Institutions and economic development: Taking stock and looking forward

Ernesto Dal Bó, Frederico Finan, and Yuen Ho

This version: October 18, 2021

1. Introduction

This paper serves as a companion paper to the Economic Development and Institutions' (EDI) White Paper "At the Intersection: A Review of Institutions in Economic Development" (Dal Bó and Finan, 2016). The White Paper reviewed nearly 200 publications from economics and political science on the role of institutions in development, drawing from experimental and quasi-experimental evidence whenever possible. The goal of the White Paper was threefold: (i) to present a framework for reviewing and synthesizing the existing evidence on how legal and political institutions affect development outcomes; (ii) to distill lessons learned from the literature; and (iii) to present open questions in each topic that defined the research frontier at that time.

This Green Paper revisits the themes and research priorities identified by the White Paper five years later. The purpose is to assess how the research funded by the EDI program contributes to our knowledge of how institutions affect development, and where the research frontier lies today. Whenever possible, we synthesize lessons across funded studies in cross-cutting analytic themes to identify the mechanisms that underlie observed effects.

In terms of methodology, much of the research funded by EDI Research Area 3 (RA3) is based upon randomized controlled trials (RCTs), although we also discuss observational and quasi-experimental findings from funded studies where relevant. Thus, it is important to note that this is not an exhaustive review of the literature that has been published since 2016 on the topic of institutions and development. Relatedly, given that much of the research activities funded by EDI took place between 2017 – 2021, many of the findings discussed in this Green Paper are based on working papers that have not yet undergone the peer-review publication process. As such, these findings are preliminary and may change over time. We distinguish which papers we review are working papers and which are published throughout. Lastly, we emphasize that much of the research funded by EDI is still ongoing and will produce results well after this paper is circulated and read. Therefore, the contributions reviewed here constitute a fraction of the knowledge that EDI has helped create.

As in the 2016 White Paper, we conceptualize institutions as North's (1991) 'rules of the game,' which comprise a wide range of social and economic, legal and political, and normative and procedural rules, as well as incentive structures and organizations. We organize our review into two main sections, one discussing political institutions and the second discussing legal institutions.

The evidence reviewed in this paper illustrates the substantial progress we have made in our understanding of the role of institutions in development and growth. While this new evidence sheds light on many of the open questions originally raised in the White Paper, often it also raises new questions and pushes the research frontier forward. It is clear that this is an active area of research where our knowledge should, and hopefully will, continue to evolve over time.

2. Political institutions

The White Paper identifies political institutions and development as an overarching research priority. This involves investigating the links between political institutions and growth, both in terms of causal effects as well as the details of which institutions matter and how. Broadly speaking, political forces impact outcomes where there is: i) the *will* to shape outcomes in a particular way, enacted via institutions that affect representation and accountability; and (ii) the *means* to do so, which depends on state capacity.

2. A. Representation and accountability

Political institutions may affect policies and welfare through two main channels: (i) representation, or whose preferences influence decisions; and (ii) accountability, or how decision makers are held responsible for their actions. Electoral rules, information and transparency, and political norms and cultures are three main elements that matter for representation and accountability.

2. A. i. Electoral rules and information and transparency

The findings from the White Paper suggest that electoral rules affect the composition of the electorate, and ultimately policy. On the representation channel, electoral rules affect voter turnout (Pande, 2003; Chattopadhyay and Duflo, 2004; Fujiwara, 2015) and political selection (Beath et al. 2016). Changes in the composition of the electorate, in turn, affect which policies are selected and enacted and, subsequently, citizen welfare (Fujiwara, 2015; Chattopadhyay and Duflo, 2004; Pande, 2003; Olken, 2010). Through the information channel, the reviewed literature shows that information matters for a wide variety of political behaviors, including turnout and preferences (for examples see DellaVigna and Kaplan, 2007; Kendall, Nannicini, and Trebbi, 2015), electoral accountability (Ferraz and Finan, 2008), and political strategy (Bidwell, Casey, and Glennerster, 2020).

While the existing literature establishes that electoral rules matter for representation and accountability, it is less clear whether there is a tradeoff between broadening representation and politician quality. On the one hand, selection by experts or officials, "elites," may be beneficial if their preferences are representative, or if they are better informed about candidate quality. On the other hand, elites may have preferences that are not representative of the broader population, or they may value traits unrelated to performance in office, such as loyalty, corruptibility, or willingness to pay to hold office.

