Individuals in the group that paid bribes received benefits in a median of 82 days, 38 days less than those in the groups that filed an information request. The groups that neither paid a bribe nor requested information only obtained benefits after 343 days. The results suggest that requesting information under the freedom of information law is a reasonable, though imperfect, substitute for bribing an official.

In a follow-up study, Peisakhin (2012) estimates the effect of the freedom-of-information law in the process of voter registration. The results suggest that the information law is an effective, free and legal substitute to bribery for middle class applicants. Once the time it takes to process the request for information is considered, paying bribes does not reduce the number of days needed to register for voting. The experiment is then replicated for poor applicants with similar results, suggesting the information act can also empower under-privileged groups.

Despite these recent studies, the academic literature still does not necessarily have the answer on the impact of some of the more macro-transparency questions that are the focus of much international advocacy. For example, there is a movement in the corruption and leakage field to increase overall budget transparency, through such organizations as the International Budget Partnership. There has also been a movement to increase transparency of natural resource revenues, through the Extractive Industries Transparency Initiative and groups such as Revenue Watch. Another important innovation is the requirement that politicians declare their wealth. To the degree that these macro initiatives operate at the country level, identifying their impact is harder, and an important area for future research.

Open research questions:

- *How does transparency about politician wealth decrease corruption and leakages?*

3.3.2.3.   Judiciary

Judicial corruption may pose a particular challenge because the judicial system is a key component in the efforts necessary to fight other sorts of corruption. After all, if judges or prosecutors are corrupt, then it will be difficult to impose punishments against corrupt officials, since they can bribe the prosecutors or judges and avoid punishment. (Of course, they would still have to pay the cost of bribing the prosecutor or judge, which would pose a disincentive for corruption, but as discussed above these types of disincentives in practice seem flatter than the actual punishments imposed.)

Judicial corruption can also have important efficiency implications for civil contract enforcement. If judges are bribable, then in the event of dispute contracts will be awarded to the party who is able to bid more. This will be ex-post efficient (i.e., the outcome will be in favor of the party with the most to gain ex-post), but not necessarily ex-ante efficient, and may therefore discourage certain types of contracts. For example, in high-income countries, companies often sign long-term sourcing contracts for the fuel of power plans before the plant is built, to ensure a steady stream of inputs to the plant. If the price of fuel rises, ex-post the fuel provider might wish to renege on the contract and sell fuel at a higher price to an outside party, which an outside (bribable) judge might let them do. But knowing that this might happen, the power plant might not get built, or if it did, it might have a different ownership structure where it also owned the fuel provider, to prevent this type of dispute.

Despite the importance of judicial corruption, there is very little rigorous literature on the topic, both in terms of the consequences of judicial corruption and how to control it. Indeed, there is relatively little empirical literature on courts in LMICs at all, with most literature focusing on delays. In one study, Visaria (2009) use the introduction of debt recovery tribunals to study the effect of judicial quality on the repayment behavior of borrowers and the lending decisions of banks in India. In a cross-country study, Djankov et al. (2003) explore the effectiveness of courts as mechanisms of resolving simple disputes and examine how the characteristics of the legal procedure affect courts' efficiency and the ability to deliver justice. More recently, Sadka et al. (2018) estimate the effects of providing information to parties in randomly selected court cases in Mexico on predicted case outcomes and conciliation services. They find that bridging this information gap is welfare-improving for the plaintiffs, nearly doubling the overall settlement rate—though only when the plaintiff receives the information directly.

We regard these areas as important directions for future research.

Open research questions:
- *What are the impacts of judicial corruption on economic efficiency?*
- *How can the extent and effects of judicial corruption be measured?*
- *What approaches work best in reducing the levels and efficiency costs of judicial corruption?*
- *Is it possible to insulate the judiciary from corruption that is endemic in the rest of the government?*

3.3.2.4.   Demographics and corruption

Individual and group demographics—including ethnicity, gender, education, or socioeconomic status—may be a salient factor in the bargaining process around corruption. For example, the equilibria for bribes may differ according to the demographic characteristics between a citizen and a government official.

With regards to electoral participation, demographic factors associated with the ethnic and religious identity of election officials might introduce biases in the management and administrative process during elections. In turn, such biases can have important impacts on election results and policy choices. During the 2014 parliamentary elections in Bihar, India, Neggers (2018) leveraged a national policy that randomly assigned of election officers to polling stations to examine whether the identity of polling station officers affected voter behavior and voting outcomes. Researchers compared polling stations with a "diverse" team of officers—consisting of a team with at least one officer with a minority ethnic and religious background—with polling stations with a "homogenous" team in terms of ethnicity and religion. The identity of polling station officers significantly impacted voting outcomes: more diverse officer teams increased the vote share for the minority-oriented coalition. The effects were concentrated in areas where officers had wider discretion over determining each voter's eligibility. Researchers found suggestive evidence that the effects were driven by identity-based bias in the decision-making of election officials.

Biases on the basis of individual and group demographics may also affect the prevalence or type of corruption. We encourage future research that examines the role of demographics in reducing corruption and leakages.

Open research questions:
- *Are demographics a salient factor in the bargaining process around corruption? If so, through what channels?*
- *How do biases based on gender, ethnicity, education and socio-economic status affect the prevalence or type of corruption?*

3.3.3. Anticorruption policy in the long run

Much of the evidence discussed above shows the short-run effects of anticorruption policies. But there is ample evidence to believe that the long-run impacts could be quite different. For example, it could take corrupt officials time to learn how to manipulate a new system, so the long-run effects of an anticorruption policy could be smaller than the short-run effects. Alternatively, it could take time for a new group of civil servants to select into the system, so an anticorruption policy could be more effective over time if it encourages more low-corruption types to select into the civil service.

There are several examples that suggest that the long-run effect of anticorruption policies may be smaller than the short-run effect as officials adapt. One of the examples already described is Banerjee et al. (2008). In this study, an incentive program on nurse attendance in India was found effective only during the first 6 months of the intervention, when the program was correctly in place. Later, however, the system was undermined by the local health administration and the program was no longer able to improve nurse attendance.

Bobonis et al. (2013) also provide an example where the disclosure of information about corruption practices only induced a reduction in corruption levels in the short run. They compared the level of reported corruption for municipalities audited before and after an election in Puerto Rico. The results seemed to confirm the positive effect of the program showing corruption was lower in municipalities audited before the election. However, in subsequent terms corruption levels increased, especially among those who refrained from rent-seeking activities in the first audit.

In the case of Brazil, Ramalho (2007) uses the 1992 impeachment of President Fernando Collor to evaluate the impact and persistence of corruption on the market value of politically connected companies. The results suggest that the market perceived the decrease in the president's probability of staying in power as affecting the value of politically connected companies, but only temporarily. According to the

results, family-connected companies had on average daily abnormal returns 2 to 9 percentage points lower during bad event days, with the effect reversing completely within one year. One interpretation is that over the course of the year, these previously politically connected firms were able to form new connections.

In Colombia, Camacho and Conover (2011) study the manipulation of the poverty index score as an eligibility requirement to gain access to social programs. They show strategic behavior in the timing of household interviews around local elections, and direct manipulation of scores given by a sharp discontinuity on the scores exactly at the eligibility threshold. The results suggest that in total three million people had their score changed, which accounts for about 40 percent of the beneficiaries. Also, they provide some evidence that manipulation increased in cases where the elections were more competitive and decreased with more community organizations and higher newspaper circulation.

Another example is Burgess et al. (2011). They examine the determinants of illegal logging in Indonesia using satellite data to track logging and find that when a district's oil and gas revenue increases, providing an alternate source of rent extraction for local district officials, illegal logging falls. However, within 3 years, the effect reverses and illegal logging is back to its previous level. Burgess et al. provide suggestive evidence that the mechanism is a change in the political equilibrium—the higher oil and gas rents change the nature of the governing coalition towards a type of coalition associated with higher rent extraction. This new political coalition presumably extracts rents not just from oil and gas but also from the forest sector.

In sum, the evidence in this section provides a word of caution: in many cases, reductions in corruption in the short run do not always persist in the medium run, with effects often being undone within a period of 1-3 years. This suggests that, to the extent possible, anticorruption actions should be tracked over a period of several years to determine whether the impacts are persistent.

Open research questions:
- *How can corruption be effectively combatted in the long run?*
- *How can the impacts of anticorruption programs be made to persist?*

4. Public service delivery: Selection, incentives, and monitoring

In this section, we examine a particular determinant of government performance: the individuals who perform government functions. We focus on two groups of public employees: appointed civil servants, which we broadly construe to include administrators with effectively permanent government appointments, and frontline service providers (e.g., teachers, nurses, firefighters, and trash collectors) who may have either permanent appointments or temporary contracts.

In many contexts, policy actors and researchers attribute poor governance to public employees being lazy, corrupt, or both. Yet, it may be that the poor institutional structure within which public employees work in these countries is the dominant cause. The two may also interact: poor institutional structures may cause the lazy and the corrupt to select into public service. Since both institutional structures and personnel selection are endogenously determined, establishing causality is hard. By providing a clean empirical method for identification, field experiments can help cut the Gordian knot and identify problems in personnel selection and management, elucidate the causes of these problems, and suggest potential solutions. These contributions of field experiments are the focus of this section, in particular the selection and recruitment of personnel, incentives, and monitoring.

Several key features of the state distinguish its personnel practices from those of the private sector, particularly in LMICs, and here we identify five that are particularly relevant to our analysis.

First, the state has a long horizon. Absent significant civil conflict, most states anticipate collecting and spending revenues indefinitely. This allows states to make long-lived promises to its employees, such as pensions, which may be difficult for the private sector to make. The long-lived reputation of the state as an employer also means that it may be reluctant to renege on such promises; in fact, many promises of the state, again such as pensions, survive radical regime change, at least in nominal terms.

Second, the set of contracts a state can offer its employees is limited. Whereas shareholders can create strong incentives for CEOs to maximize returns, the mechanisms that the ultimate principals—citizens—have at their disposal, namely elections, are coarser and more limited. Thus, politicians may seek to use jobs, and the wages associated with them, to reward their political supporters, cronies, and friends. Politicians could also be tempted to use promotions or incentives to exert undue influence on civil servants. These issues are more prevalent in the public sector than in the private sector because of the lack of discipline from the profit motive: the politician only indirectly bears the cost of inefficiency to the extent that voters are less likely to re-elect him because of it, whereas the owner of a firm directly feels the financial losses associated with these types of inefficiencies (Shleifer and Vishny 1994, Boycko, Shleifer, and Vishny 1996). To counteract the tendency for each new politician to replace large numbers of government employees with his political supporters, over time governments have enacted rigid civil service rules that restrict the discretion politicians have over hiring and firing (Evans 1995).[7] In fact, these civil service systems are typically much more rigid than their private sector counterparts, with strict formulas defining the hiring criteria, promotion patterns, and wage levels. The need to isolate the employment decision from political influence—which underlies a substantial amount of public sector personnel policy—suggests that the personnel economics of the state are likely to substantially differ from those of the private sector.

Another restriction on contracts is that public sector compensation usually does not include pay for performance. Performance pay for bureaucrats can create severe multi-tasking problems, where bureaucrats focus on the incentivized dimension of their job at the expense of the non-incentivized dimension (Holmstrom and Milgrom 1991). While multi-tasking is an issue in many contexts, it can be particularly severe in public sector contexts where agents wield substantial authority (e.g., police and judges) and it is hard to find an objective measure of the "truth" on which to incentivize them. In practice, while financial incentives for government workers were historically quite common, they had a tendency to lead to overzealous and unpopular bureaucrats who were perceived to abuse their positions in order to over-extract from the population (see, e.g., Parrillo 2013), which led the populace to demand less strongly incentivized civil servants.

Third, the nature of goods exchange between the state and citizens is substantially different than with the private sector. Very often, services provided by the state—like schooling and health—are heavily subsidized, thus limiting the competition the state faces from other providers. Because there is little competition from other providers, it becomes harder to base worker incentives on simple metrics like volume of services, and the lack of competitive pressure makes direct monitoring of service providers

---

[7] The United States at the turn of the nineteenth century provides a nice illustration. Prior to the Pendleton Act, federal employment was an important source of patronage. Under this system, these employees did not have tenure and turnover rates were high during changes in administration. After the passage of the Pendleton Act, which restricted the number of patronage positions, civil service reform started to take hold and federal employees began to get hired based on merit and public service exams. With these reforms, the job took on a different form. Federal employees were granted tenure and dismissals became more difficult and costly. Compensation became more formulaic which resulted in more wage compression.

more important for the public sector (relative to the private sector where competition will naturally weed out less productive firms).

Fourth, government careers differ from nongovernment careers in the mission of the organization: government organizations often aspire to public service; private sector ones to profit. This, arguably, implies that different types of individuals are potentially drawn to the public and private sector careers and personnel practices should account for this. The state may also seek different types of individuals who are able to balance the multifaceted objective functions inherent to a public service organization. (For a theoretical discussion of how matching mission organization to agent preferences can improve efficiency, see Besley and Ghatak (2005).)

Finally, a fifth key feature of the state is that it self-regulates to a much larger extent than the private sector. Monitoring structures are often embedded within the bureaucracy and there are relatively few instances of third-party private auditing of government services. Moreover, in several cases workers often switch between service delivery and monitoring roles, which potentially leads to conflicts of interest.

To quantify some of these differences between the public and private sector, we use survey micro-data from 32 countries to establish some facts about the relative compensation and fringe benefits of public and private sector workers. [8] The per capita GDP of our study countries ranges from $264 to $45,710 (constant 2005 USD$). The World Bank classifies nineteen of our sample countries as low-income or lower-middle-income. We use these surveys to document the stark difference between the public—private gaps in worker pay and worker tenure in LMICs versus wealthy countries: the public sector enjoys a large wage premium in poor countries, but only a small—or, in some cases, zero—premium in high-income countries. Public sector jobs are also more likely to provide fringe benefits like health insurance and pensions; again, the difference is much starker in poor as opposed to rich countries. These stylized facts point to a large pay premium for public sector employees in low-income countries. The premium could reflect the more complex nature of public sector jobs in low-income countries, elite capture of the public sector, efficiency wages designed to prevent corruption a la Becker and Stigler (1974), differences in job preferences across societies, or some combination of these factors. Whatever the underlying cause, this pay premium—in combination with weak information flows within and across government agencies—has important implications for how individuals select into the public sector and their subsequent performance and incentive structure.

The key point established here is that the public sector is substantially different, in terms of the level of wages, composition of the labor force, fringe benefits, and tenure. But understanding the impact of these personnel policies—and ways they can be altered to improve government performance—is challenging. Field experiments offer an attractive way of providing evidence on these outcomes, by examining what happens when these practices are altered or changed on various dimensions. To explore these issues in detail, and shed light on how to think about various aspects of personnel policy in the government sector, in the remaining sections we explore the evidence on three dimensions through which performance of government employees is determined and potentially can be improved: selection and recruitment, incentives, and monitoring. This section of the review paper is largely drawn from Finan, Olken, and Pande (2017)'s handbook chapter.

---

[8] For a more in-depth quantitative analysis of these differences between the public and private sector, please see Finan, Pande, and Olken 2017.

4.1. The selection and recruitment of public officials

The first questions we consider are: who are government employees, how are they recruited, and are there ways of improving the recruitment process? After all, individuals choose their career paths given the options available to them. In addition to the differences in compensation and job tenure documented above, a key difference between government and nongovernment careers that we identified earlier is mission differences: government organizations often aspire to public service; private sector ones to profit. On the other hand, in LMICs, corruption in the public sector may attract those who are interested in pursuing corrupt activities.

Do mission-driven organizations, such as public bureaucracies or private nonprofit organizations, attract employees with high levels of prosocial motivation? The idea is that some individuals care about benefiting others and thus feel drawn to organizations that provide them with the opportunity to do so. In general, the literature in public administration and economics supports this idea. For example, Cowley and Smith (2014) use data from the World Values Survey to measure the intrinsic motivation of public and private sector employees in 52 countries. They find that public sector workers are on average much more intrinsically motivated than private sector workers, even after adjusting for differences in basic socio-economic characteristics. Similarly, Banuri and Keefer (2013) sampled about 1,700 individuals from the government and private sectors in Indonesia and had them play a dictator game. They also find that subjects in the governmental sector are more prosocial.[9]

Besides the prosocially motivated, public bureaucracies can also attract individuals with less desirable personality traits. Organizations that offer low-powered incentives or are unable to hold their employees accountable can attract individuals with limited aspirations and a poor work ethic. Widespread corruption may attract dishonest or venal individuals. Recent evidence also supports this view. In the Cowley and Smith (2014) study for example, although public sector workers tend to be more intrinsically motivated than private sector workers on average, this difference depends on the corruption level of the country. In countries with high levels of corruption, intrinsically motivated individuals are not more likely to join the public sector. Two recent laboratory studies are consistent with this association. Hanna and Wang (2015) have students from a university in India play a series of experimental games designed to measure various personality traits, such as cheating and prosocial behavior. The authors find that dishonest students who cheated in a random dice game are more likely to express interest in a public sector job. In a corruption experiment with private sector job aspirants and aspirants of the Indian bureaucracy, Banerjee, Baul, and Rosenblat (2015) examine embezzlement of resources in which "supervisors" evaluate the performance of "workers" and then pay them. They find that aspirant bureaucrats are more corrupt than private sector aspirants, but their likelihood of being corrupt is similar across sectors.

This tradeoff in vocational profiles has also received theoretical attention regarding how best to design personnel policy. In addition to the intrinsically motivated, governments and other mission-driven organizations value individuals of high quality as well. But if higher quality candidates demand more compensation, then higher wages may be needed to attract these individuals. But as a small theoretical literature in economics (e.g., Delfgaauw and Dur (2007), Francois (2000), Prendergast (2007)) has pointed out, offering higher wages may come at the cost of attracting individuals who are more corruptible or care less about the mission. Whether this tradeoff exists is ultimately an empirical question, which has recently led several scholars to explore the extent to which financial incentives can affect a government's ability to recruit publicly motivated and high quality individuals. Below, we review a

---

[9] Other examples include: Dohmen and Falk (2010) find that German teachers are more trusting and less negatively reciprocal than employed non-teachers. Lagarde and Blaauw (2014) find in an adapted dictator game that giving to patients predicts student nurses' subsequent decisions to take rural hardship posts in South Africa. See Perry and Hondeghem (2008) for additional studies examining the role of prosocial motivation in selection in public service.

nascent experimental literature on how certain job attributes, including compensation, affect two aspects of the recruitment process into the public sector: who applies for the job and who accepts the job. We then conclude with a short discussion of the empirical evidence on whether intrinsic motivation does in fact lead to higher job performance in the public sector.

4.1.1. Wages and compensation

4.1.1.1. Effects of financial incentives on the applicant pool

If higher quality candidates, as priced by the market, demand higher compensation, then higher wages in the public sector are necessary to attract those candidates. Does offering higher wages come with the cost of attracting candidates with weaker public service motivation? This question has motivated four recent experimental studies. Dal Bó, Finan, and Rossi (2013) implemented a field experiment as part of an official program of Mexico's federal government called the Regional Development Program (RDP). The program, which sought to enhance state presence in 167 of Mexico's most marginalized municipalities, conducted a recruitment drive to hire 350 community development agents who were tasked with the responsibility of identifying areas where public good provision is deficient and working with existing public programs and local authorities to remedy such deficiencies.