Casey, Kamara, and Meriggi (AER 2021) shed light on these questions by partnering with both traditional political parties in Sierra Leone to experimentally vary the scope for voters to influence parliamentary primaries. In the status quo, party officials chose among potential candidates with no direct participation by voters. For a random selection of races for MP seats, in contrast with the status quo, the political parties hosted publicly broadcasted conventions where candidates presented their qualifications and policy platforms. The researchers then conducted representative opinion polling of voter preferences over candidates and shared this information with party officials. While parties were still free to choose a candidate regardless of voter preferences, the intervention significantly increased whether parties chose the voters' first choice candidate, from 38 to 61 percent. Moreover, candidates chosen under the new selection method compared to the status quo had stronger records of previously providing public goods, which is something voters valued. Overall, the results indicate that broadening representation benefited voters, as not only were parties more likely to select their preferred candidates, but those candidates had stronger records of public goods provision.

In this case, party officials were willing to voluntarily accommodate voter preferences, which raises the question of why they often did not select the most popular candidate in the status quo. The results suggest officials lacked *information* on which candidates were most popular absent the intervention. Survey results showed that "90 percent of constituency-level party officials presume local voters share their first choice over aspirants, when voters in fact only agree with them about one-half of the time; and in one-third of races, not a single party official accurately guessed which aspirant ranked first among voters" (Casey, Kamara, and Meriggi, AER 2021). Political parties likely faced supply-side barriers to collecting this information themselves as they lacked affordable ways of polling voters. These results highlight the importance of understanding where barriers to acquiring information arise, either on the side of voters or politicians, which can inform more targeted design of information interventions.

These findings also raise new questions. Many of the races where parties were willing to experiment were in their political strongholds, where general election risks were minimal. It is still an open question of how information affects outcomes when politician and voter preferences are in direct conflict. While the results shed light on how broadening representation affects political selection on the intensive margin -- that is, which candidates get selected conditional on running -- they raise a new question of what effects would be on the extensive margin, that is, who chooses to run for office in the first place. Since voter preferences were collected from a representative sample of voters, which may not reflect actual turnout in a direct vote primary, understanding effects under selective turnout is also an important area for future research.

Open questions

- How do electoral rules affect political selection on the extensive margin, that is who decides to run for office?
- Is there a tradeoff between broadening representation and competence once effects on endogenous turnout are taken into consideration?

- How do different selection rules affect performance in office, conditional on being elected?
- What are the supply and demand factors that affect information provision, acquisition, and use?
- How does information disclosure affect accountability when voter and politician interests are in direct conflict?

2. A. ii. Political norms and culture

Societal outcomes depend both on formal institutions and on equilibrium behavior that is affected by culture and norms. These informal institutions may greatly amplify or mitigate the effects of formal institutions. Evidence reviewed in the White Paper shows that civic education campaigns can alter modalities of political engagement and political outcomes (Vicente, 2014; Aker, Collier, and Vicente, 2017; Collier and Vicente, 2013) and education curricula shapes political attitudes (Cantoni et al., 2017). Given gender bias, policies that reserve a certain number of leadership positions for women can improve perceptions of women as leaders (Beaman et al., 2009).

Contributing to this line of research, Prillaman (ongoing) is conducting a field experiment in India to study whether gender-oriented civics education can increase women's political participation in the face of gender-biased social norms. Within existing women's self-help groups, women receive training over several months on gender norms in society, their political rights and entitlements, gender biases in political institutions, and avenues for political participation, including conducting visits to existing local institutions. Across 225 villages, 150 villages were randomly selected to receive the women's civics education training while the remaining 75 villages served as a comparison group. Researchers will study the impact of the civics training program on women's and men's gender views; political outcomes including participation in local governance, voter turnout, and running for office; and non-political dimensions of empowerment such as household decision-making, self-efficacy and well-being, and labor market status and income. If there is a positive effect on women's participation and electoral representation, researchers will further study the effects of increased female political participation on local governance and economic development. This project is ongoing and results are forthcoming.

Open questions

- How do political norms interact with formal institutions?
- In the face of gender-biased social norms, how can we effectively increase women's political participation and electoral representation?
- What are the effects on political participation and accountability of specific cultural traits, such as trust, respect for others, and individualism versus collectivism?
- What institutions can limit clientelism, patronage, and dynastic politics?

2. B. State capacity

Understanding what makes governments work better is a central part of a pro-growth, institutions-centered research agenda. If one conceptualizes the state as operating a production function of public goods, the inputs can be divided into three categories: personnel, resources, and procedures.

2. B. i. State personnel
2. B. i. a. Selection

The most basic aspect in personnel policy is selecting the individuals who will become public employees. Since governments are responsible for providing public goods, recruiting and employing applicants with high prosocial motivation may be especially beneficial. Indeed, a large empirical literature reviewed in the White Paper suggests that prosocial motivation is associated with higher levels of performance in the public sector in a wide variety of contexts (for examples see Perry and Hondeghem, 2008; Petrovsky, 2009).

The White Paper also describes concerns in the literature that there may be trade-offs between candidate quality and motivation and integrity. Theoretically, while offering higher wages may attract higher quality candidates, it may also attract individuals who are more corruptible or care less about the mission. The evidence reviewed in the White Paper examined the impact of unconditional higher wages and career incentives on selection into public service jobs, with mixed results. In Mexico, Dal Bó, Finan, and Rossi (2013) find that higher wages attract higher quality and publicly motivated applicants and help fill vacancies in less desirable locations. Similarly, Ashraf et al. (2020) find that emphasizing career incentives for community health worker positions in Zambia resulted in higher quality hires with equal prosociality, leading to better performance on the job and improved health outcomes. In contrast to these two studies, Deserranno (2019) finds that for village health promoter roles in Uganda, while the expectation of higher earnings helped fill vacancies, it also discouraged prosocial applicants, resulting in worse performance and retention.

New evidence from Leaver, Ozier, Serneels, and Zeitlin (AER 2021) sheds light on the effects of performance pay on selection and performance for Rwandan primary school teachers. Ex-ante, the effects of performance pay on selection could go either way: performance pay may attract those with lower intrinsic motivation, who may perform worse on the job, or it may attract individuals who anticipate performing well. The researchers first randomized at the labor market level whether teachers were recruited under performance pay or fixed wages, and then re-randomized at the school level which contract teachers actually experienced. This research design allowed them to separately identify selection effects, or how performance pay affects who decides to apply for the job, and effort effects, or how performance pay affects teacher effort on the job. The results suggest recruiting with performance pay did not change the quality of applicants or hires but did result in recruits with lower prosocial motivation. Despite the change in selection, teachers recruited under performance pay were as effective, if not more effective, in promoting learning compared to teachers recruited under a fixed wage.

These results are consistent with performance pay attracting teachers who can deliver better outcomes. Overall, the results suggest performance pay does change selection, but these changes do not negatively affect student learning and, if anything, are beneficial.

These findings also raise new questions. First, while attracting teachers with lower prosocial motivation did not negatively affect measured student learning outcomes, what are the effects on other, not measured, dimensions, such as students' own prosocial motivation or soft skills? Second, while this study focused on effects within a two year time frame, what would the effects of a long-term performance pay policy be on selection into the teaching profession more generally? Third, screening mechanisms that school districts could adopt in principle, such as teacher training scores, did not positively select on the attributes associated with better learning outcomes, highlighting the continued challenge of identifying the key correlates that could be used to attract and screen applicants.

2. B. i. b. Monitoring and Incentives

The White Paper identified four important questions surrounding incentive schemes in the state administration. The first is whether incentive schemes such as pay for performance can improve performance along the intended dimension. The second is whether using more high-powered incentive schemes triggers distortions along dimensions that are not incentivized, or "multi-tasking." The third is whether principals tend to design schemes optimally, given the trade-offs with providing incentives or interactive effects between policies. A fourth question is whether extrinsic incentives crowd out intrinsic incentives. The evidence reviewed in the White Paper shed light on the first two questions, namely that performance pay tends to improve service delivery (Banerjee et al., 2008; Basinga et al., 2011) and tax collection (Khan, Khwaja, and Olken, 2016). New evidence from EDI funded projects further shed light on the third and fourth questions.

In terms of incentivizing effort, Leaver et al. (AER 2021) find that performance pay significantly improved pupil learning compared to incentivizing teachers with fixed wage contracts. Specifically, by the second year of the study, teachers who received a performance pay contract produced student learning that was 0.16 standard deviations higher than student learning under fixed wage contracts. The results suggest that learning gains under performance pay were most plausibly driven by increased teacher effort, particularly in teacher presence and pedagogy. Taken together, the results corroborate the existing evidence on the effectiveness of performance pay to improve service delivery and suggest that fears of extrinsic incentives crowding out intrinsic motivation are overstated.