A unique feature of this recruitment drive was the exogenous assignment of wage offers across recruitment sites. Two different wage offers were randomly assigned across 106 recruitment sites. In one set of recruitment sites, the program offered 5,000 pesos per month while in the other sites the program offered a wage of 3,750 pesos. Candidates who were interested in this position were then required to undertake a screening exam that was designed to measure various dimensions of quality and motivation. Dal Bó, Finan, and Rossi (2013) find that higher wages do help attract a higher quality candidate pool. In the places that announced a higher salary, the average applicant was smarter, had better personality traits, had higher earnings, and had a better occupational profile (e.g., more experience and white collar background). Moreover, contrary to theoretical concerns, these effects do not come at the cost of attracting less publicly motivated candidates, as measured by their performance on a public service motivation inventory.[10]

In another study, Ferraz and Finan (2011) analyze the impact of wages on political performance through the use of a constitutional amendment that introduced salary caps, which varied by municipal population levels, for legislators in Brazil. They find that higher remuneration attracts better candidates, though the effects are relatively modest — a 20 percent increase in wages only leads to a 0.2 increase in the legislators' average years of schooling and a 0.05 increase in the number of terms of experience. The study also shows that higher wages boost the performance of a politician while in office. A 20 percent increase in wages leads to an increase of 25 percent in the number of bills submitted. Evidence on the selection of politicians, the impact on corruption and leakages, and the effect on public welfare are important areas for future examination.

An important design feature of these studies is the ability to offer different wages for the exact same position, which, in the case of Dal Bo, Finan, and Rossi (2013), was in large part due to the

---

[10] Perry's Public Service Motivation Index (Perry 1996) is the most commonly used measure of intrinsic motivation in the public sector. This index is constructed based on a questionnaire in which the subject must express agreement or disagreement with each of thirty-two statements. The questionnaire elicits opinions on the attractiveness of politics, public service, and prosocial activities. The questionnaire is subdivided into six modules labeled "Attraction to Policy Making", "Commitment to Public Interest", "Social Justice", "Civic Duty", "Compassion", and "Self-Sacrifice". Each dimension is an average of responses to several statements that are measured on a 5-point Likert scale, where a 5 represents strong agreement with the statement, and a 1 denotes strong disagreement.

size of the program and its expansive geographical coverage. In other settings, where offering different wages has not been feasible, researchers have had to adopt alternative, creative approaches to addressing this question. One such example is a study by Ashraf, Bandiera, and Lee (2019). In this study, the authors partnered with the Government of Zambia to hire approximately 330 community health care workers. Instead of offering different wages, the authors introduced experimental variation in how the position was advertised. In 24 of the 48 districts, potential candidates saw a job advertisement that highlighted the job's promotion prospects and the opportunity for career advancement. In the other districts, applicants saw a poster that emphasized the social importance of the job. This recruitment process led to over 2,400 applicants. In the districts where the job ads stressed career incentives, applicants were much more qualified, as measured by their high-school test scores and past performance in their natural science courses. These applicants also displayed a high degree of prosocial motivation, with levels that were similar to the applicants that applied under the social incentive treatment. While the applicants who applied under the career incentives treatment did place a higher weight on career benefits, the authors conclude that making career versus social incentives salient did not induce a tradeoff between a higher quality applicant pool and a prosocially motivated one.

In contrast to these three studies, Deserranno (2015) finds that financial incentives can lead to a less socially motivated applicant pool. Her field experiment was conducted in rural villages of Uganda in collaboration with the NGO, BRAC. The recruitment drive was for health promoters, which was a position that did not previously exist and whose remuneration was uncertain since it depended on the sales of health products. The experiment exploits these two features of the position to introduce variation in how the financial aspects of the job were advertised. In one treatment arm, the job advertisement mentioned the minimum amount that health promoter was expected to earn (low-pay treatment) and, in another treatment arm, advertised the maximum amount a health promoter was expected to earn (high-pay treatment). A third treatment advertised the mean of the expected earnings distribution (medium-pay treatment).

The study finds that while the high-pay treatment attracted 30 percent more applicants relative to the low-pay treatment, the applicants had much less experience as health volunteers, and were much more likely to state "earning money" as the most important feature of the job. Applicants under the medium- and high-pay treatments were also 24 percentage points less like to make a donation to a public health NGO in the context of a dictator game. Although the author finds large effects on these various measures of intrinsic motivation, she does not find treatment effects on candidate quality, as measured by the applicant's education and income.

In sum, the conclusions from a scant experimental literature are mixed, but this is not at all surprising. Putting aside differences in the actual treatments, as Dal Bó, Finan, and Rossi (2013) point out in their model, whether financial incentives crowd out the intrinsically motivated will depend on how these personal traits (e.g., intrinsic motivation versus quality) are correlated within the broader population—correlations which the literature has yet to document in a systematic way. These four studies, which were conducted in very different contexts, represent a step forward in the literature, but without more information on how personality traits vary across broader populations it is difficult to make general conclusions about the exact tradeoffs that financial incentives induce.

4.1.1.2. *Effects of financial incentives on recruitment*

The power of wages is not limited to attracting a larger and better applicant pool. Higher wages also increase an organization's ability to fill vacancies. In the Dal Bó, Finan, and Rossi (2013) study, the authors found that the Mexican government was 35.2 percent more likely to fill the vacancy when offering the higher wage, which corresponds to a short-run labor supply elasticity of 2.15. This elasticity

is similar to other quasi-experimental estimates found in the literature (see, for example, Manning 2011), as well as one reported in Deserranno (2015). Even though the Deserranno (2015) study manipulates earning expectations (as opposed to actual earnings), the author finds an experimental elasticity of 1.8 when comparing take-up in the low-pay treatment group to take-up in the high-pay treatment group.

Part of the reason why higher wages lead to higher recruitment rates is because they help to compensate for aspects of the job that a candidate dislikes. This mechanism was on clear display in Dal Bó, Finan, and Rossi (2013). Although the applicants for the RDP position were all applying for the same job, the jobs were located in different municipalities throughout the country. At the time of the application, the candidates did not know where the job was located and were only told this information during the offer stage. As a result, jobs that were ex-ante quite similar became quite different ex-post depending on where the job was located and the characteristics of the municipality. The authors show that distance to the municipality (from their current residence) and attributes such as the level of drug violence and the lack of public goods in the municipality were all important hurdles to filling the vacancies. Fortunately, however, higher wages proved to be an effective instrument in clearing these hurdles.

4.1.2. How should governments screen?

Wage offers affect who applies for government jobs, but government jobs are typically oversubscribed, so government—like all employers—needs to winnow down the set of applicants to those they hire. Governments vary in the way they screen their public servants. Some rely on the passage of civil service exams or attainment of university degrees, while others adopt more discretionary approaches that, while permitting them more flexibility, can also be prone to corruption and patronage. These different screening strategies have important implications for not only the quality and performance of the bureaucracy, but also for the type of person who applies.

An important consideration for any organization when screening and selecting personnel is match quality. If employers and employees share a common vision and objectives, then this positive match quality increases organizational efficiency and diminishes the need for high powered incentives. For governments, who are responsible for providing public goods that are difficult to price in the market, the ability to recruit public-service motivated individuals might be especially beneficial (Besley and Ghatak 2005).[11]

Arguably, the recruitment of publicly motivated individuals has other benefits as well. Individuals with high levels of intrinsic motivation are less likely to shirk in an environment where incentives are low-powered and/or when non-contractible elements of the service provision exist (Francois 2000).

In support of these theoretical arguments, a large empirical literature in public administration shows that intrinsic motivation—and specifically public-service motivation—is associated with higher levels of performance in government work (Perry and Hondeghem 2008).[12] Recently economists have begun to contribute to this literature. For example, as part of a monitoring experiment of health clinics in the district of Punjab, Callen et al. (2015) examine the job performance of clinic doctors. They find that those

---

[11] Before we can answer the question whether or not government should screen on intrinsic motivation, we must take a step back and ask whether it even measurable and quantifiable. There is a rich and growing literature in psychology and economics that suggest that personality traits including intrinsic motivation can be measured (see, for example, Almlund et al. 2011).

[12] In the public administration literature, recent meta-studies suggest that public-service motivation is positively correlated with job performance in the public sector, broadly defined (Petrovsky 2009). Naff and Crum (1999) use a sample of over 8,000 U.S. federal employees and find that public-sector motivation correlates with individuals' last performance evaluations. Park and Rainey (2008) analyze data from 22 federal agencies in the U.S. and find that public service motivation is positively correlated with self-reported measures of job productivity and quality of work. Similar results are found using government data from Switzerland (Ritz 2009) and the Netherlands (Steijn 2008).

who scored higher on the public service motivation index are much less likely to shirk and falsify health reports. In the same study discussed above, Deserranno (2015) also finds in the case of the Ugandan health promoters that prosocial motivation is a strong predictor of job performance. Health promoters who had donated a greater share of their endowment to a local NGO visited a larger number of households, provided more prenatal checks, and organized more public presentations. Dizon-Ross, Dupas, and Robinson (2015) conduct a survey on nurses of antenatal care centers in Uganda, Ghana, and Kenya as part of an audit study on bednet distribution programs. They find that nurses not only exhibit high levels of prosocial motivation, but that it is predictive of job performance.

Despite the mounting evidence linking public-service motivation to job performance, establishing causality has proven challenging. Given the difficulties in directly generating experimental variation in a person's level of intrinsic motivation, researchers have had to rely on indirect approaches. One approach has been to introduce experimental variation in who applies for the *same* job. The experiment induces a selection effect, while keeping any potential incentive effect constant. Both the Ashraf, Bandiera, and Lee (2015) study and the Deserranno (2015) study provide examples of this approach. In the Ashraf, Bandiera, and Lee (2015) study, the authors used the two different recruitment strategies to create variation in the type of health promoters that were recruited. Once employed, all of the health workers were tasked with the same responsibilities and faced the same incentives. Based on this design, they find health workers attracted by career incentives are much more effective at delivering health services, as measured by home visits and the organization of community meetings. These health promoters were also more likely to remain in their posts over the course of eighteen months. Although these results imply a negative relationship, if any, between intrinsic motivation and performance, it is worth noting that the level of prosocialness among the health promoters who were recruited in the career incentives treatment was also quite high.

The Deserranno (2015) study provides stronger evidence in support of the relationship between prosocialness and job performance. Among the agents recruited under the low-pay treatment who were measured to be more prosocial, she finds higher aggregate performance in the first year of work. Compared to the high-pay treatment, the health promoters recruited under the low-pay treatment visited a larger number of households, organized more public presentations in the village, and provided more natal checks. She also finds that they were more likely to target the most vulnerable households.

While these studies can credibly identify the effects of "selection" on job performance, what this selection effect comprises is not entirely clear. Prosocial motivation is frequently found to be correlated with various other personal traits, including the Big Five[13]. Short of randomly assigning individuals based on a specific attribute, it is difficult to separate the effects of intrinsic motivation from other positive personality traits.

Another experimental approach that studies have explored has been to test whether a particular intervention is more effective among individuals with high levels of intrinsic motivation. For example, the goal of the experiment in Callen et al. (2015) was to reduce high levels of absenteeism among clinic doctors and staff in Punjab.[14] One source of this absenteeism was the fact that these clinics were rarely inspected, and when inspections did occur, doctors and inspectors would collude and falsify the report. In collaboration with senior health officials of the Department of Health, the authors introduced a new monitoring program in 18 of the 35 districts constituting their experimental sample. In the treatment districts, the traditional paper-based monitoring system for clinic utilization and worker absence was replaced with a smartphone application. The new system allowed health system inspectors to upload the

---

[13] Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism.

[14] Based on unannounced visits to clinics at baseline, they found that only 56 percent of clinics had been inspected in the prior two months and that 32 percent of clinics had no doctor present.

results of their assigned visit to a central dashboard which instantly updated reports at different levels of aggregation. The data, which included geo-tagged, time-stamped facility staff photos, made it difficult for the inspector to falsify his report. While the study finds that the monitoring technology did increase the number of inspections, there was significant heterogeneity in the treatment effects by the personality type of the inspector. Higher quality inspectors responded much more positively to the treatment. A Big Five index one standard deviation higher, for example, is associated with a differential 35 percentage point treatment effect in terms of health inspections.

While these studies establish that the characteristics of individuals are an important determinant of performance, and they suggest how governments could change the applicant pool, they do not necessarily tell us how governments should screen among the candidates who apply to further improve selection. Hanna and Wang (2015), for example, show that current Indian civil-service type screening exams would not eliminate the negative selection on dishonesty they find in their setting.

Ashraf, Bandiera and Lee (2016) conducted a randomized experiment in Zambia to understand the messages that will be most effective at recruiting the highest-performing public services providers. They find that recruitment drives of public health workers emphasizing civil career prospects are more effective at attracting agents with more skills and ambition than those emphasizing the "doing good" aspect of the job. Agents attracted by a career in the civil service provide more inputs (29% more household visits), increase facility utilization rates, and improve health practices, ultimately leading to improved health outcomes, with a 5-percentage point drop in in underweight under five children.

Ravanilla (2016) conducted a randomized evaluation to test the impact of an all-expense-paid three-day leadership training workshop in the Philippines on incentivizing honest and competent youth to run for a village youth council. He finds that providing conditional or unconditional nonfinancial rewards may be an effective means to screen out unmotivated youth from public sector work. Amongst youth who were found to have a low public sector motivation, those provided with a reward were 13 percentage points less likely to be assigned to take a leadership role, compared to 34 percent of youth assigned in the comparison group.

Given the large public sector premia we observe for low-income countries, an important friction in improving the human capabilities of the state may lie in its screening technologies. To understand how to most effectively screen seems an important direction for future research.

Open research questions:
- *What are the tradeoffs associated with higher wages?*
  - *Selection: Do they improve the qualifications or prosocial motivation of public sector job applicants?*
  - *Performance: Do higher wages reduce corruption and, if yes, through which channel?*
- *What are the possible selection and promotion criteria to reduce the propensity of civil servants to be corrupt?*

4.2. Using incentives to improve performance

Once selected, it may be possible to use incentives to further improve workers' performance and reduce corruption. Public sectors careers, however, typically feature a relatively flat incentive structure. As discussed above, in the standard civil service model, adopted almost universally since the early twentieth century to limit politician discretion over appointments and salaries, the public sector is staffed by salaried civil servants whose salaries are based on rigid and formulaic pay scales. These pay scales feature compressed wages relative to that in the private sector (see, for example, Borjas (2015) for the

United States). The combination of formulaic pay systems and wage compression limits the degree to which financial incentives can be used to reward the performance of public servants: the formulas are largely based on seniority and position, allowing little room for discretion, and the wage compression is such that even promotions within the civil service are less of an incentive than in the private sector.

While this type of salary structure may be appropriate for governments in many contexts, it may not be effective in LMICs where government officials are often thought to have poor job performance. Therefore, scholars have begun to examine the costs and benefits of providing additional incentives to government workers in LMICs. In this section, we review recent experimental evidence that seeks to shed light on these issues in a variety of sectors. We begin first with evidence on civil service wages and then highlight the promise of using financial incentives to reward good performance. The evidence sheds light on the degree to which such incentives can improve performance, but also highlights the challenges with using such incentives in practice, particularly those that arise when using them in the public sector.

Given the constraints in financial payments imposed by civil service systems, we then go on to consider nonfinancial incentives, which are prevalent in government contexts. We examine one type in particular that is quite common in government practice: using transfers to more or less desirable postings as an incentive device. We then examine other types of nonfinancial incentives.

4.2.1. Compensation and corruption[15]

Despite the attention often given to civil service wages, there is relatively little evidence on their impact. Several cross-country studies find that higher public wages are associated with lower corruption, though these studies are essentially cross-sectional in nature. For instance, in a cross-section of 31 low-income countries, Van Rijkenghem and Weder (2001) find that a doubling of government relative to manufacturing wages is associated with only 0.5 point reduction in ICRG corruption index measured on a scale from 0 to 6. Meanwhile, Rauch and Evans (2000) find that the level of bureaucratic wages are significant in explaining only one of the five measures of bureaucratic performance, namely that a 1 standard deviation increase in salary is associated with an improvement of 0.5 standard deviation in the bureaucratic delay index measured on a range from 1 to 4.

With regard to more micro evidence, Di Tella and Schargrodsky (2003) test the efficiency wage idea by looking at a corruption crackdown in Buenos Aires hospitals' procurement departments. They examine the impact of increasing the probability of detection and examine heterogeneous impacts on the prices paid for basic inputs based on the level of wages. They find that prices paid by hospitals for basic, homogeneous inputs decrease by 15 percent during the first 9 months of the crackdown, and following period prices increase, but remain 10 percent lower than those prevailing before the crackdown. During the first phase of the crack-down, when audit intensity can be expected to be maximal, higher wages have no effect on inducing lower input prices. Meanwhile, higher wages do have a negative effect in the last phase of the crackdown, when audit intensity can be expected to take intermediate values -the wage elasticity of input prices exceeds 0.2 (controlling for fixed effects).

Niehaus and Sukhtankar (2013) examine the theoretical idea that the rents from keeping one's job can deter corruption today in order to preserve tomorrow's opportunities. The rents they examine come from corruption, not wages. Specifically, they examine a setting where officials have two types of corruption available to them. In their setting, corruption is measured as the gap between official and actual quantities—including over-reporting of days and under-payment of wages in the Indian National Rural Employment Guarantee Act. They are able to identify the effect of future rents on the level of corruption today because the program features two types of projects; some projects pay fixed daily wages, while

---

[15] This subsection is largely based on Olken and Pande (2012).

others pay piece rates. They examine how corruption in the two types of projects varies with anticipated rent-extraction opportunities using an exogenous increase in the wage rate for daily wage projects. Their results show an 80 percent reduction in the daily theft on piece rate projects in the period post-wage increase. Hence, when the opportunities for theft from daily wage projects increase, theft on piece rate projects goes down. In addition, they find reduced over-reporting of days worked on daily wage projects in areas where the proportion of future daily wage projects is higher.

Aside from wage increases, other types of resource windfalls, such as federal transfers, could influence corruption by local officials. Brollo et al. (2013) use regression discontinuity design to study the impact of increased federal revenues transferred to Brazilian municipalities on local corruption and candidate selection. Since the level of federal transfer received by each municipality is determined by population, they use population discontinuities as an instrument for transfers actually received, looking at seven thresholds for receiving varying levels of funds and comparing municipalities just above or below these population thresholds. They find that a 10 percent increase in federal transfers to a municipality raises the incidence of broad corruption by 17 percent and narrow corruption by 24 percent, decreases the fraction of opponents with at least a college degree by 7 percent, and increases the reelection probability of incumbents by 7 percent. They hypothesize that the mechanisms behind these results are that increased funding means local politicians have more room to collect rents without being noticed by voters, and those candidates who have worse outside options are more likely to benefit from greater opportunities for corruption, thus the lower quality candidates attracted to municipalities with higher levels of corruption.

While these various studies all suggest that there might be a relationship between the future returns to employment and the amount of corruption chosen, they are by no means dispositive, and this remains an important area for future research.

4.2.2. Financial incentives

Government officials do many types of jobs, and some are easier to incentivize than others. In some cases, what we refer to as "agents of government authority" are those who are tasked with ensuring citizens comply with government laws and regulations. For such officials—such as police, judges, prosecutors, tax inspectors, building inspectors, and so on—there is a natural tension between what the government would like the agent to do (for example, to make people pay taxes that are due under the law) and what the targets of government enforcement would like the agent to do (e.g., to allow them to avoid paying taxes). This tension invites opportunities for corruption between the agent and the citizen (e.g., reducing taxes in exchange for a bribe), and, as we will see, complicates the incentive problem for the government. In other cases, such as front-line government service providers, the government and the citizen's incentives are aligned: both would like the agent (e.g., the teacher) to provide more or better services. Providing incentives may therefore be more straightforward in the second case. We consider incentives in both contexts in turn.