In addition to financial incentives, non-financial incentives such as transfers can be a cost-effective way to motivate personnel. Cohen, Dal Bó, Finan, Omala, and Schönholzer (ongoing)[1] are implementing a RCT to study whether incorporating teacher preferences into the existing transfer protocol can improve teacher motivation and performance in Uganda. In

[1] This research was supported by EDI Research Area 4, but is included in this discussion as a randomized evaluation supported by EDI on related topics.

the status quo, primary school teachers in Uganda are transferred to a new school every five years. These transfers are undirected and do not take teacher preferences into account, exacerbating issues with low-morale, low engagement, and high absenteeism among teachers. Researchers will study a simple, low-cost modification to the existing transfer system that rewards high-performing teachers with their choice of future placement. This project is ongoing and results are forthcoming.

One important dimension of designing optimal personnel policy is considering how different incentive tools, such as performance pay, promotions, and transfers, interact with one another. Deserranno, Kastrau, and León-Ciliotta (Working Paper 2021) explicitly study the interactions between meritocratic promotions and steeper pay progression among community health workers (CHWs) in Sierra Leone. In the status quo, promotion decisions were decided by the local health authority and were perceived, correctly, to be largely based on personal connections rather than productivity. Furthermore, only 30% of CHWs accurately guessed the amount of increased pay associated with promotion. To study how these two factors interact, CHWs were randomly assigned to either a new meritocratic system that promoted the best performing CHWs, to receive accurate information about the pay associated with promotion, both treatments, or neither. The results suggest important complementarities exist between meritocracy and pay progression. Meritocratic promotions lead to higher productivity, but only when workers believed they would be sufficiently compensated upon promotion. Likewise, higher perceived pay raises improved worker productivity, but only when promotions were meritocratic. Importantly, when promotions were non-meritocratic, a perceived steeper pay progression *reduced* worker productivity, most likely through negative morale effects. In addition to highlighting the importance of how different incentive tools interact, sometimes with unintended consequences, these findings further strengthen the evidence that extrinsic incentives do not significantly crowd out intrinsic motivation.

In a companion paper, Deserranno (r), Caria (r), Kastrau (r), and León-Ciliotta (Working paper 2021) examine interactions between performance pay at different levels of a bureaucracy. Again working with community health workers (CHWs) in Sierra Leone, performance pay was randomly assigned to either only CHWs, only their supervisors, or to be shared equally between CHWs and their supervisors. Performance bonuses were linked to the number of CHW services provided and reported, including routine household visits, pre-natal and post-natal check-ups, and the diagnosis and treatment of diseases such as malaria, diarrhea, and pneumonia. Supervisors received bonuses based on the number of health services the CHWs under their supervision performed. In practice, supervisors could affect the quantity and quality of health services provided by training, advising, and monitoring health workers in their teams, as well as improving awareness and trust of CHW services in the community. The results suggest that all three performance pay schemes were effective in increasing the number of services provided by CHWs. However, incentivizing both workers and their supervisors simultaneously generates an increase in health visits that is 61% larger than the impact of incentivizing either workers or their supervisors in isolation, which in turn leads to a larger improvement in health outcomes. . The results are explained by the existence of effort

complementarities between CHWs and their supervisors, where the return to worker effort increases in the amount of effort provided by the supervisor and vice versa.

Within the context of a large Indian bureaucracy, Dodge, Neggers, Pande, and Troyer-Moore (Working Paper 2021) examine similar themes by randomizing which levels of the bureaucracy received improved information tools for monitoring. Specifically, researchers provided bureaucrats access to PayDash, a mobile app which made information from the digitized MGNREGA wage payment system more accessible by allowing bureaucrats to see the status of pending payments, identify responsible officers, and immediately follow up by phone or text. Access to PayDash was randomly provided to either the district level, the sub-district level, or both levels simultaneously. The results show PayDash reduced payment processing times by 11% (or roughly 1.2 days) on average, and by 17% (or roughly 3.4 days) in the areas with the highest wage payment delays. Providing PayDash to two bureaucratic levels led to no greater improvements than giving access to only a single level. The substitutability is likely due to officials sharing information with each other to improve performance, not just to monitor staff, suggesting information costs were a key binding constraint in this context. As a result of receiving access to PayDash, the bureaucracy also significantly substituted away from using officer transfers as a costly, alternative incentive tool.

Taken together, these studies highlight the importance of understanding how different monitoring and incentive tools interact with one another when designing optimal personnel policy. Explicitly modeling and estimating such interactions remain an important avenue for new research.