4.2.2.1.   Incentives for agents of government authority

4.2.2.1.1.   Incentives for tax collection

Several recent experiments explore the risks and rewards of financial incentives in the public service. Khan, Khwaja, and Olken (2015) conduct a field experiment in urban Punjab, Pakistan, to study performance pay for tax inspectors. The experiment involved high-powered financial incentives for property tax inspectors who are in charge of assessing properties, collecting property taxes, and levying sanctions on those who fail to pay. The basic treatment gave the team of tax staff in an area, which

consisted of three people, an incentive payment equal to an average of 30 percent of tax revenues collected above a historically predicted benchmark, enough to double their baseline wages. Tax inspectors are exactly the sort of government worker where one might be concerned ex-ante about the efficacy of incentives; while there is substantial scope for improvement through either increased effort or reduced corruption, incentives also have the potential to increase bribes by raising the bargaining power of tax inspectors (who now must be paid a higher bribe to compensate them for their foregone incentive payment), or to lead to over-taxation as was thought to be the case historically.

Khan, Khwaja, and Olken (2015) find evidence for both the positive and negative aspects of incentives. On the plus side, the incentives raised revenue substantially. On average, treated areas had revenue growth that was 9.4 log points greater than control, which translates to a 46 percent higher growth rate in revenue. Incentive schemes that only rewarded on revenue did best, increasing revenue growth by 12.9 log points (64 percent higher growth), whereas those that attempted to control multi-tasking problems through incentive schemes that also rewarded taxpayer satisfaction and accuracy of tax assessments had less impact on revenue, yet did not improve these dimensions. The revenue gains substantially exceeded the costs of the incentives. The incentives did not appear to reduce taxpayer satisfaction, in part because the increased tax collection was concentrated among a small number of taxpayers. On the negative side, however, they find evidence that bribe rates did increase in incentive areas, potentially to compensate incentivized tax inspectors for foregone incentive payments.[16]

4.2.2.1.2. *Incentives for policing and justice*

The Khan, Khwaja, and Olken (2015) study is relatively rare in focusing on the tax sector, where the government has a potentially adversarial role against the taxpayer, which leads to opportunities for collusion and where incentives can have perverse effects, such as the increase in bribes they document. The police force is another area with agency problems, where one might be concerned that financial incentives (e.g., on the number of citations issued, arrests made, or the like) could lead to overzealous or inaccurate enforcement, or simply a reallocation of resources from non-incentivized to incentivized tasks. Baicker and Jacobson (2007), for example, document that when police agencies in the United States are allowed to keep the revenue they obtain from assets they seize in drug arrests, they increase drug arrests, but do so by reducing enforcement of other petty crimes, suggesting that multi-tasking is an important issue. A commonly voiced concern about these laws that their paper does not address is whether these types of laws lead to unjustified seizures and abuses (Miller and Selva 1994). Other similar areas where financial incentives have been tried, but not rigorously evaluated, include incentives for prosecutors in New York City to ensure speedy disposition of cases (Church and Heumann 1989), as well as historical examples from the United States where prosecutors were paid incentives based on conviction rates (Meares 1995). Exploring the impacts of incentives in these areas in a more rigorous and careful way, and seeing whether they are effective in LMICs with more corrupt and generally less effective police forces, seems an important area for future work.

4.2.2.2. *Incentives for front-line service providers*

A more common area of focus has been incentives for health and education service providers. These have taken two broad forms, incentives on outcomes (e.g., test scores, immunizations given) and incentives on inputs (e.g., provider attendance).

4.2.2.2.1. *Incentives on outcomes: Test scores*

---

[16] Other non-experimental studies of tax agencies also find increases in revenue (e.g., Kahn, Silva, and Ziliak 2001, Burgess et al. 2010), but are unable to examine the potential downside in terms of overenforcement or bribery.

One commonly considered type of outcome-based incentive is teacher incentives based on student test scores.[17] Muralidharan and Sundararaman (2011) report the results of a large-scale teacher incentive program run by the Indian state of Andhra Pradesh as a school-level randomized trial. Public school teachers were paid incentives based on test scores, with both group and individual incentives considered. Incentives were substantially smaller than in the Khan, Khwaja, and Olken (2015) study discussed above, as they were calibrated to be around 3 percent of a typical teacher's annual salary. They find that the incentives were effective in promoting learning: after two years, students in incentivized schools had test scores that were 0.27 standard deviations higher in math and 0.17 standard deviations higher in language. They find no evidence of multi-tasking; in fact, students also do better in non-incentivized subjects, such as science and social studies. Incentives appear to have worked by increasing effort conditional on attendance, not by increasing teacher attendance. The individual incentives outperformed the group incentives by the end of the second year. Muralidharan (2012) reports that the effects increase even more with time: after five years, students in treatment schools had test scores 0.54 standard deviations higher in math and 0.35 standard deviations higher in language, and still had higher test scores in non-incentivized subjects.

Duflo et al. (2012b) show that financial incentives seem to work for teachers with good monitoring. They design a randomized experiment in which they varied the exposure to monitoring and financial incentives of teachers across 113 schools in rural Rajasthan. Teachers in treatment schools were exposed to daily monitoring (using cameras) and were remunerated based on a nonlinear function of their attendance, while teachers in the control group were neither monitored nor given financial incentives. To disentangle the effects of monitoring from financial incentives, the authors estimate a structural dynamic model of teacher labor supply using the daily attendance data in the treatment schools, and find that teachers respond to the financial incentives: estimates suggest that the elasticity of labor supply with respect to the level of the financial bonus is between 0.20 and 0.30.

On the other hand, Glewwe, Illias, and Kremer (2010) find somewhat less encouraging results. They conduct a randomized trial of teacher incentives in Kenya, where an NGO provided in-kind prizes to teachers in Kenyan government schools on the basis of school-level performance on district exams, where those who did not take the exam were imputed a low score. They find that incentivized schools had more people taking the government exam and higher scores on the government exam used for the incentives. However, unlike the Muralidharan and Sundararaman (2011) example that found positive spillovers to non-incentivized subjects, they find no evidence of higher scores on an independent exam administered by the NGO that was not linked to performance incentives. The authors conclude that multi-tasking was a real issue in their context and that teachers may have emphasized test-taking skills, as opposed to general instruction, in response to the incentives. In both cases, incentives improved targeted indicators but understanding why there were positive spillovers to non-incentivized contexts in India but not in Kenya seems an important area for future research.

While the above literature discussed financial incentive for teachers, Behrman et al. (2015) conduct a randomized evaluation in Mexico to study the impact of different performance incentive schemes. The authors randomly assigned 88 high schools with over 40,000 students into three treatment groups or a comparison group. The treatment groups received incentives based on students' performance in math for either students, teachers, or both, and schools in the comparison group received no incentives. The incentives were designed to prevent teachers from focusing only on top students. At any given level, the monetary incentive for teachers for when their weaker students reached a specific level was higher than when their stronger students reached the same level. The authors find that giving incentives to both

---

[17] In addition to the experimental studies reviewed here, there is also a large literature on teacher performance pay in the United States. See Neal (2011) for a review.

teachers and students yields the highest impact, leading to 0.3 to 0.6 standard deviation increase in math performance. Giving incentives to only students improves math performance by 0.2 to 0.3 standard deviation. In contrast, provision of monetary incentives to only teachers has no significant impact on student math performance. One caveat to note is that since the test is 'low-stakes' for students in comparison schools, their test-taking effort might differ with that of students in treatment schools.

4.2.2.2.2.   *Incentives on outcomes: Health*

While incentives based directly on health outcomes are rare, one notable example is Miller et al. (2012). The researchers conducted a randomized trial in 72 Chinese primary schools in which school principals received performance payments based on reduction in anemia among their students. Specifically, principals were paid 150 RMB per student who changed from anemic to non-anemic over the course of the intervention. This implied a payment of roughly two months' salary for reducing anemia by half. Comparison groups were given the same information and subsidies as the incentive treatment, but no direct financial incentives. They find that the incentives reduced anemia compared with the pure control group by about 5 percentage points (23 percent). The non-incentivized comparison groups did not achieve statistically detectable reductions in anemia, suggesting a role for incentives, but confidence intervals are such that they cannot statistically distinguish between the incentive group and the non-incentive information and subsidy groups. They report that incentivized school principals were more likely to use subsidies for iron-focused supplements, whereas non-incentivized school principals used subsidies for supplements that could affect both iron and overall calorie intake. Depending on one's perspective, this could be considered a multi-tasking issue as well to the extent that one is interested in both types of supplementation.

4.2.2.2.3.   *Incentives on service delivery*

In health, incentives for providers have tended to focused instead on measures of service delivery, such as the number of immunizations given. These service delivery metrics can be thought of as somewhere between ultimate outcomes (e.g., learning, lack of disease) and provider inputs (e.g., attendance). One reason for focusing on this level is that for ultimate health outcomes, the signal to noise ratio may be high; that is, in a given context, most of the variance in health outcomes is idiosyncratic rather than due to provider effort. If one believes there is a clear mapping from health service delivery to health, these types of incentives may make sense.

Basinga et al. (2011) and Gertler and Vermeersch (2012) examine this approach. The intervention they study in their experiment took place in Rwanda and provided incentives to primary care facilities, which were in turn used to compensate facility personnel. The incentives were based on the quantity of visits to the facility for various services (e.g., childbirth in the facility, prenatal care) and the content of services provided in those visits (e.g., pregnant women receiving tetanus vaccines and malaria prophylaxis during prenatal care, immunizations given during postnatal care, etc.), weighted by an overall quality index of the facility. The incentive payments were substantial, equal to a 38 percent increase in total compensation for facility personnel. They found that the incentives led to a substantial increase in pre- and postnatal services, which translated into increased health: infant weight-for-age increased by 0.53 standard deviations and height-for-age for children age two to five by 0.25 standard deviations, with increased breastfeeding and reductions in infant illness hypothesized to be important channels.

Celhay et.al (2015) also evaluated an impermanent incentive scheme that paid clinics to perform early pre-natal care in Argentina. In the study, the clinics were authorized to allocate up to half of the incentives received for employees' bonuses. The authors found positive treatment effect; clinics randomly

assigned to the monetary incentive treatment delivered initiation of early pre-natal care at a rate 34 percent greater than that of the control clinics, and the effect lasted for a minimum of 15 months after the intervention concluded. However, the authors did not observe any impact of the incentive scheme on health, as measured by birth weight and number of low and premature births.

In contrast, Huillery's and Seban's (2016) study in the Democratic Republic of Congo shows that providing performance-based incentives for healthcare service providers can have negative impacts. The authors evaluated the effects of performance-based group payment that awarded incentives to health workers based on the number of services performed in health clinics. Workers in the treatment group received financial incentives based on output, whereas the control group received a fixed salary. Results show that while the performance-based incentives led to greater effort by the workers, it had negative impacts on facility revenue, service utilization, and child health. The average weight and height of newborn children in the treatment groups was 0.17 standard deviation lower than in the control group and the average number of prenatal visits in the treatment group was 0.4 standard deviations lower than in the control group. Thus, while workers in the treatment group provided more effort, they evidently misplaced this effort towards ineffective strategies to attract more patients. Workers in the treatment group responded to the incentive scheme by reducing costs but did nothing to enhance service quality or build awareness about the benefits of health clinics, two key barriers to service take-up in this setting. This study shows that performance-based rewards need to be carefully implemented in an environment where workers' capabilities are limited and demand for healthcare services is complex.

Olken, Onishi, and Wong (2014) report the results of a large-scale field experiment in Indonesia in which villages were provided with incentive payments based on health service delivery (similar to that in Rwanda), school enrollment, and education. Specifically, villages received a block grant each year that they could use for any purpose related to health or education. In incentivized areas, 20 percent of the total amount set aside for block grants in a subdistrict was allocated to villages based on their performance on the targeted health and education indicators; in non-incentivized areas, the block grant was allocated based only on population. The incentive was to the community as a whole, and unlike the previous examples, was generally not passed on to service providers, but was instead used for programming (e.g., nutritional supplements, subsidies for childbirth, etc.). Comparing the incentivized to non-incentivized areas, they found the incentivized areas performed better on the targeted health indicators. On average, the eight targeted health indicators were about 0.04 standard deviations higher in the incentivized than non-incentivized areas. These effects were about twice as large in areas with low initial levels of performance, but the relative gain of incentivized to non-incentivized areas declined over time as non-incentivized areas improved. The main health reduction was a 15 percent (2.6 percentage points) decline in malnutrition rates, though again this effect became more muted over time. There were no detectable differences between incentivized and non-incentivized areas on educational outcomes. The program suggests that the incentives sped up improvements on the targeted health outcomes, with no detectable multi-tasking effects.

In their study in Pakistan, Andreoni, et al. (2016) explore the role of individual time preferences in designing performance-based incentives. They conducted a two-phased randomized evaluation to examine the effects of personalized bonus contracts or "tailored contracts" on health workers' performance in administering vaccines across a two-day vaccination drive. In the first stage of the study, the authors measured health workers' time preferences and patience based on how they allocated vaccinations over time. Workers were asked to set targets for the number of vaccinations they would complete in each day and, if they reached their target on both days, they would receive a fixed bonus of around USD$10. Before making this decision, health workers were randomly assigned an "interest rate" ranging from 0.9 to 1.25. For each vaccination allocated to day two, the number of vaccinations required on day one was reduced by the interest rate. A higher interest rate made it more attractive to allocate more vaccines to the second day. The authors then used the workers' daily targets to estimate time preferences for each worker to create

personalized bonus contracts. In the second stage, the authors randomly assigned personalized bonus contracts to health workers. The results indicate that workers assigned to personalized bonus contracts provided significantly smoother service and were one third closer to achieving their targets compared to workers who received random interest rate contracts.

4.2.2.2.4.   *Incentives on inputs: Provider attendance*

The final category of financial incentives that we consider is incentives based on attendance. Given the problems with provider attendance highlighted by Chaudhury et al. (2006), this is clearly an important issue—but key questions are whether attendance responds to financial incentives, and, if so, if it translates into ultimate outcomes. Here the evidence is mixed. In an experimental study for teacher attendance at single-teacher schools run by NGOs in India, Duflo, Hanna, and Ryan (2012) provided linear incentives based on the number of days (above ten per month) that teachers could submit time-stamped photos of themselves with students to prove they had been attending. They found that the incentives not only increased attendance, but also led to increased learning, with students in schools where teachers were incentivized to attend having test scores about 0.17 standard deviations higher than in control schools after one year.

On the other hand, a similar study by Banerjee, Glennerster, and Duflo (2008) of financial incentives for nurses' attendance in India provides a more cautionary note. In that experiment, the incentives for attendance were broadly similar to those in Duflo, Hanna, and Ryan (2012): nurses who were recorded absent more than 50 percent of the days in a month would have their pay reduced by the number of days they were recorded absent, and nurses who were absent more than 50 percent of the days in two consecutive months would be suspended from government service. Nurses used a protected time/date stamp machine to verify attendance. In their study, while there was initially a substantial treatment effect, the effect diminished over time and was zero at the end of their study. Although they do not have the data to confirm this, anecdotal evidence suggests the decline was due to nurses learning how to exploit loopholes in the systems and recording more exempt absences over time. One possible difference is that the Duflo, Hanna, and Ryan (2012) schools were run by an NGO, which may have had more independence in enforcing the incentives than the government. A key question then is understanding which of these effects is more likely to generalize: the positive long-run effects found in Duflo, Hanna, and Ryan (2012) or the rapid decline in effectiveness found in Banerjee, Glennerster, and Duflo (2008).

4.2.2.2.5.   *Unconditional financial incentives*

While previous literature provides evidence that providing financial incentive based on test-scores lead to improvement in students' performance, De Ree et.al (2015) conducted a large-scale randomized evaluation in Indonesia with more than 3,000 teachers and 80,000 students to study the effects of unconditional increases of teachers' salaries. In an effort to improve teacher quality and learning outcomes, the government of Indonesia enacted a national teacher certification program in 2005. Teachers who successfully completed this certification had their salaries doubled. In the study, all eligible teachers in 120 randomly-selected schools were granted immediate access to the certification process and the resulting doubling of pay, while teachers in control schools experienced status quo access to the certification process through a gradual phase in over time. The results show that this reform benefitted teachers; doubling salaries increased teacher satisfaction towards their salary and reduced the number of teachers who had a second job. In contrast, this policy had no impact on students' test scores or teachers' effort. The rationale behind this policy reform is the widely-accepted assumption that the cause of teachers' poor performance is their low pay. This study provided evidence that providing unconditional increases in remuneration might not be the most cost-effective way to improve service delivery.

4.2.3. Nonfinancial incentives

While the majority of research has focused on financial incentives, non-pecuniary incentives are potentially important. While civil service regimes typically place many restrictions on hiring and firing, they have much more flexibility in assigning bureaucrats to postings within the civil service and these postings can be used as reward and punishment devices. Many bureaucracies informally recognize high achievers (i.e., "employee of the month" type awards). Public sector jobs in particular may seek to take advantage of the fact that their employees may be public-spirited, and use this as a way of creating rewards. While much less extensively studied than pecuniary incentives, several studies suggest that it is a promising direction for further exploration.

4.2.3.1.   Transfers and postings

Civil service regimes typically feature much more flexibility in where people are posted than in whether people are fired or how much they are paid. This is perhaps natural, in that there are a wide variety of positions that need to be filled and these positions are heterogeneous in many dimensions, both in terms of the skills needed to complete the job effectively and their desirability as a place to work. In both cases, there can be a mix of common and idiosyncratic rankings. For example, as we discussed in the case of Mexico RDP program, civil servants were much more willing to work in a safe community, as opposed to one with high incidences of drug violence; this would be a common preference. The civil servants also preferred to work near their place of residence; since people are from different places, this creates idiosyncratic preferences. The same can be said for job attributes: a common attribute would be the need to put the cleverest tax inspectors on the most complicated corporate tax cases; an idiosyncratic attribute would be the need to match police with areas where they have social connections they can use to gather information. To the extent that there are common components to preferences, this creates scarcity for posting in the most desirable locations, and such plum postings can be used as an incentive device.

One problem with transfers as an incentive device is that politics often gets in the way. Iyer and Mani (2012) examined a comprehensive dataset that tracks the careers of elite Indian Administrative Service personnel. They show that transfers are likely right after a new Chief Minister is elected, particularly for those bureaucrats who were not at the very top of their initial class in terms of performance. However, even though ability is predictive of future success, caste affinity to the politician also plays an important role. Bureaucrats who share the same caste as the chief minister's party are just as likely to be assigned to important posts as the high ability bureaucrats. These results, while not definitive, suggest that while transfers are quite common, they are not entirely based on performance, which may dampen their usefulness as a performance tool.

Banerjee et al. (2014) explore several aspects of these issues in the context of the police force in Rajasthan, India, a context in which transfers are frequent: one third of all policemen were transferred during a typical 18-month period. As in the elite civil servants examined by Iyer and Mani (2012), anecdotal evidence suggests that police transfers are frequently imposed by politicians, often for reasons that may reduce their use as incentive device (e.g., for partisan or corrupt motives). They explore a treatment where all transfers were frozen during a two-year period, except for well-documented cases of police misconduct. The idea was to remove arbitrary transfers and leave only transfers being used as an incentive device. They find that this freeze had no effect on outcomes such as whether decoy surveyors were treated differently or community satisfaction. One potential reason for the lack of effect is that, if the exceptions were sufficiently difficult to implement, the freeze could have eliminated both transfers used

as incentives as well as politically motivated transfers. The elimination of the transfer-as-incentive could have then offset the positive effect of removing political transfers.