Open research questions

- How do different incentive schemes affect self-selection, i.e. individuals' early career decisions to choose a particular occupation?
- While the tradeoff between extrinsic and intrinsic motivation does not seem acute empirically, how does performance pay affect other outcomes related to soft skills, norms, and aspirations?
- When does performance pay lead to multi-tasking distortions?
- How does the interaction between different incentive and monitoring tools affect the design of optimal personnel policy?

2. B. ii. State resources

Taxes are the main way for the state to obtain resources. Therefore, how to improve tax compliance in settings with weak institutional capacity is a priority area for research. The evidence reviewed in the White Paper suggests tax compliance depends on various aspects of the tax regime, including third-party reporting (Pomeranz, 2015; Kleven et al., 2011), tax levels (Kleven et al. 2011), the salience of tax penalties and audits (McGraw and Scholz, 1991; Blumenthal, Christian, and Slemrod, 2001; Fellner, Sausgruber, and Traxler, 2013; Castro and Scartascini, 2015); and the social norms around compliance (Perez-Truglia and Troiano, 2018).

Enforcing taxation can be especially difficult in contexts where many transactions are informal or not observed by the government. The value-added tax (VAT), which adds tax to a product at every point in the supply chain, facilitates tax collection by generating paper trails on transactions between firms (Pomeranz, 2015; Kleven et al., 2011). Under a VAT, firms have incentives to request receipts from their suppliers to lower their own tax liability, generating third-party information that can be used by the tax authority to catch evasion. Indeed, the majority of countries around the world have some form of VAT, which represents the largest source of tax revenue in many developing countries. Despite these self-enforcing properties, VAT systems can still be plagued by tax evasion and imperfect compliance. Two new non-experimental studies, resulting from pilot and data collection activities to design a RCT, shed light on potential sources of noncompliance within a VAT system and strategies to identify and reduce evasion.

Using data on VAT filing returns from the state of Delhi in India, and exploiting naturally occurring variation in policy, Luksic and Mittal (Working paper 2019) study the effects of policies that reduce reporting requirements for smaller firms. They find that size-dependent filing policies lead to bunching of firms' reported turnover around the filing thresholds. The bunching disappears when thresholds are removed. Simple graphical analysis suggests the bunching behavior likely stems from firm underreporting. The reported revenue losses due to the bunching behavior by firms is substantial, ranging from 2-8% of the bunchers' total yearly VAT contributions. However, a back of the envelope social welfare analysis suggests the benefits of less frequent reporting to smaller firms more than offset the revenue losses to the tax authority. Taken together, the results suggest that frequent reporting requirements generate high compliance costs that can distort firm behavior, with unclear benefits to the tax authority from more frequent information and payments. Tax authorities can improve welfare and reduce evasion by simplifying reporting requirements.

VAT systems also create a demand for fraudulent firms that sell fake receipts, which genuine firms can use to evade taxation by generating false paper trails. A key challenge in improving tax compliance is to identify such fraudulent firms cheaply and reliably. Using existing data from the Delhi tax authority in India, including the universe of tax returns and the results of past audits, Mittal, Reich, and Mahajan (SIGCAS, 2018) develop a machine learning (ML) algorithm to identify bogus firms. The data, while typical of the tax compliance data available in many developing countries, raises significant challenges compared to the standard datasets used to train ML algorithms. The authors develop a ML algorithm that addresses these challenges, which include: i) selective labels, or not being representative of the population; ii) one-sided labels, where the tax authority only records bogus firms that were inspected and caught, but not firms that were inspected and found to be legitimate, and iii) the fact that the authority never identified all bogus firms, making the standard classification approach unviable. Using simulations, they estimate that this ML tool could prevent up to US$15-45 million worth of tax fraud by identifying fraudulent firms more quickly than the status quo system of manually targeted audits. In future work, the research team plans to conduct a RCT to test the field

efficacy of this ML tool and to study the extent to which firms substitute their fraudulent activity to new or undiscovered bogus firms.

Property taxes are another form of instrument which can facilitate compliance, at least in theory, due to their tangible and immobile tax base. Cogneau, Gurgand, Knebelmann, Pouliquen, and Sarr (Working paper 2021) document the property tax gap, or the difference between the property taxes paid and the property taxes owed by law, in Dakar, Senegal. In Senegal, the national tax administration (Directorate General of Taxes and Domains) is responsible for maintaining the property valuation roll while the Treasury is responsible for distributing tax notifications, managing payments, and enforcing compliance. To keep the valuation roll up to date, property owners are required to make annual declarations of their properties and corresponding values. Owners also need to make tax payments in person at the Treasury offices. Working in partnership with the cadaster department of the national tax administration, researchers helped digitize, extend, and edit the existing cadastral dataset using drone and satellite images. They also used administrative data on tax assessments and tax revenue; conducted a survey of property owners, including property valuations by real estate experts; and collected data from Treasury agents on the outcomes of tax notifications. They estimate that less than 20% of property owners are in the tax net, and only 9% of potential tax is being collected. Approximately 75% of foregone revenue is on the extensive margin, due to the exclusion of taxpayers, while 16% is on the intensive margin, due to actual tax payments being lower than what they should be based on market property values. Based on suggestive evidence, they hypothesize that the most important impediments to effective property taxation in this context are i) the reliance on voluntary owner declarations, which leads to low compliance; ii) an incomplete and outdated valuation roll; and iii) misaligned incentives between different central government departments and between central and local governments.

Comparing the observed distribution of the tax burden to the theoretical one under full compliance, Cogneau et al. (Working paper 2021) find that "weak enforcement leads to a tax profile that is more regressive than what is provided for in the legal framework." The gap between the effective rate and the theoretical rate is significantly wider, and the effective tax rate faced by payers is significantly lower, for the top deciles of property value. In other words, conditional on paying taxes, taxpayers in lower deciles of property value are paying more than they should, while taxpayers in higher deciles are paying less than they should. Under simulations, they estimate that the tax administration would collect four times as much revenue for the same number of payers if enforcement was targeted at the highest valued properties. These results highlight the importance of understanding how laws are enforced de facto versus what is provided for de jure, as this can have important implications for the realized welfare impacts of a given policy.

Complementing this tax gap analysis, Cogneau, Gurgand, Knebelmann, Pouliquen, and Sarr (ongoing) are implementing a field experiment in partnership with Senegal's tax administration to study the introduction of a new modernized property tax management system. The modernized system uses a new application - developed collaboratively throughout the project -

that automates field data collection, the computation of tax liabilities, and the production of tax notifications using digitized cadastral and address information. Specifically, 194 cadastral sections have been randomly assigned to either continue with the status quo system based on the existing valuation roll or to receive a comprehensive fiscal census along with the new data management system. Within the treatment group, cadastral sections will be cross-randomized to either conduct property valuation using the status quo discretionary method or a rule-based method integrated into the software that estimates property value based on visible characteristics. Researchers will measure the effects of the new system on fiscal capacity, the relative advantages of a rule-based versus discretionary valuation method, the effects of increased taxation on local governance dynamics, and the extent to which the tax burden is shifted to tenants. Results from the field experiment will inform the scale-up of the new tax management system nationwide. This project is ongoing and results are forthcoming.

Open questions

- What areas of taxation are most affected by corruption?
- How much do improvements in enforcement generate displacement in wrongdoing?
- Machine learning approaches seem promising at detecting underreporting, and they can be low cost provided the data exists. Generating the data, however, can be expensive in low state-capacity areas, raising a chicken-and-egg problem. What are low-cost ways of generating/improving data and what are low-cost (in terms of tax-payer compliance) ways to raise tax revenue?
- What are the tradeoffs between effectiveness and transparency when adopting sophisticated data analysis tools to improve tax collection?
- What are the distributional consequences of weak or heterogenous enforcement?

2. B. iii. State procedures

The allocation of decision-making authority to different agents and levels of a hierarchy is one of the key parameters determining organizational performance. A central debate in this area is whether it is better to have strict rules and monitoring or to give bureaucrats discretion and autonomy. On the one hand, leveraging the local knowledge and creativity of individual bureaucrats may improve productivity. On the other hand, bureaucrats may abuse discretion and shirk their responsibilities, engage in corruption, or make errors in judgement.

Evidence reviewed in the White Paper suggests autonomy, compared to incentives or increases in resources, can improve service delivery. In the Nigerian Civil Service, Rasul and Rogger (2018) find that management practices that grant greater autonomy to bureaucrats are correlated with project completion rates, while practices related to incentives and monitoring are negatively correlated. Corroborating these descriptive findings with experimental evidence, Duflo et al. (2018) find that increasing the frequency of random, non-discretionary environmental inspections did not significantly reduce pollution relative to less frequent, discretionary audits in India. Results from structural estimation suggest that regulators use their

discretion to target the most extreme polluters, while random inspections mostly identify smaller violators, making discretionary targeting potentially more cost-effective.