Consistent with this idea, a second treatment suggests that transfers have the potential to be used as an incentive device. In a second experiment, Banerjee et al. (2017) examined an anti–drunk-driving reform, where police were supposed to conduct sobriety effects. Two groups of police ran sobriety checks. In the first group, they worked with police in the central reserve 'police lines' group, who are outside of typical station assignments and was given the incentive that they would be transferred back to the regular police unit if they performed well. The second group consisted of police from normal stations, who had no transfer incentives. They randomized which units were sent to which areas and found that those in the first group performed better in terms of whether the roadblock actually occurred, the number of people stopped, and so on. Of course, the composition of personnel was different in the two groups, so one cannot know if it is the transfer incentive per se that is driving the results or some other factor (for example, maybe the police lines teams had nothing else to do with their time, while the regular police teams were juggling many other tasks), but the results are suggestive that this could be important.

Khan, Khwaja and Olken (2016) conducted a randomized evaluation to study the effect of performance-based transfers among property tax inspectors in Punjab, Pakistan. The tax inspectors in the treatment group were informed that they would be ranked based on performance and the top performers would get their first posting preference, the second-rank inspectors would get their first preference from the remaining postings, and so on. The authors find that the treatment of performance-based transfers substantially raised revenues, by 44 percent in the first year and 80 percent in the second year. This effect is of the same magnitude as the performance-pay scheme (Khan, Khwaja and Olken, 2014) but is much more cost effective since applying the performance-based scheme does not incur any cost for the government. The authors conclude that there are substantial gains to be made to performance of bureaucracy by using transfers as an incentive. However, the authors also note that the effects disappear if an inspector is exposed to the scheme for two consecutive years—suggesting that this scheme can only be applied periodically.

4.2.3.2.   *Intrinsic motivation*

Beyond explicit incentives, it may be possible to use other types of intrinsic rewards as a motivational tool. One experimental study that examines this idea is Ashraf, Bandiera, and Jack (2014). In their study, public health extension workers in Zambia who are tasked with selling condoms are randomly assigned to either different financial rewards (margins of 10 percent or 90 percent on each condom sale) or a nonfinancial reward that gives agents additional stars for each sale on a thermometer-type display. They find remarkable evidence of the effectiveness of nonfinancial rewards: the thermometer treatment agents sell twice as much as those in the financial rewards treatment, and the effect for both nonfinancial and financial treatments is greater for "prosocially motivated" sellers. The authors suggest that the "star" treatment is effective partially because it creates a comparison between sellers.

Another recent, nonrandomized study also suggests that dimensions other than incentives may be important to job performance. Rasul and Rogger (2015) use a survey to measure management practices in the Nigerian bureaucracy and find that autonomy is positively correlated with job performance, whereas performance incentives are negatively correlated. Of course, there could be endogeneity problems: one might choose to give performance incentives to those bureaucrats who behave badly and to reward high performers with autonomy.

4.2.3.3.   In-kind transfers

Whereas the abovementioned studies highlight the promise of incentives, one issue is that programs designed to incentivize good behavior by certain actors can sometimes incentivize corrupt behavior by others. Linden and Shastry (2012) study the implementation of a conditional school nutrition program in India in which students receive 3kg of uncooked rice for every month of school in which they have at least 80 percent attendance. The idea behind the program is to both improve student nutrition as well as incentivize school attendance. Whether or not a student receives the transfer is completely dependent on the attendance records kept by the teacher of each classroom. When comparing attendance records of the teachers to attendance records taken once a week by external monitors, however, the authors find evidence that teachers use their discretion to inflate the attendance of certain students to put them over the 80 percent threshold. Of the 15,000 students in the sample, at least 40 percent received grain at least one month during the sample period in which they technically had not met the attendance requirements. Teachers were more likely to inflate the attendance records of girls, lower caste students, and high performing students but less likely to inflate for Muslim students. Patterns of attendance inflation suggest that teachers were doing so specifically in response to the nutrition program as they were more likely to inflate for children just below the 80 percent threshold and more likely to inflate at the end of the month when they could tell who needed the boost. Interviews with teachers suggest their actions were motivated by compassion in response to the difficult situation of extreme poverty and that they were not seeking benefits for themselves. This example suggests that corruption does not always come from malevolent intent nor does it necessarily decrease efficiency. As the authors suggest, the behavior of the teachers, as it was likely to be based on greater information about individual students, may have increased the efficiency of targeting students in need for the nutritional component of the program than if the implementation of the transfer had not been decentralized. At the same time, the actions of the teachers likely distorted incentives to attend school that were built into the program and may have been discriminatory in nature.

4.2.3.4.   Contract employment

Aside from looking at outcome-based incentives for teachers to boost students' learning outcomes, Muralidharan and Sundararaman (2013) also studied the effect of hiring contract teachers in Andhra Pradesh public schools. In this context, contract teachers are paid much less and possess lower qualifications relative to regular teachers. Their findings indicate that schools with contract teachers outperformed the comparison schools by 0.16 standard deviations in language and 0.15 standard deviations in math. Contract teachers also showed more effort: they had higher attendance and conducted more teaching activities in class. In this context, contract teachers are rarely hired as regular teachers. Therefore, it is unlikely that contract teachers applied more effort in the hopes of securing a permanent job. The authors also found that reducing student teacher ratios by 10 percent with contract teachers lead to an increase in average test scores by 0.03 standard deviations per annum. Similarly, decreasing the student teacher ratio with regular teachers lead to 0.02 standard deviations increase in overall test scores. The authors conclude that contract teachers are at least as effective as regular teachers in improving learning outcomes, even though their remuneration is one-fifth that of regular teachers.

Duflo et al. (2015) studied the impact of an extra-teacher program in Kenya in combination with a governance program that aimed to strengthen parents' roles in school committees. There are three important findings from this study. They found no significant improvement for students assigned to civil-service teachers, despite a sharp decline in the student-teacher ratio. In comparison, scores increased for students who were assigned to be taught by locally-hired contract teachers by 0.244 standard deviations. One reason may be that civil-service teachers responded to the policy by reducing their effort and hiring their relatives for contract teacher positions. Researchers also found that the intervention aimed at improving school

governance did lessen the negative effects caused by these two responses. Similar to the findings in Muralidharan and Sudararaman (2013), contract teachers in this setting have higher attendance than civil-service teachers. However, unlike the context of India in the Muralidharan and Sundararaman (2013) study, contract teachers in Kenya have formal education in teaching and could be hired as civil-servant teachers. Therefore, their efforts could stem from the prospect of permanent employment, although the authors could not attribute the increase in test scores to pure incentive effects in this study. Indeed, after the two-year period ends, contract teachers with higher performing students are more likely to be employed as civil servant teachers. This indicates that the program was able to identify and select effective teachers. The authors estimate that promoting contract teachers that performed well to civil-servant teaching positions could lead to a substantial increase in students' test scores, although this depends in part on the duration for which teachers remain on the job.

4.2.4. Summary

Several themes emerge from the evidence reviewed in this section. First, there is robust evidence that financial incentives matter: across a wide variety of settings, financial incentives in a government context seem to increase performance on the incentivized dimension. This is not particularly surprising. In fact, given these robust results, the question becomes why do governments not use financial incentives more often? Part of the reason can be attributed to the simple fact that unlike in the private sector where firms can contract on profits, performance in the public sector can be hard to measure. Also, as we discussed, there is some evidence that multi-tasking issues can be a problem—yet we do not have a clear understanding of when multi-tasking issues will or will not be present. Finally, as Benabou and Tirole (2006) have highlighted, financial incentives may even reduce effort among the prosocially motivated. While a number of studies have documented such adverse effects in such activities as volunteer work (e.g., Gneezy and Rustichini 2000) or blood donation (e.g., Mellstrom and Johannesson 2008), it remains to be seen whether this is a first-order issue in the context of governments. Developing a clearer understanding of when these issues will be present seems an important direction for future research. Finally, the bulk of the work on financial incentives has focused on front-line service providers where the incentives of the agent's principal (e.g., bureaucrats at the central government ministry) and the citizens the agent serves are aligned. With just a few exceptions, there has been much less work on the more complex case where these interests are unaligned, such as tax, police, procurement, and so on. Understanding the degree to which incentives can be effective in this context without further empowering these officials to collect more bribes or overenforce the law seems an important area for future work.

Open research questions:
- *When are those in charge of enforcing regulations (auditors, police, prosecutors, anticorruption commissions, etc.) effective at reducing corruption, and when do they themselves become corrupt and only add to the problem?*
- *How can incentives be effectively designed for non-frontline service providers, like tax, police, or procurement officials?*
- *Can output-based incentives for government officials such as police and tax inspectors reduce corruption, or will they lead to overenforcement and extortion?*
- *Can outcome-based incentives improve quality of service delivery? Does a narrow focus on incentivized outcomes reduce performance on other dimensions?*
- *Can nonfinancial motivations, such as shame, intrinsic motivation, and mechanisms to internalize the greater good reduce corruption?*

4.3. Monitoring Mechanisms and Public Service Delivery

Incentives focus on tying rewards—typically financial rewards—to easily observable and verifiable measures of performance. Taxes is a canonical example: the state easily observes the amount of taxes collected by each tax inspector, and so can base rewards on that. In many cases, however, monitoring performance itself requires costly effort on the part of either state or non-state actors. We now turn to whether improved monitoring can improve the performance of civil servants.

Increased monitoring can improve program performance via multiple channels. First, in cases where outcomes are not observed without some effort, increased monitoring can allow managers to directly enforce punishments and rewards based on program outcomes (e.g., firing or transferring poor performers). Second, monitoring can play an important deterrence role. Third, access to monitoring results can empower citizens to demand and obtain better services by threatening to report on or vote out poor performers.

However, there are also reasons to believe that information alone may not suffice. First, in situations where state capacity is weak, managers' or regulators' ability to impose punishments is limited, and improving information flows may do relatively little by itself. Second, those in charge of collecting information or monitoring based on available information may themselves be susceptible to corruption and misuse this information. One may worry that allowing discretion to managers in collecting and using information may have the perverse effect of increasing rather than reducing program leakage. Thus, a key dimension of heterogeneity surrounding the role of information will be the extent to which those who receive the information—be they supervisors or workers—have the incentives and ability to act on it.

4.3.1. Information Flows and Monitoring

Information on project and intermediaries' performance arises in multiple ways. The classic method remains via government auditing and inspection units that are required to monitor government programs. More recently, the rise of e-governance has meant that government agencies have access to large administrative datasets on funds flow, intermediaries' behavior (typically attendance), and monitored program outcomes. These data directly allow managers to obtain better real time data on program performance and, in many cases, public availability of these data (aided in part by the rise of freedom of information acts) increases citizen monitoring.

In addition, the ease of access to information at different levels of a bureaucracy can also play a key role in ensuring that government programs monitoring takes place. Exploiting the randomized rollout of a mobile app in India that reduced search costs for identifying the agents who delay the processing of wages for a workfare program, Dodge et al. (2017) find that lowering the costs of information acquisition reduce payment processing time by up to 21 percent. The impact of the program was stronger when information was made available at both the intermediate and senior levels.

Below we first discuss findings from experiments that evaluate government monitoring processes and then we turn to citizen monitoring.

4.3.2. Government Monitoring

4.3.2.1. Does more information on performance improve outcomes?

4.3.2.1.1.   Studies of audits

Several studies examine the role of government audits. Olken (2007) conducted a field experiment in Indonesia where a local village body implemented a road construction program and audits were conducted by the government agency. The key finding is that audits have a significant deterrence impact. Before villages began building road projects, some were randomly selected for a high audit intensity group, where they faced an audit by the government agency with 100 percent probability as opposed to a 4 percent probability in the control group. Olken (2007) found substantial effects of the government audits, reducing unaccounted for expenditures by about 8 percentage points or about 30 percent from the baseline level.

In the Ferraz and Finan (2008) study discussed in section 2.2.3.3.2, the authors examined audits of municipal accounts in Brazil where small municipalities were randomly chosen to be audited by government auditors. They examine the impact of the timing of auditing on the probability that the mayor is re-elected. They find that, conditioning on the actual number of corruption violations found by the auditors, those audited before the election were less likely to be reelected than those who were audited after the election for those with an intermediate number of violations.

In a follow-up paper, Ferraz and Finan (2010) use the same audit report from their 2008 study for municipalities that were audited prior to the 2004 municipal elections to test whether electoral accountability affects the corruption practices of incumbent politicians. They compare mayors serving in a first term (eligible for re-election) to mayors in their second term (non-eligible for re-election) to identify the effects of re-election as an incentive to avoid corruption. Using the share of total federal resources transferred to municipalities that are associated with fraud in the public procurement of goods and services, diversion of funds, and over-invoicing of goods and services as a measure of corruption, they find that mayors with re-election incentives are significantly less corrupt than mayors without re-election incentives. In terms of magnitudes, they estimate that in municipalities where mayors are in their first term the share of stolen resources is, on average, 27 percent lower than in municipalities where mayors were non-eligible for re-election. Although this result suggests a two-term period is more effective than a one-term period as an anticorruption policy, it does not mean politicians should be re-elected indefinitely. Term limits could also produce benefits if politicians in the absence of the pressure of being re-elected have better incentives to implement socially optimal policies with a long-term horizon.

Gerardino et al. (2014) provide evidence that public entities respond to a temporary drop in audit risk following an audit by temporarily decreasing transparency in the procurement process, which also provides evidence that increased audit risk reduces corruption. The authors identify the effects of being audited on public entities' subsequent procurement practices in Chile. They find that less transparent procurement modalities were used during and for several months after an audit, in response to the reduced risk of being audited in the near future. The study uses a fuzzy regression discontinuity design based on a scoring rule used to identify public entities for audits to determine the effects of the audits on medium-risk entities above and below the threshold in 2011 and 2012. During the audit, use of public auctions, a more transparent modality, decreases by 14 percentage points and use of direct negotiation, a less transparent modality, increases by 10 percentage points. This shift towards less transparent modalities persists for several months after an audit but disappears completely by the subsequent fiscal year.

An open question is whether a higher likelihood of punishment also had a deterrence effect in this setting. Bobonis, Fuertes, and Schwabe (2016) examine this question in the context of Puerto Rico. Puerto Rico has established an independent body that systematically conducts municipal government audits, the findings of which are made publicly available and disseminated to media sources. Bobonis, Fuertes, and Schwabe (2016) exploit two features of the audit process. First, municipalities are audited in a pre-

established order, making the timing of audits and their dissemination pre-determined. Second, audits are "timely audits," such that reports released in the period leading up to an election are more likely to inform on the incumbent mayor's activities than those reports published shortly after an election due to a high independent turnover rate of politicians. They find that timely audits induce a significant short-term reduction in municipal corruption levels of approximately 67 percent, as well as an increase in incumbent mayors' electoral accountability. However, in contrast to these desirable short-run consequences of the audits, municipal corruption levels in the subsequent round of audits are, on average, the same in municipalities audited preceding the previous election and those whose audits became publicly available afterwards. They also find that incumbent re-election rates in the subsequent election are significantly higher in municipalities in which there was an earlier timely audit. The presence of selection effects in future re-election rates, but not in corruption, is prima facie evidence in favor of the view that the information contained in the audits helps voters select competent but opportunistic politicians, rather than honest or virtuous ones.

Borcan et al. (2015) examine the role of monitoring and the threat of punishment on performance. In an effort to curb rampant corruption in the public education system, the government of Romania introduced an anticorruption campaign with two elements: exam monitoring using closed-circuit TV (CCTV) and threat of punishment for corruption behavior—both teachers and students involved in any form of cheating or bribery in a high-stakes high school exam, or the "baccalaureate exam," would face serious repercussions. The authors used differences-in-differences to estimate the impact of these policies and found that while both unmonitored and monitored areas displayed lower test scores after implementation, the latter showed a larger decline. The results show that CCTV monitoring, combined with the threat of being punished, lead to a 0.31 standard deviation drop in overall exam scores. Since the "punishment" element of the program was being rolled out at the same time in the whole country, the authors could not estimate its individual effect. Instead, they used a low-stakes oral exam as a placebo. Overall, they find that the threat of punishment reduced corrupt practices and that monitoring enhanced the effectiveness of punishment.

There is, however, the potential for monitoring to backfire. Lichand et al. (2015) studies the introduction of the municipal audits as in Ferraz and Finan (2008) in a differences-in-differences framework. They show that in municipalities that expected they might be audited, procurement went down, with negative consequences for health. To the extent that the problems we observe are due to incompetence or laziness rather than corruption, as suggested by Bandiera et al. (2009), too much of a focus on corruption could backfire.

Aside from evaluating the impact of audits on performance, audits can be utilized as a tool to compare service quality across sectors. Das et. al (2016) used audits to gain insights on the quality of public versus private healthcare providers in rural Madhya Pradesh, India. They sent fake "patients", people who had been trained to show specific symptoms, to both public and private clinics. In order to single out the effects of practicing at private clinics, they also identified doctors with medical degrees who practiced in both public and private clinics and sent fake patients to the same doctor in both practices. The results indicate that private healthcare providers, while often lacking formal medical training, allocated more time to patients and completed more essential checklist items for diagnosing patients compared to public providers. Public health care providers also did not perform better than private providers in terms of diagnosis and treatment;[18] private healthcare providers were equally likely to provide a correct treatment. Interestingly, the study shows that doctors that work in both public and privates practice performed better when they treated patients in private clinics, as measured by time allocated to patients, likelihood of giving appropriate treatments, and number of completed checklist items. In addition, in the private sector there was a positive correlation between prices charged and

---

[18] Measured by likelihood to give correct diagnosis, correct treatments, and unnecessary treatments.

provider effort, but there was also a positive correlation with unnecessary treatment. We regard continued explorations of this issue as an important area for future work.

4.3.2.1.2.   Mobile monitoring and biometric/time stamp studies

As e-governance becomes more widespread, countries have increased their investment in, and use of, e-monitoring systems. Several studies described in our discussion of incentives based on provider attendance (Section 4.2.2.2.4) utilize such systems. Duflo et al. (2012) based incentives on information obtained via time-stamped photographs. However, as Banerjee et al. (2008) study highlights, the robustness of such monitoring systems is sensitive to how tamper-proof the monitoring mechanism is. If nurses can destroy the monitoring system (here, by literally breaking the time stamping mechanism), then they will do so. More importantly, if the incentive system allows for loopholes, then improved information may do little—a general lesson that exists above and beyond the nature of monitoring mechanism. It is worth noting that monitoring alone, without incentives, can lead to positive results in some cases. An experimental study by Aker and Ksoll (2015) indicates that regular phone calls to teachers, students, and village head as a form of monitoring lead to an improvement in learning outcomes. The treatment effect of this mobile monitoring for an adult education program in Niger is estimated to be between 0.14 and 0.3 standard deviations in reading and between 0.08 and 0.15 standard deviations in math.

Recent studies expand our understanding of the issues at stake. Dhaliwal and Hanna (2014) studied the rollout of biometric monitoring of the staff at primary health centers in South India. Health worker attendance increased by 14.7 percent in clinics with improved monitoring and was driven by lower-level staff in these centers (nurses and pharmacists), rather than by doctors. Improved monitoring had important health impacts: there was a 16 percent increase in the delivery of infants by a doctor and a 26 percent reduction in the likelihood that infants are born under 2500 grams. However, they also find lower staff satisfaction and widespread attempts by the staff to circumvent the system. Taken together, the results show both how improved monitoring can improve service delivery and also the rents at stake in the system. To the extent that staff dissatisfaction and delays in implementation tend to be more visible, this study also points to the importance of measuring impacts. In the absence of careful measurement of health outcomes, it would have been easy to focus on the partial implementation and to deem the system a failure. Similarly, Dal Bo et al. (2018) find that introducing a new monitoring technology aimed at improving the performance of agricultural extension agents in rural Paraguay increases the share of farmers visited by 6 percentage points (a 22 percent increase), without compromising the quality of the visit or the farmers' satisfaction. They add that the allocation of the monitoring technology can be complemented by information held by the principal who is able to identify the agents who would be more responsive to such monitoring.

If poor monitoring opens the door for shirking and potentially outright corruption by service providers, then we may expect these impacts to vary by the personality of the service provider. As discussed above, Callen et al. (2015) reports on a monitoring experiment where the traditional paper-based monitoring system for clinic utilization, resource availability, and worker absence was replaced by an Android smartphone application. In the new system, data generated by health inspections are transmitted to a central database using a mobile data connection in real time. Data are then aggregated and summary statistics, charts, and graphs are presented in a format designed in collaboration with senior health officials to effectively communicate information about health facility performance. Especially relevant given the Dhaliwal and Hanna (2014) study, these authors find that senior health inspectors who score higher on the Big Five personality inventory are more likely to respond to a report of an underperforming facility by compelling better subsequent staff attendance. More surprisingly, they also find that inspectors who score higher on personality tests are more likely to reduce absenteeism when

dashboards are implemented. This paper provides one way into understanding the observed heterogeneity in responsiveness to better monitoring individual personality characteristics.

In a related study, Callen et al. (2016) find that low political competition and patronage reduces the effectiveness of smartphone absence monitoring technology in Pakistan. Results suggest that persistence of absenteeism is linked to systems of political patronage, where politicians reward their supporters with in-kind benefits such as public sector jobs. The authors find that politicians routinely interfere when bureaucrats sanction doctors for absenteeism and that doctors are more absent when they are connected to politicians or when they are located in uncompetitive constituencies. The effects of a smartphone monitoring technology that nearly doubles inspection rates and an intervention that provides real time information on doctor absence to policymakers are attenuated in uncompetitive constituencies and for politically connected doctors.

4.3.2.2.   *Who collects information and does that matter?*

A common concern in the literature is the veracity of information collected by monitors. For instance, both Banerjee et al. (2008) and Dhaliwal and Hanna (2014) report on how service providers seek to reduce the functionality of monitoring devices.

Incentives to provide poor quality information may also arise if inspectors and auditors are corruptible by those they are intended to monitor. This possibility is, arguably, particularly stark in the case of private sector auditors paid by the firms or institutions they audit. Duflo et al. (2013) examine the implications of corrupted information flows for regulatory efficacy. In a large field experiment conducted with the environmental regulator in Gujarat, they altered the assignment and payment mechanism for third-party environmental auditors of industrial plants. Under the status quo, the auditors were hired and paid for by the plant they audited. In the treatment group, auditors were instead randomly assigned to plants and paid a fixed salary from a central pool of funds. The experiment demonstrated that the status quo system was largely corrupted, with auditors systematically reporting plant emissions just below the standard, although true emissions were typically higher. Second, the treatment caused auditors to report more truthfully and significantly lowered the fraction of plants that were falsely reported as compliant with pollution standards. Third, treatment plants, in turn, reduced their pollution emissions.

A different margin of potential corruption in information acquisition is providing higher-level officials discretion in whom to inspect or audit and when. Another is the choice of intermediaries who have monitoring responsibilities.

Duflo et al. (2015) examine this issue in the context of environmental inspections in India. They examine whether raising the frequency of inspections in a rule-bound manner changes regulator behavior and improves plant compliance. They find that more inspections lead the regulator to send more warnings but not to increase incidence of punishments. They use detailed information on regulator plant interactions to show that the regulator uses his discretion to target information collection and punishment efforts at a smaller subset of highly polluting plants. In this case, regulatory discretion is valuable as it allows the regulator to best target his scarce inspection resources.

Finally, a different way to improving monitoring is by increasing private sector incentives to report the truth. A commonly cited example is Value Added Tax (VAT), which generates paper trails on transactions between firms. Pomeranz (2015) analyzes the role of third-party information for VAT enforcement through two randomized experiments among over 400,000 Chilean firms. She shows that announcing additional monitoring has less impact on transactions that are subject to a paper trail, indicating the paper trail's preventive deterrence effect. This leads to enforcement spillovers up the VAT

chain. We return to the theme of how technology of service provision or revenue collection can be harnessed for better monitoring when we discuss the promise of e-governance.

4.3.3. Citizen Monitoring

The last decade has seen increased interest in monitoring undertaken directly by citizens. In part, this reflects the increasing incidence of freedom of information acts and in part the greater ease of obtaining already digitized data on program performance.

4.3.3.1.   Does information on program performance matter?

As discussed in section 3.3.2.2, Banerjee et al. (2010a) randomly provided Indian slum dwellers pamphlets and free newspapers containing information on candidate qualifications and legislator performance during the run-up to the 2008 Delhi election. The information increased voter turnout and reduced the incidence of vote buying. Banerjee et al. (2016) conduct a randomized evaluation to examine the impacts of giving communities the option to privatize rice distribution for a government rice subsidy program in Indonesia. Out of a sample of 572 villages, 191 villages were randomly assigned to have the option to privatize the distribution of subsidized rice through a bidding process. A significant fraction chose to do so, and roughly 17 percent of villages switched distributors and reduced price markup by 8 percent, improving overall program performance.

Positive impact of information provision is also found in Andrabi et al. (2014) in Punjab, Pakistan. The authors find that providing parents with information about children's academic performance, as well as the performance of all schools within the village, through report cards led to an increase of 0.11 standard deviations in test scores. The intervention also improved enrollment to primary schools by 4.5 percent and lowered private schools' tuition by 17 percent. In contrast, Banerjee et al. (2010b) report limited results from a report card intervention in which village volunteers prepared a report card on student learning. They interpret this as reflecting the absence of mechanisms to hold teachers accountable. This study points to the importance of understanding how monitoring mechanisms interact with underlying incentives for government workers.

Another method of providing information on program performance is through creating committees made up of citizens and program providers. Björkman and Svensson (2010) found that informing Ugandan citizens of the dismal state of local health service delivery and holding meetings between citizens and health workers to agree on action plans significantly reduced provider absenteeism, increased utilization, and improved health. In a second randomized evaluation, Björkman, de Walque, and Svensson (2016) examined a less expensive version of the program where they did not provide information on health worker performance and found no impacts, suggesting that access to information was key. The participation intervention alone resulted in the community identifying issues that primarily required third-party actions, rather than issues that local dispensaries could control. These results indicate that lack of information can make it harder to pin responsibility on the local provider.

These ideas are explored further in Pradhan et al. (2014) who examine the role of school committees in improving education quality in Indonesia. The study is a randomized evaluation covering 520 schools in Central Java from 2007 to 2008. They have four main treatments. The first treatment facilitated democratic elections of school committee members. The second treatment linked school committees to the village council by facilitating joint planning meetings (which they describe as linkage). They benchmark these two treatments against more common treatments: providing block grants and providing training. Two years later, test scores increased by 0.17 standard deviations for linkage and 0.23 standard deviations for linkage plus elections. In contrast, training did not impact learning, and the effect

of grants, while positive, was typically statistically indistinguishable from zero. Taken together, the contrasting results from India and Indonesia point strongly to the importance of institutional reforms that lead to positive interactions between monitoring and incentive mechanisms.

4.3.4. Summary

The evidence reviewed in this section leads to several conclusions. First, monitoring can help to resolve two agency problems. The first of these is the asymmetrical information that exists between the employer and the employee. With better monitoring, employees have less of an incentive to shirk, and employers also have the flexibility to offer high-powered contracts. The second is the information asymmetry between the service provider and the citizen. With better monitoring (and the disclosure of this information) citizens can hold their service providers accountable, by applying both bottom-up pressure, as well as inducing top-down pressure. Of course, improving monitoring capabilities is unlikely to induce much change without accountability mechanisms in place. Second, third-party reporting matters. With more monitoring, employees may try to game the system. Third-party reporting creates conflicting interests; this helps establish the veracity of the information.

Open research questions:
- *Does providing government information on its intermediaries' (civil servants) performance via audits or time-stamp technology improve outcomes?*
- *How can community-monitoring programs be designed to make them more effective in monitoring civil servants?*

4.4. Crossing cutting factors: the promise of e-governance and demographics

4.4.1. The promise of e-governance

A common theme that emerges from the body of experimental evidence is the sensitivity of individual behavior to the incentives they face. This, in turn, points to the importance of the structure within which government workers function. Below, we discuss how an emerging body of evidence suggests that e-governance and other technological changes in how the government functions may well help improve governance in low income settings, perhaps by replacing some of the functions played by personnel—who are subject to all the various problems discussed above—with technological solutions.

There are a number of ways in which technology can constrain the discretion of local officials and improve performance. For example, in low income countries the rural poor—an important target group for government transfer programs—are often less informed about state services available to them. In addition, limited state presence in rural areas implies that it is often harder to deploy traditional personnel-intensive monitoring mechanisms to ensure that intended program beneficiaries get their due. As a result, traditional modes of delivery tend to provide the village-level service provider significant discretion in who gets the transfer and when. In cases where payment is supposed to be conditional on the beneficiary undertaking certain activities (e.g., working in a workfare program or children going to school) the village provider often receives funds ahead of the activity having occurred. The monitoring system in these cases often focuses on reviewing the funds request system. Two recent papers examine how e-governance and the use of biometrics can help improve both the fund flow system from central coffers to village-level coffers and the transfer of resources from the village-level provider to the final beneficiaries.

Banerjee et al. (2016) report on a randomized evaluation which evaluated an e-governance reform of the fund-flow system for the federal workfare program in the Indian state of Bihar. In the status quo system, transfers were based on projected expenditures and passed through several intermediate

administrative tiers. The reform changed the traditional fund flow practice by conditioning fund disbursement for wage payments on incurred expenditure as reflected in worker detail entry on a new electronic platform. This reform also reduced the number of administrative tiers associated with wage disbursement and changed the informational requirements for requesting and disbursing program funds. It did not alter the flow of funds from the village fund to workers. The authors find that program expenditure and reported employment declined by 16 percent and 14 percent respectively, with no discernible impact on actual employment as measured by an independent survey. These changes reflected a reduction in corruption and leakage; the authors find a 5 percent reduction in fake households on program rolls and a 14 percent decrease in self-reported median personal assets for local program officials in reform areas relative to comparison areas. Thus, the financial reform was effective in reducing corruption and program costs but did not change actual demand met by the program.

Muralidharan et al. (2016) provide evidence on the impact of new payments technology on the last step of the payment process in India. They evaluate how "Smartcards," which coupled electronic transfers with biometric authentication, affected the functioning of two government welfare schemes in India, a workfare program (NREGS) and a pension program (SSP). Their experiment randomized the rollout of Smartcards over 157 subdistricts and 19 million people in the southern state of Andhra Pradesh. They find that even though the new Smartcard system was not fully implemented, it resulted in a faster and less corrupt payments process without adversely affecting program access. For each of these outcomes, treatment group distributions first-order stochastically dominated those of the control group. For example, there was a significant reduction in the "leakage" of funds between the government and beneficiaries in both the NREGS and SSP programs. NREGS recipients in treatment areas reported weakly earnings that were Rs. 35 higher (a 24 percent increase from Rs. 146 in the comparison group) with no corresponding change in government expenditures, indicating a 35 percent reduction in leakages. Furthermore, the Smartcards system was cost-effective, as time savings to NREGS beneficiaries alone were equal to the cost of the intervention, and beneficiaries overwhelmingly pthe new payment system. The results suggest that state capacity to implement welfare programs can be significantly enhanced by investing in secure payments infrastructure.

Barnwal (2014) reports supportive quasi-experimental evidence for the staggered rollout and subsequent pulling back of a biometric based scheme for fuel subsidies in India. He finds that the biometric-based transfer policy reduced fuel purchases in the domestic fuel sector by 11 percent to 14 percent, suggesting a reduction in subsidy diversion. In addition, after the policy is terminated, fuel purchases in the domestic sector revert to levels similar to before the policy was introduced.

A different way in which technology can help delivery of transfer programs is the use of mobile money. Aker et al. (2016) use data from a randomized experiment of a mobile money cash transfer program in Niger, and find evidence of benefits of this new system: Households receiving mobile transfers had higher diet diversity and children consumed more meals per day. These results can be partially attributed to increased time saving, as program recipients spent less time traveling to and waiting for their transfer, as well as increased intra-household bargaining power for women.

Technology can also play a role in reducing malfeasance in elections and promoting electoral accountability. Callen and Long (2015) implement an experiment to estimate the causal effects of photo quick count—a technology used to reduce electoral fraud—and its announcement on aggregation fraud. Photo quick count announcement reduced damaging of election materials by candidate representatives from 18.9 to 8.1 percent, and reduced votes for politically powerful candidates at a given polling location from about 21 to about 15 percent.

In a quasi-experimental study, Fujiwara (2015) studies the introduction of electronic voting on voter enfranchisement in Brazil. Electronic voting was introduced at scale in 1998 elections. But because of a

limited supply in devices, only municipalities with more than 40,500 registered voters used the new technology. Using a regression discontinuity design, he finds that electronic voting reduced residual voting in state legislature elections by a magnitude larger than 10 percent of total turnout, thus enfranchising millions of voters. By enfranchising a poorer and less education population, the introduction of electronic voting led to an increase in the number of pre-natal visits by health professionals and lowered the prevalence of low-weight births (below 2500g) by less educated women.

A third area where technology shows some promise is in government procurement. Government procurement accounts for an enormous amount of government expenditures, and despite many regulations put in place to ensure that procurement is conducted fairly and with limited corruption, the fact that procurement regulations must be implemented by officials allows scope for discretion. For example, officials can withhold detailed bidding documents from bidders outside the favored cartels. Lewis-Faupel et al. (2016) study the introduction of electronic procurement systems for public works projects in India and Indonesia using a differences-in-differences design that takes advantage of the differential rollout of electronic procurement by states/provinces over time. These electronic procurement systems allow greater access to information for all potential bidders and ensure that the procurement rules are followed correctly. Lewis-Faupel et al. (2016) find that electronic procurement leads to contracts being more likely to be won by providers from outside the region where the project is being executed, suggesting that an important role for e-procurement is increasing access to information. They also find that it leads to quality improvements, though not lower prices paid by the government.

4.4.2. Demographics and public service delivery

Gender may be a salient factor in the selection, recruitment, and performance of public officials. Gender bias may affect the recruitment of civil servants and the wages offered to applicants. Gender may also play a role in an individual's prosocial motivation and intrinsic motivation, with implications for performance. The design of personnel policy and incentive structures to improve performance may also have differential impacts by gender.

Regarding caste differences, Nagavarapu and Shekri (2016) seek to determine the role of informal monitoring and enforcement by social networks in reducing corruption and leakages in India's subsidized food program. They study India's targeted public distribution system, which provides poor households with subsidized grains and fuel through "Fair Price Shops" (PDS Shops). The study finds that for some subsidized items, distribution is higher among historically disadvantaged Scheduled Castes (SCs) when they buy from an SC shopkeeper at the PDS shop. For example, grain purchases increase by 51 percentage points when purchased by SC community members from an SC shopkeeper. However, there was no difference in take up for kerosene, where there is universal eligibility for benefits. At the same time, purchases decrease for non-SCs when there is an SC shopkeeper. A household demand and supply model constructed by the authors provides further evidence that SC pairing contributes to the ability of the community to monitor and enforce distribution.

An important area for research is whether expanding state capacity can impact existing gendered power imbalances. In post-conflict Papua New Guinea, young male combatants completely took over certain areas of the autonomous region of Bougainville; Cooper (2019) tests how assigning permanent police officer presence in villages through a program that devolves policing powers to carefully selected community members can empower women by enforcing a more equal rule of law. He identifies a strategic reaction of the existing system as men seek to preserve their advantage by increasing their reliance on local chiefs and institutions.

The interaction of demographics with public service delivery is an important area for future research.

Open research questions
- *Can e-governance limit corruption and leakages by automating processes previously conducted by individuals with discretion?*
- *Can e-procurement improve the efficiency of government investment?*
- *Do certain screening strategies reduce or exacerbate bias by gender, ethnicity, religion, or other demographic dimensions?*
- *How do biases based on demographic factors affect the wage premiums of civil servants?*
- *How does affirmative action affect service delivery?*
- *Are there differential responses to performance incentives by demographic factors?*

4.5. Concluding thoughts

In countries where the quality of government is low, public servants tend to be paid relatively well. While it is difficult to assign causality this relationship, it does hint at some important directions for future research.

That public servants earn on average significantly more than their private sector counterparts does not suggest that financial incentives do not matter for bureaucratic performance. The evidence we have reviewed here suggests that they clearly do. But rather than differences in the levels, it is the nature of the incentives that perhaps matter most for performance. In most countries, wages in the public sector tend to exhibit a high degree of compression, which combined with long tenure rates can make it hard to incentivize an individual once employed. But in settings where government can offer high-powered incentives (e.g., tax administration or education), the evidence suggests that public servants do in fact perform better.

High-powered incentives come with tradeoffs: these incentive schemes have the potential to discourage effort among the prosocially motivated, as well as to create issues of multi-tasking, though the empirical evidence suggests that in many cases these concerns are not as strong as one might initially have thought. The literature has begun to identify some of these tradeoffs, but much more research is needed to better understand in which settings these issues are most likely to arise. For example, task complexity might provide such a setting. The multi-tasking concerns associated with performance pay are more likely to arise when bureaucrats are tasked with complex jobs. At the same time, complex jobs are more difficult to monitor. Whether the benefits of lower monitoring costs outweigh the costs associated with multi-tasking is an interesting question with important implications for how bureaucracies should be organized.

Financial incentives also matter for selection. Organizations that offer higher wages will attract more qualified applicants. Given the large public-sector wage gaps, one might be tempted to reason that selection issues are not a first-order concern. But high wages are only a necessary condition for attracting talent, not a sufficient one. The type of individuals who work in the public sector will ultimately depend on how candidates are screened. The screening mechanism can easily undo any positive selection created by higher wages, as is perhaps the case when countries hire based on patronage, and not for meritocratic reasons. Governments vary in the way they screen their public servants. To understand how governments should screen among the candidates who apply to further improve selection is another exciting avenue for future research. It is of course difficult to consider optimal screening mechanisms without raising the question of what personality traits we should screen for. While there is an extensive empirical literature

arguing that individuals who exhibit high degrees of prosocialness perform better in the public sector, experimental evidence of this relationship is virtually non-existent.

The usefulness of financial incentives can also be limited if internal accountability mechanisms do not exist or function. One issue that bureaucracies typically face is the inability to perfectly monitor their employees. Fortunately advances in technology appear to be a step forward. The use of smartphone technology and e-governance platforms not only promote transparency and accountability, but also serve as disciplining devices. Importantly, these technological advances may also create a feedback loop on the compensation structure of employees. As governments increasingly adopt these new technologies, thus enabling them to better monitor and evaluate its employees, the set of contracts that it can offer its employees expands. The relationship between technology adoption and compensation scheme is another exciting area of future research.

As this survey documents, there is plenty to do and a lot to learn. We are excited to see what the next several years will bring for this research agenda, which has the potential to unlock some of the doors to efficient service delivery and good governance.

5. Bibliography

Acemoglu, D. 2010. "Theory, General Equilibrium and Political Economy in Development Economics." *Journal of Economic Perspectives* 24(3): 17-32.

Acemoglu, D., Cheema, A., Khwaja, A. I., & Robinson, J. A. (2018). Trust in State and Non-State Actors: Evidence from Dispute Resolution in Pakistan (No. w24611). *National Bureau of Economic Research*.

Acemoglu, D. and S. Johnson. 2005. "Unbundling Institutions." *Journal of Political Economy* 113(5): 949-995.

Acemoglu, D., S. Johnson, and J. A. Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review* 91(5): 1369-1401.

Acemoglu, D., T. Reed, and J.A. Robinson. 2014. "Chiefs: Economic Development and Elite Control of Civil Society in Sierra Leone." *Journal of Political Economy* 122(2): 319-368.

Acemoglu, D., J.A. Robinson, and R. Santos. 2013. "The Monopoly of Violence: Evidence from Colombia." *Journal of the European Economic Association* 11: 5-44.