While the previous literature compared discretionary targeting to random targeting, how would discretion compare to a data-driven, rules-based approach? This is precisely what Bachas, Brockmeyer, Ferreira, and Sarr (Working paper 2021) study in the context of selecting firms for tax audits in Senegal. Using data from self-reported tax declarations, as well as third-party data from customs, procurement contracts, and transacting partners, the research team constructed rule-based scores for firms that reflected their risk of tax evasion. Firms with higher risk scores had a higher number of observed discrepancies, where information reported in their tax returns did not match information from third-parties, and anomalies, where their displayed behavior was unusual relative to similar firms. Since the tax administration wanted the risk-score to be transparent and easily understood, the risk-score algorithm systematically analyzed the available data but did not use machine learning or fine-tuned parametrization. Each inspector selected 45% of their audits using the traditional discretionary method, an additional 45% of audits were assigned based on the risk-score, and the remaining 10% were randomly selected. To study whether better information can improve the effectiveness of discretionary targeting, a subset of audits were randomly selected to include information on the top three risk indicators detected by the risk-score, the top three risk indicators plus third party data, or no additional information. The research team is in the process of acquiring data on audit results for their final analysis. Nevertheless, preliminary results suggest that "discretionary methods select larger firms on average and uncover equivalent evasion rates as the risk-score algorithm, thus outperforming it in terms of fines." The information treatments do not seem to significantly affect audit outcomes. Tax inspectors were more likely to investigate cases selected by the discretionary method, compared to those selected by the risk-score method, suggesting the buy-in of personnel to use new technology is key for that technology to be effective.

Azulai, Rasul, Rogger, and Williams (2020) take as given that empowering bureaucrats is desirable and study the effects of a new training focused on improving work culture within the Ghanian Civil Service.[2] The training aimed to improve the skills, motivation, and norms around civil servants taking individual initiative to identify and implement work process improvements, in contrast to a more hierarchical status quo. Civil servants were randomly assigned to either receive the new training or a pre-existing productivity training. To study whether coordinated group action is necessary for changing organizational norms, researchers also cross-randomized whether bureaucrats received the training with their entire work team or individually with trainees from other departments. The results show that while both treatments improved knowledge in the short-term, individual-level trainings also improved organizational culture, such as teamwork and fostering new ideas, and team performance, including task completion and the quality of administrative processes. In contrast, the team-based trainings had no significant effects on culture or performance, potentially because they reinforced existing team culture or diffused responsibility for leading change. These patterns complement the evidence

[2] This research was supported by EDI Research Area 4, but is included in this discussion as a randomized evaluation supported by EDI on related topics.

that greater autonomy and discretion can improve bureaucratic performance, and highlight the importance of targeting interventions at individuals who can champion reforms.

Open questions

- To what extent do changes in managerial practices affect public goods provision?
- What are the effects of decentralization on public service delivery and accountability?
- What level of autonomy to bureaucrats is optimal for delivering high quality public services?
- How does improving the quality of information available to bureaucrats affect the costs and benefits of bureaucratic discretion?

3. Legal institutions

Personal security and the security of claims to property are thought to be a basic driver of investment and development. Indeed, it is well documented that rule of law correlates with economic growth (for a review, see Haggard, MacIntry, and Tiede, 2008). Rule of law can be seen as a function of legal institutions, broadly understood. For legal institutions to function well, four elements are required. First, legally sanctioned rights must be specified. Second, the beneficiaries or holders of such rights must be aware of their entitlement and be able to access the means to have their rights upheld. Third, there must be processes for mediating disputes, such as through courts. Finally, conditions must exist so that rights and court rulings can be enforced, such as through police and auditors, and informal enforcement through social norms. As in the White Paper, we organize this section according to these four elements.

3. A. Legally sanctioned rights

Theoretically, property rights can reduce expropriation risk and improve access to credit, since land can be used as collateral for loans (Besley and Case, 1995; De Soto, 2000). These improvements can in turn enable households to invest more in agricultural production and assets and pursue off-farm economic opportunities. The quasi-experimental evidence reviewed in the White Paper shows property rights increase investments in housing and land productivity, primarily through improving tenure security rather than through relaxing credit constraints (Field, 2005; Ali et al., 2014). The White Paper therefore identified a need for experimental evidence on whether credit access plays a key role as theorized.

Ghatak, Goldstein, Habyarimana, Montalvao, O'Sullivan, and Udry (ongoing) address this evidence gap by implementing a randomized experiment to study how improved access to land titling, particularly for women, affects household investment decisions in Uganda. Households were randomly assigned to receive financial and information incentives to promote joint spousal land titling. To study whether credit constraints prevent households from investing in opportunities unlocked by formal titling, women were cross-randomized to receive cash transfers of approximately US$200. The research team will study effects on household investment, productivity, and welfare, and whether complementarities exist between improving

land rights and relaxing liquidity constraints. The project is ongoing, and results are forthcoming.