Adida, C., J. Gottlieb, and E. Kramon, and G. McClendon. 2016. "How Coethnicity Moderates the Effect of Information On Voting Behavior: Experimental Evidence from Benin." Working Paper.

Aker, J., P. Collier, and P. Vicente. 2016. "Is Information Power? Using Mobile Phones and Free Newspapers During an Election in Mozambique." *Review of Economics and Statistics*.

Aker, J., and C. Ksoll. 2015. "Call Me Educated: Evidence from a Mobile Monitoring Experiment in Niger."

Aker, J. C., Boumnijel, R., McClelland, A., & Tierney, N. (2016). Payment mechanisms and antipoverty programs: Evidence from a mobile money cash transfer experiment in Niger. *Economic Development and Cultural Change*, 65(1), 1-37.

Alatas, V., A. Banerjee, R. Hanna, B. Olken, and J. Tobias. 2012. "Targeting the Poor: Evidence from a Field Experiment in Indonesia." *American Economic Review* 102(4): 1206-1240.

Alatas, V., A. Banerjee, R.A., Hanna, B.R., Olken, R.B. A., Purnamasari, and M.R., & Wai-Poi. , M. (2013.). "Does Elite Capture Matter? Local Elites and Targeted Welfare Programs in Indonesia." *National Bureau of Economic Research*, Working Paper 18798.

Alesina, A., R. Baqir, and W. Easterly. 1999. "Public Goods and Ethnic Divisions." *The Quarterly Journal of Economics* 114(4): 1243-1284.

Alesina, A. and B. Weder. 2002."Do Corrupt Governments Receive Less Foreign Aid?" *American Economic Review* 92(4): 1126-1137.

Almlund, Mathilde, Angela L. Duckworth, James Heckman, and Tim Kautz. 2011. "Personality Psychology and Economics." In *Handbook of the Economics of Education*, Volume 4, edited by Erik Hanushek, Stephen Machin, and Ludger Woessman, 1-181. Elsevier.

Anderson, S., P. Francois, and A. Kotwal. 2015. "Clientelism in Indian villages." *The American Economic Review* 105(6): 1780-1816.

Andrabi, T., J. Das, and A. Khwaja. 2014. "Report Cards: The Impact of Providing School and Child Test Scores on Educational Markets."

Andreoni, J. M. Callen, K.H. Jaffar, M.Y. Khan, and C. Sprenger. 2016. "Using Preference Estimates to Customize Incentives: An Application to Polio Vaccination Drives in Pakistan." *National Bureau of Economics* Working Paper 22019.

Aronson, E. 1999. "The Power of Self-Persuasion." *American Psychologist* 54(11): 875-884.

Asher, S., & Novosad, P. (2017). Politics and local economic growth: Evidence from India. *American Economic Journal: Applied Economics*, 9(1), 229-73.

Ashraf, N., O. Bandiera, and K. Jack. 2014. "No Margin, No Mission? A Field Experiment on Incentives for Public Service Delivery." *Journal of Public Economics* 120: 1-17.

Ashraf, N., Bandiera, O., & Lee, S. S. (2014). Awards unbundled: Evidence from a natural field experiment. Journal of Economic Behavior & Organization, 100, 44-63.

Ashraf, N., O. Bandiera, and S. Lee. 2016. "Do-Gooders and Go-Getters: Career Incentives, Selection, and Performance in Public Service Deliver." STICERD, LSE.

Ashraf, N., Bandiera, O., & Lee, S. S. (2019). Losing Prosociality in the Quest for Talent. Sorting, Selection, and Productivity in the Delivery of Public Services.

Avis, Eric, Claudio Ferraz, and Frederico Finan. (2016.). "Do Government Audits Reduce Corruption? Estimating the Impacts of Exposing Corrupt Politicians." *National Bureau of Economic Research*, Working Paper 22443.

Bai, J., Jayachandran, S., Malesky, E. J., & Olken, B. A. (2017). Firm growth and corruption: empirical evidence from Vietnam. *The Economic Journal*, 129(618), 651-677.

Baicker, K. and M. Jacobson. 2007. "Finders Keepers: Forfeiture Laws, Policing Incentives, and Local Budget." *Journal of Public Economics* 91: 2113-2136.

Baland, J.M., and J. A. Robinson. 2008. "Land and Power: Theory and Evidence from Chile." *American Economic Review* 98(5): 1737-1765.

Bandiera, O., A. Prat, and T. Valletti. "Active and Passive Waste in Government Spending: Evidence from a Policy Experiment." *American Economic Review* 99(4): 1278-1308.

Bandiera, O., and G. Levy. 2011. "Diversity and the power of the elites in democratic societies: Evidence from Indonesia." *Journal of Public Economics* 95(11-12): 1322-1330.

Bandiera, O., A. Prat, and T. Valletti. 2009."Active and Passive Waste in Government Spending: Evidence from a Policy Experiment." *American Economic Review* 99(4):1278-1308.

Banerjee, A. 1997. "A Theory of Misgovernance." *The Quarterly Journal of Economics* 112(4): 1289-1332.

Banerjee, A., R. Banerji, E. Duflo, R. Glennerster, and S. Khemani. 2010b. "Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in Education in India." *American Economic Journal: Economic Policy* 2(1): 1-30.

Banerjee, A., Duflo, E., Keniston, D., & Singh, N. (2017). The efficient deployment of police resources: theory and new evidence from a randomized drunk driving crackdown in India. Working paper.

Banerjee, A., A. Deaton, and E. Duflo. 2004. "Wealth, Health, and Health Services in Rural Rajasthan." *American Economic Review* 94(2): 326-330.

Banerjee, A. and E. Duflo. 2003. "Inequality and Growth: What Can the Data Say?" *Journal of Economic Growth* 8(3): 267-99.

Banerjee, A., E. Duflo, and R. Glennerster. 2008."Putting a Band-Aid on a Corpse: Incentives for Nurses in the Indian Public Health Care System." J*ournal of the European Economic Association* 6(2-3): 487-500.

Banerjee, A. E. Duflo, C. Imbert, S. Matthew, and R. Pande. 2016. "E-governance, Accountability, and Leakage in Public Programs: Experimental Evidence from a Financial Management Reform in India." (No. w22803). *National Bureau of Economic Research*.

Banerjee, A., D. Green, J. Green, and R. Pande. 2010. "Can Voters Be Primed to Choose Better Legislators? Evidence from Voter Campaigns in India." *World Bank* Working Paper.

Banerjee, A., Hanna, R., Kyle, J., Olken, B. A., & Sumarto, S. (2018). Tangible information and citizen empowerment: Identification cards and food subsidy programs in Indonesia. *Journal of Political Economy*, 126(2), 451-491.

Banerjee, A., R. Hanna, J. Kyle, B. Olken, and S. Sumarto. 2015. "Contracting out the Last-Mile of Service Delivery: Subsidized Food Distribution in Indonesia." *National Bureau of Economic Research* Working Paper 21837.

Banerjee, A., R. Hanna, and S. Mullainathan. 2012b. "Corruption." *National Bureau of Economic Research* Working Paper 17968.

Banerjee, A., S. Kumar, R. Pande, and F. Su. 2011. "Do Informed Voters Make Better Choices? Experimental Evidence from Urban India."

Banerjee, A. V. and R. Pande. 2009. "Parochial Politics: Ethnic Preferences and Politician Corruption."

Banerjee. R. T. Baul, and T. Rosenblat. 2015. "On Self Selection of the Corrupt into the Public Sector." *Economics Letters* 127: 43-46.

Banuri, S. and P. Keefer. 2013. "Intrinsic Motivation, Effort, and the Call to Public Service." *World Bank* Policy Research Working Paper 6729.

Bardhan, P. and D. Mookherjee. 2006. *Decentralization and Local Governance in Developing Countries: A Comparative Perspective*, Ed. 1 Vol. 1. Cambridge: The MIT Press.

Bardhan, P, and D. Mookherjee. 2012. "Political Clientelism and Capture: Theory and Evidence from West Bengal." *The Bureau for Research and Economic Analysis of Development* Working Paper.

Barnwal, P. 2014. "Curbing Leakage in Public Programs with Biometric Identification Systems: Evidence from India's Fuel Subsidies."

Basinga, P., P. Gertler, A. Binagwaho, A. Soucat, J. Sturdy, and C. Vermeersch. 2011. "Effect on Maternal and Child Health Services in Rwanda of Payment to Primary Health-Care Providers for Performance: An Impact Evaluation." *Lancet* 377(9975): 1421-1428.

Basu, K. 2011. "Why, for a Class of Bribes, the Act of Giving a Bribe Should be Treated as Legal."

Beaman, L., R. Chattopadhyay, E. Duflo, R. Pande, and P. Topalova. 2009. "Powerful Women: Does Exposure Reduce Bias?" *The Quarterly Journal of Economics* 124(4): 1497-1540.

Beaman, L., E. Duflo, R. Pande, and P. Topalova. 2011. "Political Reservation and Substantive Representation: Evidence from Indian Village Councils." India Policy Forum, Brookings, and NCAER.

Beath, A., F. Christia, and R. Enikolopov. 2013a. "Empowering Women through Development Aid: Evidence from a Field Experiment in Afghanistan." *American Political Science Review* 107(3): 540-557.

Beath, A., F. Christia, and R. Enikolopov. 2013b. "Direct Democracy and Resource Allocation: Experimental Evidence from Afghanistan."

Beath, A., F. Christia, and R. Enikolopov. 2013c. "Do Elected Councils Improve Governance Outcomes? Experimental Evidence on Local Institutions in Afghanistan." *World Bank* Policy Research Working Paper 6510.

Beath, A., F. Christia, G. Egorov, and R. Enikolopov. 2016. "Electoral Rules and Political Selection: Theory and Evidence from a Field Experiment in Afghanistan." *Review of Economic Studies* 83(3): 932-968.

Becker, G.S. and G. J. Stigler. 1974. "Law Enforcement, Malfeasance, and Compensation of Enforcers." *Journal of Legal Studies* 3(1): 1-18.

Behrman, J. S. Parker, P. Todd, and K. Wolpin. 2015. "Aligning Learning Incentives of Students and Teachers: Results from a Social Experiment in Mexican High Schools. No. 123. *Journal of Political Economy* 123(2): 325-364.

Bellé, N. 2012. "Experimental Evidence on the Relationship between Public Service Motivation and Job Performance." *Public Administration Review* 73 (1): 143-153.

Benabou, R. and J. Tirole. 2006. "Incentives and Prosocial Behavior." *American Economic Review* 96 (5): 1652-1678.

Berge, L. I. O., Bjorvatn, K., Galle, S., Miguel, E., Posner, D. N., Tungodden, B., & Zhang, K. (2015). How strong are ethnic preferences? (No. w21715). National Bureau of Economic Research.

Besley, T. and M. Ghatak. 2005. "Competition and Incentives with Motivated Agents." *American Economic Review* 95 (3): 616-636.

Berman, E., M. Callen, C. Gibson, and J. Long. 2014. "Election Fairness and Government Legitimacy in Afghanistan." *National Bureau of Economic Research* Working Paper 19949.

Bertrand, M., S. Djankov, R. Hanna, and S. Mullainathan. 2007. "Obtaining a Driver's License in India: An Experimental Approach to Studying Corruption." *The Quarterly Journal of Economics* 122(4): 1639-1676.

Besley,T. and R. Burgess. 2002. "The Political Economy of Government Responsiveness: Theory And Evidence From India." *The Quarterly Journal of Economics* 117(4): 1415-1451.

Besley, T., R. Pande, and V. Rao. 2005. "Participatory Democracy in Action: Survey Evidence from South India." *Journal of the European Economic Association* 3(2-3): 648-657.

Besley, T. and M. Reynal-Querol. 2011. "Do Democracies Select More Educated Leaders?" *American Political Science Review* 105(3): 552-566.

Besley, T., M. Reynal-Querol, and J. G. Montalvo. 2011. "Do Educated Leaders Matter?" *The Economic Journal* 554(121): F205-227.

Bhalotra, S. and I. Clots-Figueras. 2014. "Health and the Political Agency of Women." *American Economic Journal: Economic Policy*, 6(2): 164–197.

Bhalotra, S., I. Clots-Figueras, and L. Iyer. 2016. "Path-Breakers? Women's Electoral Success and Future Political Participation." Women's Electoral Success and Future Political Participation" H*arvard Business School* BGIE Unit Working Paper 14-035.

Bhavnani, R. 2009. "Do Electoral Quotas Work after They Are Withdrawn? Evidence from a Natural Experiment in India." *American Political Science Review* 103: 23-35.

Bhavnani, R. R. (2017). Do the effects of temporary ethnic group quotas persist? Evidence from India. *American Economic Journal: Applied Economics*, 9(3), 105-23.

Bidwell, K., K. Casey, and R. Glennerster. 2016. "Debates: Voting and Expenditure Responses to Political Communication." *London School of Economics* Political Science and Political Economy Seminar 17.

Björkman, M. and J. Svensson. 2009. "Power to the People: Evidence from a Randomized Field Experiment on Community-Based Monitoring in Uganda." *The Quarterly Journal of Economics* 124(2): 735-769.

Björkman, M. and J. Svensson. 2010. "When Is Community-Based Monitoring Effective? Evidence from a Randomized Experiment in Primary Health in Uganda." *Journal of the European Economic Association* 8(2-3): 571-581.

Björkman, M. D. de Walque, and J. Svensson. 2017. "Experimental Evidence on the Long-Run Impact of Community Based Monitoring." *American Economic Journal: Applied Economics*.

Blais, A. and A. Dobrzynska. 1998. "Turnout in Electoral Democracies." *European Journal of Political Research* 33(2): 239-262.

Blaydes, L. 2006. "Who Votes in Authoritarian Elections and Why? Political Mobilization, Turnout, and Spoiled Ballots in Contemporary Egypt."

Blum, F. (2016). Price-caps, discretion, and service delivery. *International Growth Centre Working Paper*.

Bobonis, Gustavo J., Luis R. Cámara Fuertes, and Rainer Schwabe. (2016.). "Monitoring Corruptible Politicians." *American Economic Review* 106(8): 2371-2405.

Borcan, O., Lindahl, M., & Mitrut, A. (2017). Fighting corruption in education: What works and who benefits?. *American Economic Journal: Economic Policy*, 9(1), 180-209.

Borjas, G.J. 2015. "The Wage Structure and the Sorting of Workers into the Public Sector." *National Bureau of Economic Research* Working Paper 9313.

Boycko, M., Shleifer, A., & Vishny, R. W. (1996). A theory of privatisation. The Economic Journal, 106(435), 309-319.

Bowles, J., & Larreguy, H. (2018). Who debates, who wins? at-scale experimental evidence on debate participation in a liberian election. Working Paper.

Bratton, M. 1999. "Political Participation in a New Democracy: Institutional Considerations from Zambia." *Comparative Political Studies* 32(5): 549-588.

Bratton, M. 2008. "Vote Buying and Violence in Nigerian Election Campaigns." *Electoral Studies* 27(4): 621-632.

Brollo, F., T. Nannicini, R. Perotti, and G. Tabellini. 2013. "The Political Resource Curse." *American Economic Review* 103(5): 1759-1796.

Brollo and U. Troiano. 2016. "What Happens When a Woman Wins an Election? Evidence from Close Races in Brazil." *Journal of Development Economics* 122: 28–45.

Bruhn, M. 2008. "License to Sell: The Effect of Business Registration Reform on Entrepreneurial Activity in Mexico." *Review of Economics and Statistics* 93(1): 382-386.

Brunetti, A. and B. Weder. 2003. "A Free Press is Bad News for Corruption." *Journal of Public Economics* 87(7-8): 1801-1824.

Burgess, R., M. Hansen, B. Olken, P.Potapov and S. Sieber. 2011. "The Political Economy of Deforestation in the Tropics." *National Bureau of Economic Research* Working Paper 17417.

Burgess, R., R. Jedwab, E. Miguel, A. Morjaria, and G. Miquel. 2015. "The Value of Democracy: Evidence from Road Building in Kenya." *The American Economic Review* 105(6): 1817-1851.

Callen, M., S. Gulzar, A. Hasanain, and Y. Khan. 2016. "The Political Economy of Public Sector Absence: Experimental Evidence from Pakistan."

Callen, M., and J. D. Long. 2015. "Institutional Corruption and Election Fraud: Evidence from a Field Experiment in Afghanistan." *American Economic Review* 105(1): 354-81.

Camacho, A. and E. Conover. 2011. "Manipulation of Social Program Eligibility: Detection, Explanations and Consequences for Empirical Research." *American Economic Journal: Economic Policy* 3(2): 41-65.

Carrillo, P., Pomeranz, D., & Singhal, M. (2017). Dodging the taxman: Firm misreporting and limits to tax enforcement. *American Economic Journal: Applied Economics*, 9(2), 144-64.

Caselli, F. and G. Michaels. 2009. "Do Oil Windfalls Improve Living Standards? Evidence from Brazil." *American Economic Journal: Applied Economics* 5(1): 208-238.

Casey, K., R. Glennerster, and E. Miguel. 2012. "Healing the Wounds: Learning from Sierra Leone's Post-War Institutional Reforms." *National Bureau of Economic Research* Working Paper 18368.

Casey, K., 2018. Radical decentralization: Does community-driven development work? *Annual Review of Economics,* 10, pp.139-163.

Castro, L. and C. Scartascini. 2015. "Tax Compliance and Enforcement in the Pampas: Evidence from a Field Experiment." *Journal of Economic Behavior and Organization* 116: 65-82.

Celhay, P. P. Gertler, P. Giovagnoli, and C. Vermeersch. 2015. "Long Run Effects of Temporary Incentives on Medical Care Productivity. *National Bureau of Economic Research* Working Paper 21361.

Chattopadhyay, R. and E. Duflo. 2004. "Women as Policy Makers: Evidence from a Randomized Policy Experiment in India." *Econometrica* 72(5): 1409-1443.

Chauchard, S. 2014. "Can Descriptive Representation Change Beliefs about a Stigmatized Group? Evidence from Rural India." *The American Political Science Review* 108(2): 1-20.

Chaudhury, N., J. Hammer, M. Kremer, K. Muralidharan and F. H. Rogers. 2006. "Missing in Action: Teacher and Health Worker Absence in Developing Countries." *Journal of Economic Perspectives* 20(1): 91-116.

Chin, A., and N. Prakash. 2011. "The Redistributive Effects of Political Reservation for Minorities: Evidence from India." *Journal of Development Economics* 96(2): 265-77.

Chong, A., A. L. De La O, D. Karlan, and L. Wantchkon. 2013. "Looking Beyond the Incumbent: The Effects of Exposing Corruption on Electoral Outcomes."

Church, T. and M. Heumann. 1989. "The Underexamined Assumptions of the Invisible Hand: Monetary Incentives as Policy Instruments." *Journal of Policy Analysis and Management* 8 (4): 641-657.

Clots-Figueras, I. 2011. "Women in Politics: Evidence from the Indian States." *Journal of Public Economics* 95(7-8): 664-690.

Clots-Figueras, I. 2012. "Are Female Leaders Good for Education? Evidence from India." *American Economic Journal: Applied Economics* 4(1): 212-244.

Cole, S., A. Healy, and E. Werker. 2012. "Do Voters Demand Responsive Governments? Evidence from Indian Disaster Relief." *Journal of Development Economics* 97(2): 167-181.

Collier, P. and P. Vicente. 2014. "Votes and Violence: Evidence from a Field Experiment in Nigeria." *The Economic Journal* 124(574): F327-355.

Cooper, J. (2018). When Power Does Not Corrupt: Evidence from randomized police recruitment. Working paper presented at EPSA, Vienna.

Cooper, J. (2019). State capacity and gender inequality: Experimental evidence from Papua new Guinea. New York, NY: Columbia University.

Cowley, E. and S. Smith. 2014. "Motivation and Mission in the Public Sector: Evidence from the World Values Survey." *Theory and Decision* 76: 241-263.

Dal Bo, E., P. Dal Bo, and J. Snyder. 2008. "Political Dynasties." *Review of Economic Studies* 76(1): 115-142.

Dal Bo, E., F. Finan, and M.A. Rossi. 2013. "Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service." *Quarterly Journal of Economics* 128 (3): 1169-1218.

Dal Bó, E., Finan, F., Li, N. Y., & Schechter, L. (2019). Government decentralization under changing state capacity: Experimental evidence from Paraguay (No. w24879). National Bureau of Economic Research.

Da Silveira, B. and J. De Mello. 2011. "Campaign Advertising and Election Outcomes: Quasi-natural Experiment Evidence from Gubernatorial Elections in Brazil." *Review of Economic Studies* 78(2): 590-612.

DeBacker, J. M., B. T. Heim, and A. Tran. 2012. "Importing Corruption Culture from Overseas: Evidence from Corporate Tax Evasion in the United States." *National Bureau of Economic Research* Working Paper 17770.

De Janvry, A., F. Finan, and E. Sadoulet. 2012. "Local Electoral Incentives and Decentralized Program Performance." *Review of Economics and Statistics* 94(2): 672-685.

De Janvry, A., M. Gonzalez-Navarro, and E. Sadoulet. 2014. "Are Land Reforms Granting Complete Property Rights Politically Risky? Electoral Outcomes of Mexico's Certification Program." *Journal of Development Economics* 110: 216-225..

De Ree, J., K. Muralidharan, M. Pradhan, and H. Rogers. 2015. "Double for Nothing? Experimental Evidence on the Impact of an Unconditional Teacher Salary Increase on Student Performance in Indonesia. *National Bureau of Economic Research* 21806.

Delfgaauw, J. and R. Dur. 2007. "Signaling and Screening of Workers' Motivation." *Journal of Economic Behavior & Organization* 62 (4): 605-624.

Deserranno, E. 2015. "Financial Incentives as Signals: Experimental Evidence from the Recruitment of Health Workers."

Dhaliwal, I. and R. Hanna. 2014. "Deal with the Devil: The Successes and Limitations of Bureaucratic Reform in India." *National Bureau of Economic Research* Working Paper 20482.

Di Tella, R., and I. Franceschelli. 2011. "Government Advertising and Media Coverage of Corruption Scandals." *American Economic Journal: Applied Economics* 3(4): 119-151.

Di Tella, R., and E. Schargrodsky. 2003. "The Role of Wages and Auditing during a Crackdown on Corruption in the City of Buenos Aires." *Journal of Law & Economics* 46 (1): 269–92.

Dizon-Ross, Rebecca, Pascaline Dupas, and Jonathan Robinson. (2015.). "Governance and the Effectiveness of Public Health Subsidies." *National Bureau of Economic Research*, Working Paper 21324.

Djankov, S., R. La Porta, F. Lopez-de-Silanes and A. Shleifer. 2003. "Courts." *The Quarterly Journal of Economics* 118(2): 453-517.

Djankov, S., R. La Porta, F. Lopez de Silanes and A. Shleifer. 2010. "Disclosure by Politicians." *American Economic Journal: Applied Economics* 2(2): 179-209.

Dodge, E., Neggers, Y., Pande, R., & Moore, C. T. (2017). Having it at Hand: How Small Search Frictions Impact Bureaucratic Efficiency.

Dohmen, T., A. Falk, D. Huffman, and U. Sunde. 2010. "Are Risk Aversion and Impaction Related to Cognitive Ability?" *American Economic Review* 100(3): 1238-1260.

Dower, P. and T. Pfutze. 2015. "Vote Suppression and Insecure Property Rights." *Journal of Development Economics* 114: 1-19.

D'Souza, A. 2012. "The OECD Anti-Bribery Convention: Changing the Currents of Trade." *Journal of Development Economics* 97(1): 73-87.

Dunning, T., Monestier, F., Piñeiro, R., Rosenblatt, F., & Tuñón, G. (2017). Is Paying Taxes Habit Forming? Theory and Evidence from Uruguay.

Dunning, T. and J. Nilekani. 2013. "Ethnic Quotas and Political Mobilization: Caste, Parties, and Distribution in Indian Villages." *American Political Science Review* 107(1): 35-56.

Duflo, E., M. Greenstone, R. Pande, and N. Ryan. 2013. "Truth-telling by Third-Party Auditors and the Response of Polluting Firms: Experimental Evidence from India." *Quarterly Journal of Economics* 128(4): 1499-1545.

Duflo, E., P. Dupas, and M. Kremer. 2012a . "School Governance, Teacher Incentives, and Pupil-Teacher Ratios: Experimental Evidence from Kenyan Primary Schools." National Bureau of Economic Research Working Paper 17939.

Duflo, E., M. Greenstone, R. Pande, and N. Ryan. (2014.). "The Value of Discretion in the Enforcement of Regulation: Experiment Evidence and Structural Estimates from Environmental Inspections in India" *National Bureau of Economic Research*NBER Working Paper No. 20590.

Duflo, E., M. Greenstone, R. Pande, and N. Ryan. 2015. "The Value of Regulatory Discretion: Estimates from Environmental Inspections in India." *Harvard* Environmental Economics Program Discussion Paper 15-60.

Duflo, E., R. Hanna and S. Ryan. 2012b. "Incentives Work: Getting Teachers to Come to School." *American Economic Review* 102(4): 1241-78.

Duflo, E. and P. Topalova. 2004. "Unappreciated Service: Performance, Perceptions, and Women Leaders in India."

Dunning, T., and L. Harrison. 2010. "Cross-Cutting Cleavages and Ethnic Voting: An Experimental Study of Cousinage in Mali." *American Political Science Review* 104(1): 21-39.

Easterly, W. and R. Levine. 1997. "Africa's Growth Tragedy: Policies and Ethnic Divisions." *The Quarterly Journal of Economics* 112(4): 1203-50.

Eifert, B., Miguel, E. and Posner, D. N. 2010. "Political Competition and Ethnic Identification in Africa." *American Journal of Political Science* 54: 494–510.

Engerman S. and K. Sokoloff. 2002. "Factor Endowments, Inequality, and Paths of Development Among New World Economics." *National Bureau of Economic Research* Working Paper 9259.

Faccio, M. 2006."Politically Connected Firms." *American Economic Review* 96(1): 369-386.

Faccio, M, and D. Parsley. 2009."Sudden Deaths: Taking Stock of Geographic Ties." *Journal of Financial and Quantitative Analysis* 44(03): 683-718.

Fafchamps, M., and J. Labonne. 2016. "Family Networks and Distributive Politics." *Centre for Economic Policy Research* Discussion Paper 11245.

Faguet, J. P. 2004."Does Decentralization Increase Government Responsiveness to Local Needs? Evidence from Bolivia." *Journal of Public Economics* 88(3-4): 867-893.

Fearon, J., M. Humphreys, and J. Weinstein. 2015. "How Does Development Assistance Affect Collective Action Capacity? Results from a Field Experiment in Post-Conflict Liberia." *The American Political Science Review* 109(3): 450-469.

Ferraz, C. and F. Finan. 2008. "Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes." *The Quarterly Journal of Economics* 123(2): 703-745.

Ferraz, C. and F. Finan. 2010. "Electoral Accountability and Corruption: Evidence from the Audit Reports of Local Governments." *American Economic Review* 101(4): 1274-1311.

Ferraz, C. and F. Finan. 2011. "Motivating Politicians: The Impacts of Monetary Incentives on Quality and Performance."

Ferraz, C., F. Finan and D.B. Moreira. 2012. "Corrupting Learning: Evidence from Missing Federal Education Funds in Brazil." *Journal of Public Economics* 96(9-10): 712-726.

Finan, F., Olken, B. A., & Pande, R. (2017). The personnel economics of the developing state. In Handbook of Economic Field Experiments (Vol. 2, pp. 467-514). North-Holland.

Finan, F. and L. Schechter. 2012. "Vote-Buying and Reciprocity." *Econometrica* 80(2): 863-881.

Fisman, R. 2001. "Estimating the Value of Political Connections." *American Economic Review* 91(4): 1095-1102.

Fisman, D., R. Fisman, J. Galef, R. Khurana, and Y. Wang. 2012a. "Estimating the Value of Connections to Vice-President Cheney." *B.E. Journal of Economic Analysis & Policy* 12(3).

Fisman, R. and R. Gatti. 2002. "Decentralization and Corruption: Evidence Across Countries." *Journal of Public Economics* 83(3): 325-345.

Fisman, R., N. A. Harmon, E. Kamenica, and I. Munk. 2015 "Labor Supply of Politicians." *Journal of the European Economic Association* 13(5): 871-905.

Fisman, R. and E. Miguel. 2007. "Corruption, Norms, and Legal Enforcement: Evidence from Diplomatic Parking Tickets." *Journal of Political Economy* 115(6): 1020-1048.

Fisman, R., F. Schulz, and V. Vig. 2014. "Private Returns to Public Office." *Journal of Political Economy* 122(4): 806-862.

Fisman, R. and J. Svensson. 2007. "Are Corruption and Taxation Really Harmful to Growth? Firm level Evidence." *Journal of Development Economics* 83(1): 63-75.

Fisman, R., & Wang, Y. (2017). The distortionary effects of incentives in government: Evidence from China's" death ceiling" program. *American Economic Journal: Applied Economics*, 9(2), 202-18.

Fisman, R., and S-J. Wei. 2004. "Tax Rates and Tax Evasion: Imports in China." *Journal of Political Economy* 112(2): 471-496.

Fisman, R., & Wang, Y. (2015.). "The Mortality Cost of Political Connections." The Review of Economic Studies, rdv020.

Forbes, K. 2000. "A Reassessment of the Relationship between Inequality and Growth." *American Economic Review* 90(4): 869-887.

Fornos, C. A. 2004. "Explaining Voter Turnout in Latin America, 1980 to 2000." *Comparative Political Studies* 37(8): 909-940.

Foster, A. and M.R. Rosenzweig. 2002. "Democratization, Decentralization and the Distribution of Local Public Goods in a Poor Rural Economy." *The Bureau for Research and Economic Analysis of Development* Working Paper.

Francois, P. 2000. "'Public Service Motivation' as an Argument for Government Provision." *Journal of Public Economics* 78 (3): 275-299.

Fujiwara, T. 2015. "Voting Technology, Political Responsiveness, and Infant Health: Evidence from Brazil." *Econometrica*, 83(2): 423-464.

Fujiwara, T. and L. Wantchekon. 2013. "Can Informed Public Deliberation Overcome Clientelism? Experimental Evidence from Benin." *American Economic Journal: Applied Economics* 5(4): 241-255.

Gadenne, L. and M. Singhal. 2013. "Decentralization in Developing Economies." *National Bureau of Economic Research* Working Paper 19402.

George, S., Gupta, S., & Neggers, Y. (2018). Coordinating Voters against Criminal Politicians: Evidence from a Mobile Experiment in India. Working paper.

Gerardino, M., S. Litschig, and D. Pomeranz. 2015. "Monitoring Public Procurement: Evidence from a Regression Discontinuity Design in Chile."

Gertler, P., H. A. Patrinos, and M. Rubio-Codina. 2012. "Empowering Parents to Improve Education: Evidence from Rural Mexico." *Journal of Development Economics* 99(1): 68-79.

Gertler, P. and C. Vermeersch. 2012. "Using Performance Incentives to Improve Health Outcomes." *World Bank* Policy Research Working Paper 6100.

Ghani, E., K. William, and S. O'Connell. 2014. "Political Reservations and Women's Entrepreneurship in India." *Journal of Economic Development* 108: 138-153.

Gine, X.and G. Mansuri. 2011. "Together We Will: Evidence from a Field Experiment on Female Voter Turnout in Pakistan." *World Bank* Policy Research Working Paper 5692.

Glaeser, E., R. La Porta, F. Lopez-de-Silanes and A. Shleifer. 2004. "Do Institutions Cause Growth?" *Journal of Economic Growth* 9(3): 271-303.

Glennerster, R., E. Miguel, and A. Rothenberg. 2013. "Collective Action in Diverse Sierra Leone Communities." *The Economic Journal* 123(568): 285-316.

Glewwe, P., I. Nauman and M. Kremer. 2010. "Teacher Incentives." *American Economic Journal: Applied Economics* 2(3): 205-27.

Gneezy, U. and A. Rustichini. 2000. "A Fine is a Price." *Journal of Legal Studies* 29 (1).

Golden, M., Gulzar, S., & Sonnet, L. (2017). Improving Political Communication: Results from a Pilot Field Experiment in Pakistan.

Golden, M. A. and D. Tiwari. 2009. "Criminality and Malfeasance among National Legislators in Contemporary India." APSA 2009 Toronto Meeting Paper.

Gordon, R. and W. Li. 2009. "Tax Structures in Developing Countries: Many Puzzles and a Possible Explanation." *Journal of Public Economics* 93(7-8): 855-866.

Gorodnichenko, Y. and K. Sabirianova Peter. 2007. "Public Sector Pay and Corruption: Measuring Bribery from Micro Data." *Journal of Public Economics* 91(5-6): 963-991.

Grant, A.M., E. Campbell, G. Chen, K. Cottone, D. Lapedis, and K. Lee. 2007. "Impact and the Art of Motivation Maintenance: The Effects of Contact with Beneficiaries on Persistence Behavior." *Organizational Behavior and Human Decision Processes* 103: 53-67.

Grossman, G., M. Humphreys, and G. Sacramone-Lutz. 2014. "I wld like u WMP to extend electricity 2 our village": On Information Technology and Interest Articulation." *The American Political Science Review* 108(3): 688-705.

Gugerty, M. K. and M. Kremer. 2008. "Outside Funding and the Dynamics of Participation in Community Associations." *American Journal of Political Science* 52: 585–602.

Guiteras, R. A.M. Mobarak. 2015. "Does Development Aid Undermine Political Accountability? Leader and Constituent Responses to a Large-Scale Intervention." *National Bureau of Economic Research* Working Paper 21434.

Habyarimana, J., M. Humphreys, D. Posner and J. Weinstein. 2006. "Why Does Ethnic Diversity Undermine Public Goods Provision?" *American Political Science Review* 101(4): 709-725

Hanna, R. and S.Y. Wang. 2015. "Dishonesty and Selection into Public Service: Evidence from India." *National Bureau of Economic Research* Working paper 19649.

Hicken, A., S. Leider, N. Ravanilla, and D. Yang. 2015. "Temptation in Vote-Selling: Evidence from a Field Experiment in the Philippines."

Holmstrom, B. and P. Milgrom. 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics and Organization* 7(0): 24-52.

Hsieh, C.T., E. Miguel, D. Ortega, and F. Rodriguez. 2011. "The Price of Political Opposition: Evidence from Venezuela's Maisanta." *American Economic Journal: Applied Economics* 3: 196-214.
Hsieh, C.T. and E. Moretti. 2006."Did Iraq Cheat the United Nations? Underpricing, Bribes, and the Oil for Food Program." *The Quarterly Journal of Economics* 121(4):1211-1248.

Huillery, E., and J. Seban. 2015. "Misplaced E ort: The E ect of Pay-for-Performance in the Health Sector in the DRC."

Humphreys, M., W. Masters and M. Sandbu. 2006. "The Role of Leaders in Democratic Deliberations: Results from a Field Experiment in São Tomé and Príncipe." *World Politics* 58: 583-622.

Humphreys, M., de la Sierra, R., and P. Van der Windt. 2015. "Social Engineering in the Tropics: A Grassroots Democratization Experiment in the Congo."

Humphreys, M. and J. Weinstein. 2012. "Policing Politicians: Citizen Empowerment and Political Accountability in Uganda." *International Growth Centre* Working Paper S-5021-UGA-1.

Hunt, J. 2004. "Trust and Bribery: The Role of the Quid Pro Quo and the Link with Crime." *National Bureau of Economic Research* Working Paper 10510.

Hunt, J. 2007. "How Corruption Hits People When they are Down." *Journal of Development Economics* 84(2): 574-589.

Huntington, S. P. 1991. "Democracy's Third Wave." *The Journal of Democracy* 2(2).

Ichino, N., and N. Nathan. 2005. "Do Primaries Improve Electoral Performance? Clientalism and Intra-Party Politics in Ghana." *American Journal of Political Science* 57(2): 428-441.

Ichino, N. and N. Nathan. 2013. "Crossing the Line: Local Ethnic Geography and Voting in Ghana." *American Political Science Review* 107(2): 344-361.

Ichino, N. and M. Schundeln. 2010. "Deterring or Displacing Electoral Irregularities? Spillover Effects of Observers in a Randomized Field Experiment in Ghana." *Journal of Politics* 74(1): 292-307.

Iyer, L. and A. Mani. 2012. "Traveling Agents: Political Change and Bureaucratic Turnover in India." *Journal of Public Economics* 92(4): 693-713.

Iyer, L., A. Mani, P. Mishra, and P. Topalova. 2012. "The Power of Political Voice: Women's Political Representation and Crime in India." *American Economic Journal: Applied Economics* 4(4): 165-193.

Jack, K. and M. Recalde. 2015. "Leadership and the Voluntary Provision of Public Goods: Field Evidence from Bolivia." *Journal of Public Economics* 122: 80-93.

Javorcik, B. S. and S.J. Wei. 2009. "Corruption and cross-border investment in emerging markets: Firm-level evidence." *Journal of International Money and Finance* 28(4): 605-624.

Jayachandran, S. 2006. "The Jeffords Effect." *Journal of Law and Economics* 49(2): 397-425.

Jensenius, F.R. 2015. "Development from Representation? A Study of Quotas for the Scheduled Castes in India." *American Economic Journal: Applied Economics* 7(3): 196–220.

Jia, R., M. Kudamatsu, and D. Seim. 2015. "Political Selection in China: The Complementary Roles of Connections and Performance." *Journal of the European Economic Association* 13(4): 631–668.

Jones, B.F. and B. A. Olken. 2005. "Do Leaders Matter? National Leadership and Growth since World War II." *The Quarterly Journal of Economics* 120(3): 835-864.

Kahn, C.M., E.C.D Silva and J. P. Ziliak. 2001."Performance-Based Wages in Tax Collection: The Brazilian Tax Collection Reform and Its Effects." *The Economic Journal* 111(468): 188-205.

Kahn, M., P. Li, and D. Zhao. 2015. "Water Pollution Progress at Borders: The Role of Changes in China's Political Promotion Incentives." *American Economic Journal: Economic Policy* 7(4): 223-242.

Kaplan, D., E. Piedra, and E. Seira. 2011. "Entry regulation and business start-ups: Evidence from Mexico." *Journal of Public Economics* 95(11-12): 1505-1515.

Khan, A., A. Khwaja, and B. Olken. 2015. "Tax Farming Redux: Experimental Evidence on Performance Pay for Tax Collectors." *National Bureau of Economic Research* Working Paper 20627.

Khan, A., A. Khwaja, and B. Olken. 2016. "Making Moves Matter: Experimental Evidence on Incentivizing Bureaucrats through Performance-Based Transfers."

Khwaja, A.I. 2009. "Can Good Projects Succeed in Bad Communities?" *Journal of Public Economics* 93(7-8): 899-916.

Khwaja, A.I. and A. Mian. 2005. "Do Lenders Favor Politically Connected Firms? Rent Provision in an Emerging Financial Market." *The Quarterly Journal of Economics* 120(4): 1371-1411.

Kleven, H.J., M. B. Knudsen, C. T. Kreiner , S. Pedersen and E. Saez. 2011. "Unwilling or Unable to Cheat? Evidence from a Randomized Tax Audit Experiment in Denmark." *Econometrica* 79(3): 651-692.

Kremer, M., N. Chaudhury, F. H. Rogers, K. Muralidharan and J. Hammer. 2005. "Teacher Absence in India: A Snapshot." *Journal of the European Economic Association* 3(2-3): 658-667.

Kumler, T., E. Verhoogen, and J. Frias. 2015. "Enlisting Employees in Improving Payroll-Tax Compliance: Evidence from Mexico." *National Bureau of Economic Research* Working Paper 19385.

Kuncoro. A. 2006. "Corruption and Business Uncertainty in Indonesia." *ASEAN Economic Bulletin* 23(1): 11-30.

Labonne, J. 2013. "The Local Electoral Impacts of Conditional Cash Transfers: Evidence from a Field Experiment." *Journal of Development Economics* 104: 73–88

Lagarde, M. and D. Blaauw. 2014. "Prosocial Preferences and Self-Selection into Jobs: Evidence from South African Nurses." *Journal of Economic Behavior and Organization* 107(A):136-152.

Larreguy, H. 2013. "Monitoring Political Brokers: Evidence from Clientelistic Networks in Mexico." EPSA 2013 Annual General Conference Paper 655.

Larreguy, H., J. Marshall, and J. Snyder Jr. 2014. "Revealing Malfeasance: How Local Media Facilitates Electoral Sanctioning of Mayors in Mexico." *National Bureau of Economic Research* Working Paper 20697.

Larreguy, H., Marx, B., Reid, O., & Blattman, C. (2018). A Market Equilibrium Approach to Reduce the Incidence of Vote-Buying: Evidence from Uganda. Working paper.

Larreguy, H., J. Marshall, and P. Querubán. 2016. "Parties, Brokers, and Voter Mobilization: How Turnout Buying Depends Upon the Party's Capacity to Monitor Brokers." T*he American Political Science Review* 110(1): 160-179.

Leon, G. 2013. "Turnout, Political Preferences, and Information: Experimental Evidence from Peru." *The Bureau for Research and Economic Analysis of Development* Working Paper.

Lewis-Faupel, S. Y. Neggers, B. Olken, and R. Pande. 2016. "Can Electronic Procurement Improve Infrastructure Provision? Evidence from Public Works in India and Indonesia." *American Economic Journal: Economic Policy* 8(3): 258-83.

Lichand, G., M. Lopes, and M. Medeiros. 2015. "Is Corruption Good For Your Health?"

Linden, L., and G. Kartini Shastry. 2012. "Grain Inflation: Identifying Agent Discretion in Response to a Conditional School Nutrition Program." *Journal of Development Economics* 99(1): 128-138.

Litschig, S. 2012. "Are Rules-Based Government Programs Shielded from Special-Interest Politics? Evidence from Revenue-Sharing Transfers in Brazil." *Journal of Public Economics* 96(11-12): 1047-1060.

Lott, J. and L. Kenny. 1999. "Did Women's Suffrage Change the Size and Scope of Government?" *Journal of Political Economy* 107(6): 1163-1198.

Lowe, M. (2018). Types of contact: A field experiment on collaborative and adversarial caste integration. Unpublished Manuscript.

Malesky, E. J. C.V. Nguyen and A. Tran. 2014. "The Impact of Recentralization on Public Services: A Difference-in-Difference Analysis of the Abolition of Elected Councils in Vietnam." *The American Political Science Review* 108(1): 144-168.

Malesky, E. and K. Samphantharak. 2008. "Predictable Corruption and Firm Investment: Evidence from a Natural Experiment and Survey of Cambodian Entrepreneurs." *Quarterly Journal of Political Science* 3 (3): 227-267.

Malesky, E., & Taussig, M. (2019). Participation, government legitimacy, and regulatory compliance in emerging economies: A firm-level field experiment in Vietnam. American Political Science Review, 113(2), 530-551.

Manacorda, M., E. Miguel, and A. Vigorito. 2011. "Government Transfers and Political Support." *American Economic Journal: Applied Economics* 3 (3): 1-28.

Martinez-Bravo, M. 2014. "The Role of Local Officials in New Democracies: Evidence from Indonesia." *American Economic Review* 104(4): 1244-87.

Martinez-Bravo, M., G. Padro-i-Miquel, N. Qian, and Y. Yao. 2011. "Do Local Elections in Non-Democracies Increase Accountability? Evidence from Rural China." *National Bureau of Economic Resaerch* Working Paper 16948.

Martinez-Bravo, M., G. Padro-i-Miquel, N. Qian, and Y. Yao. 2012. "The Effects of Democratization on Public Goods and Redistribution: Evidence from China." *National Bureau of Economic Research* Working Paper 18101.

Marx, B., V. Pons, and T. Suri. 2016. "The Perils of Building Democracy in Africa."

Mauro, P. 1995. "Corruption and Growth." *The Quarterly Journal of Economics* 110(3): 681-712.

McMillan, J. and P. Zoido. 2004. "How to Subvert Democracy: Montesinos in Peru." *Journal of Economic Perspectives* 18(4): 69-92.

Meares, T. 1995. "Rewards for Good Behavior: Influencing Prosecutorial Discretion and Conduct with Financial Incentives." Faculty Scholarship Series Paper 473.

Mellstrom, C. and M. Johannesson. 2008. "Crowding Out in Blood Donation: Was Titmuss Right?" *Journal of the European Economic Association* 6 (4): 845-863.

Meyersson, E. 2014 "Islamic Rule and the Empowerment of the Poor and the Pious." *Econometrica* 82(1): 229-269.

Miguel, E. and M. K. Gugerty. 2005. "Ethnic Diversity, Social Sanctions, and Public Goods in Kenya," *Journal of Public Economics* 89 (11-12): 2325-2368.

Miller, G. 2008."Women's Suffrage, Political Responsiveness, and Child Survival in American History." *The Quarterly Journal of Economics* 123(3): 1287-1327.

Miller, G., R. Luo, L. Zhang, S. Sylvia, Y. Shi, P. Foo, Q. Zhao, R. Martorell, A. Medina, and S. Rozelle. 2012. "Effectiveness of Provider Incentives for Anaemia Reduction in Rural China: A Cluster Randomised Trial." *BMJ* 345: e4809.

Miller, M. and L. Selva. 1994. "Drug Enforcement's Double-Edged Sword: An Assessment of Asset Forfeiture Programs." *Justice Quarterly* 11 (2): 313-

Mironov, Maxim, and Ekaterina Zhuravskaya. (2016.). "Corruption in Procurement and the Political Cycle in Tunneling: Evidence from Financial Transactions Data." *American Economic Journal: Economic Policy* 8(2): 287-321.

Mocan, N. 2008. "What Determines Corruption? International Evidence from Microdata." *Economic Inquiry, Western Economic Association International* 46(4): 493-510.

Monteiro, J. and J. Assunção. 2012. "Coming Out of the Shadows? Estimating the Impact of Bureaucracy Simplification and Tax Cut on Formality in Brazilian Microenterprises." *Journal of Development Economics* 99(1): 105-115.

Monteiro, J. and C. Ferraz. 2012. "Does Oil Make Leaders Unaccountable? Evidence from Brazil's Offshore Oil Boom."

Muralidharan, K. 2012. "Long-Term Effects of Teacher Performance Pay: Experimental Evidence from India."

Muralidharan, K. and V. Sundararaman. 2009. "Teacher Performance Pay: Experimental Evidence from India." *The Journal of Political Economy* 119(1): 39-77.

Muralidharan, K., Chaudhury, N., Hammer, J., Kremer, M., & Rogers, F. H. 2011. "Is There a Doctor in the House? Medical Worker Absence in India." Working Paper, April 2011.

Muralidharan, K., Das, J., Holla, A., & Mohpal, A. (2016.). "The Fiscal Cost of Weak Governance: Evidence from Teacher Absence in India." *World Bank*, Policy Research WorkingWork Paper 7579.

Muralidharan, K., Niehaus, P., & Sukhtankar, S. (2016). Building state capacity: Evidence from biometric smartcards in India. *American Economic Review*, 106(10), 2895-2929.

Muralidharan, K. and V. Sundararaman. 2011. "Teacher Performance Pay: Experimental Evidence from India." *Journal of Political Economy* 119 (1): 39-77.

Muralidharan, K, and V. Sundararaman. 2013. "Contract Teachers: Experimental Evidence from India. *National Bureau of Economic Research* Working Paper 19440.

Naff, K. and J. Crum. 1999. "Working For America: Does Public Service Motivation Make a Difference?" *Review of Public Personnel Administration* 19(4): 5-16.

Nagavarapu, S. and S. Sekhri. 2016. "Informal Monitoring and Enforcement Mechanisms in Public Service Delivery: Evidence from the Public Distribution System in India." *Journal of Development Economics* 121(C): 63-78.

Naritomi, J. (2016). Consumers as tax auditors. London School of Economics. mimeo.

Neal, D. 2011. "The Design of Performance Pay in Education." *National Bureau of Economic Research* Working Paper 16710.

Neggers, Y. (2018). Enfranchising Your Own? Experimental Evidence on Bureaucrat Diversity and Election Bias in India. American Economic Review, 108(6), 1288-1321.

Niehaus, P and S. Sukhtankar. 2013a. "Corruption Dynamics: The Golden Goose Effect." *American Economic Journal: Economic Policy* 5(4): 230-269.

Niehaus, P., & Sukhtankar, S. 2013b. "The Marginal Rate of Corruption in Public Programs: Evidence from India." Journal of public Economics 104: 52-64.

Oliva, P. Paulina. 2015. "Environmental Regulations and Corruption: Automobile Emissions in Mexico City." *Journal of Political Economy* 123(3): 686

Olken, B. 2006. "Corruption and the costs of redistribution: Micro evidence from Indonesia." *Journal of Public Economics* 90(4-5): 853-870.

Olken, B. 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy* 115: 200-249.

Olken, B. 2009. "Corruption Perceptions vs. Corruption Reality." *Journal of Public Economics* 93 (7-8): 950-964.

Olken, B. 2010. "Direct Democracy and Local Public Goods: Evidence from a Field Experiment in Indonesia." *American Political Science Review* 104: 243-267.

Olken, B. and P. Barron. 2009. "The Simple Economics of Extortion: Evidence from Trucking in Aceh." *Journal of Political Economy* 117(3): 417-452.

Olken, B., J. Onishi, and S. Wong. 2014. "Should Aid Reward Performance? Evidence from a Field Experiment on Health and Education in Indonesia." *American Economic Journal: Applied Economics* 2014, 6 (4): 1–-34.

Olken, B. and R. Pande. 2012. "Corruption in Developing Countries." *Annual Review of Economics* 4: 479-509.

Olson, M. 1993. "Dictatorship, Democracy, and Development." *The American Political Science Review* 87 (3): 567-576.

Olken, B., Sacks, A., Endarso, G., Kelik A; Subandoro, Winoto, A; Wrobel, R. 2018. Indonesia - Long-term impact evaluation of Generasi. Washington, D.C. : World Bank Group. http://documents.worldbank.org/curated/en/481161529900287499/Indonesia-Long-term-impact-evaluation-of-Generasi

Padró I Miquel, G. 2007. "The Control of Politicians in Divided Societies: The Politics of Fear." *Review of Economic Studies* 74: 1259–1274.

Pande, R. 2003. "Can Mandated Political Representation Increase Policy Influence for Disadvantaged Minorities? Theory and Evidence from India." *American Economic Review* 93(4): 1132-1151.

Pande, R. 2011. "Can Informed Voters Enforce Better Governance? Experiments in Low-Income Democracies." *Annual Review of Economics* 3: 215-237.

Pande, R. and C. Udry. 2005. "Institutions and Development: A View from Below." *Economic Growth Center* Working Paper 928.

Papaioannou, E. and G. Siourounis. 2008. "Democratization and Growth." *The Economic Journal* 118(532): 1520-1551.

Park, S.M. and H.G. Rainey. 2008. "Leadership and Public Service Motivation in U.S. Federal Agencies." *International Public Management Journal* 11(1): 109-142.

Parrillo, N. R. (2013). Against the Profit Motive: The Salary Revolution in American Government, 1780-1940. Yale University Press.

Peisakhin, L. and P. Pinto. 2010. "Is Transparency an Effective Anti-Corruption Strategy? Evidence from a Field Experiment in India." *Regulation and Governance* 4(3): 261–280.

Peisakhin, L. 2012. "Transparency and Corruption: Evidence from India." *The Journal of Law & Economics* 55(1): 129-149.

Perry, J.L. 1996. "Measuring Public Service Motivation: An Assessment of Construct Reliability and Validity." *Journal of Public Administration Research and Theory* 6(1): 5-22.

Perry, J.L. and A. Hondeghem. 2008. "Building Theory and Empirical Evidence about Public Service Motivation." *International Public Management Journal* 11(1): 3-12.

Persson, T. and G. Tabellini. 1994. "Is Inequality Harmful for Growth?" *American Economic Review* 84(3): 600-621.

Pathak, Y. and K. Macours. 2013. "Women's Political Reservation, Early Childhood Development and Learning in India." *Young Lives* Working Paper 110s.

Petrovsky, N. 2009. "Review of Rein Taagepera's Making Social Sciences More Scientific: The Need for Predictive Models." *Public Administration* 87(4): 977-980.

Pradhan, M., D. Suryadarma, A. Beatty, M. Wong, A. Gaduh, A. Alisjahbana, and R. Prama Artha. 2014. "Improving Educational Quality through Enhancing Community Participation: Results from a Randomized Field Experiment in Indonesia." *American Journal of Economics*, 6(2): 105-26.

Prendergast, C. 2007. "The Motivation and Bias of Bureaucrats." *American Economic Review* 97 (1): 180-196.

Pomeranz, D. 2015. "No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax." *American Economic Review* 105(8): 2539-2569.

Prillaman, S., 2017. Strength in numbers: How women's groups close India's political gender gap. Working paper.
Raffler, P., Posner, D. N., & Parkerson, D. (2018). The weakness of bottom-up accountability: Experimental evidence from the Ugandan health sector. Unpublished manuscript.

Ramalho, R. 2007. "The Persistence of Corruption: Evidence from the 1992 Presidential Impeachment in Brazil." *World Bank* Working Paper 45250.

Rasul, I. and D. Rogger. 2015. "Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service."

Rauch, J. E. and P.B Evans. 2000. "Bureaucratic Structure and Bureaucratic Performance in Less Developed Countries." *Journal of Public Economics* 75(1): 49-71.

Ravallion, M., D. van de Walle, P. Dutta, and R. Murgai. 2015. "Empowering Poor People through Public Information? Lessons from a Movie in Rural India." *Journal of Public Economics* 132: 13-22.

Ravanilla, N. (2016). Attracting good people into public service: Evidence from a field experiment in the philippines.

Reid, O., & Weigel, J. (2017). Citizen Participation in Corruption: Evidence from Roadway Tolls in the Democratic Republic of the Congo. Working Paper.

Reinikka, R. and J. Svensson. 2004. "Local Capture: Evidence from a Central Government Transfer Program in Uganda." *The Quarterly Journal of Economics* 119(2): 678-70.

Reinikka, R. and J. Svensson. 2005. "Fighting Corruption to Improve Schooling: Evidence from a Newspaper Campaign in Uganda." *Journal of the European Economic Association* 3(2-3): 259-267.

Reinikka, R. and J. Svensson. 2011. "The Power of Information in Public Services: Evidence from Education in Uganda." *Journal of Public Economics* 95(7-8): 956-966.

Ritz, A. 2009. "Public Service Motivation and Organizational Performance in Swiss Federal Government." *International Review of Administrative Sciences* 75(1): 53-78.

Rogoff, K. 2016. *The Curse of Cash*. Princeton: Princeton University Press.

Rossi, M. 2010. "Political Dynasties: Evidence from a Natural Experiment in Argentina."

Rundlett, A., and M. Svolik. 2016. "Deliver the Vote! Micromotives and Macrobehavior in Electoral Fraud." *The American Political Science Review* 110(1): 180-197.

Sadka, J., Seira, E., & Woodruff, C. (2018). Information and Bargaining through Agents: Experimental Evidence from Mexico's Labor Courts (No. w25137). National Bureau of Economic Research.

Scott, J.C. 1969. "Corruption, Machine Politics, and Political Change." *The American Political Science Review* 63 (4): 1142-1158.

Sequeira, S. and S. Djankov. 2014. "Corruption and Firm Behaviour: Evidence from African Ports." *Journal of International Economics* 94(2); 277-329.

Sequeira, S. (2016). Corruption, trade costs, and gains from tariff liberalization: Evidence from Southern Africa. *American Economic Review*, 106(10), 3029-63.

Sexton, R. (2017). The Unintended Effects of Bottom-Up Accountability: Evidence from a Field Experiment in Peru.

Shleifer, A. and R. Vishny. 1993. "Corruption." *The Quarterly Journal of Economics* 108(3): 599-617.

Steijn, B. 2008. "Person-Environment Fit and Public Service Motivation." *International Public Management Journal* 11(1): 13-27.

Stokes, S. 2005. "Perverse Accountability: A Formal Model of Machine Politics with Evidence from Argentina." *American Political Science Review* 99(3):315-325.

Sukhtankar, S. 2012. "Sweetening the Deal? Political Connections and Sugar Mills in India." *American Economic Journal: Applied Economics* 4(3): 43-63.

Suryadarma, D. and C. Yamauchi. 2013. "Missing public funds and targeting Performance: Evidence from and Anti-Poverty Transfer Program in Indonesia." *Journal of Development Economics* 103: 62-76.

Svensson, J. 2003. "Who Must Pay Bribes and How Much? Evidence from a Cross Section of Firms." *The Quarterly Journal of Economics* 118(1): 207-230.

Szabó, Andrea and Gergely Ujhelyi. (2015.). "Reducing Nonpayment for Public Utilities: Experimental Evidence from South Africa." *Journal of Development Economics* 117: 20-31.

Tabellini, G., F. Brollo, T. Nannicini and R. Perotti. 2009. "The Political Resource Curse." *Innocenzo Gasparini Institute for Economic Research* Working Paper 356.

Treisman, D. 2000. "The Causes of Corruption: A Cross-National Study." *Journal of Public Economics* 76: 399-457.

Van Rijckeghem, C. and B Weder. 2001. "Bureaucratic Corruption and the Rate of Temptation: Do Wages in the Civil Service Affect Corruption, and by how Much?" *Journal of Development Economics* 65(2): 307-331.

Vicente, P. 2013. "Is Vote Buying Effective? Evidence from a Field Experiment in West Africa." *The Economic Journal* 124(574): F356-387.

Visaria, S. 2009. "Legal Reform and Loan Repayment: The Microeconomic Impact of Debt Recovery Tribunals in India." *American Economic Journal: Applied Economics* 1(3): 59-81.

Wantchekon, L. 2003. "Clientelism and Voting Behavior: Evidence from a Field Experiment in Benin." *World Politics* 55(3): 399-422.

Wantchekon, L., G. Lopez-Moctezuma, T. Fujiwara, C. Lero, and D. Rubenson. 2015. "Policy Deliberation and Voting Behavior: A Campaign Experiment in the Philippines."

Wei, S. J. 2000. "How Taxing is Corruption on International Investors?" *The Review of Economics and Statistics* 82(1): 1-11.

World Bank. 1992. "Governance and Development." Washington, D.C.

World Bank. 2002. "Building Institutions for Markets." New York: Oxford University.

World Bank. 2007. "Strengthening World Bank Group Engagement on Governance and Anticorruption." Washington, D.C.

Xu, Y. and Y. Yao. 2015. "Informal Institutions, Collective Action, and Public Investment in Rural China." *American Political Science Review* 109(2): 371-391.

Yang, D. 2008."Integrity for Hire: An Analysis of a Widespread Customs Reform." *Journal of Law and Economics* 51(1): 25-57.