Open questions

- Can property titling relax credit constraints?
- If complementary reforms are needed, what are they?
- What are the regulatory barriers for firm growth from land titling and zoning?

3. B. Awareness and access

Even when people have specific legal rights, making individuals aware of their rights, and the means they have to secure those rights, can be important for closing the gap between *de jure* rights and *de facto* practice. Some ways of expanding legal access that have been studied in the literature are *pro bono* representation and improving access to legal assistance. Indeed, evidence reviewed in the White Paper shows that the presence of paralegals in treated villages helps promote knowledge of the rights of women (Mueller et al. 2018) and *pro bono* legal assistance tends to improve outcomes for beneficiaries (Frankel et al. 2001; Greiner and Pattanayak, 2012; Sandefur and Siddiqi, 2015; Blattman, Hartman, and Blair, 2014).

Field and Vyborny (Working paper 2020) study another aspect of improving legal awareness and access, namely ensuring that those responsible for implementing the law know the details of the law on the books. In 2015, the state of Punjab in Pakistan passed a set of legal reforms to strengthen women's rights in marriage and divorce. However, most of the 60,000 marriage registrars in the province were unaware of these reforms, or of women's legal rights in marriage more generally. In 2017, the government initiated the first mandatory training of marriage registrars in partnership with the researchers "to inform registrars of these key provisions and punishments for non-compliance." Researchers exploited three sources of variation: i) a randomized phase-in design in part of the province to measure the immediate effects of the training on registrar knowledge and intentions, ii) a randomized experiment in another district that randomly assigned whether union councils received the training to measure long term outcomes on marriage contracts, and iii) quasi-experimental variation in the remaining roll-out of the training to examine province-level effects. The results suggest "the training increased registrar knowledge of basic provisions of the law, dramatically reduced procedural violations, and helped to ensure women are aware of the terms of the contract they are signing."

Open questions

- How do improved legal outcomes translate into improved welfare?
- Improving knowledge of laws is a necessary but not sufficient condition for ensuring effective rights. What complementary interventions are needed, perhaps on the citizen/demand side, to bring *de facto* and *de jure* rights more in alignment?
- Does broader legal access affect productive activities?
- Does a more extensive "rule of law" create aggregate (i.e., general equilibrium) effects?

3. C. Courts

The evidence reviewed in the White Paper largely draws on observational data and structural estimation to study the effects of incentives and selection of judges on the functioning of courts. The reviewed evidence suggests judges selected via elections versus appointments face different incentives, with implications for sentencing (Berdejó and Yuchtman, 2013). Importantly, the visibility of sentencing through media coverage matters (Lim, 2015; Lim and Snyder, 2015; Lim, Snyder, and Strömberg, 2015). Judge identity also affects sentencing: evidence from Israeli courts shows that verdicts on Arab defendants vary significantly depending on whether there is at least one Arab member in the panel of judges (Grossman et al., 2016).

Contributing to this line of work, Ash, Asher, Bhowmick, Bhupatiraju, Chen, Devi, Goessmann, Novosad, and Siddiqi (Working paper 2021) study whether judges deliver more favorable outcomes to defendants who share their same gender or religion. The authors construct a novel dataset on over 80 million legal case records from India from 2010-2018 and exploit quasi-random assignment of judges to cases and changes in judge cohorts. Given the large sample of cases, the authors precisely estimate zero effects of gender- or religion-based bias, ruling out effect sizes that are one fifth of those reported in the prior literature. The authors find a correlation between the effect sizes and the standard errors reported in published papers using similar identification strategies, suggesting previous estimates of ingroup bias among judges are likely overestimated due to publication bias. The results do, however, suggest small positive ingroup bias by caste, which is likely underestimated as the authors are only able to match judge and defendant caste if they share the same last name. While judges appear to exhibit some bias by shared caste, given the limitations in identifying the caste membership of judges and defendants, this remains an area for future work.

While the papers discussed thus far focus on the incentives of judges, an important open question in this area is how other types of court personnel impact the effectiveness of the judiciary.

Sadka, Seira, and Woodruff (Working paper 2020) shed light on this question by highlighting the importance of agency issues within the lawyer-plaintiff relationship. In partnership with the Mexico City Labor Court, the researchers randomly varied whether plaintiffs in active or not-yet-started cases received personalized outcome predictions generated from machine learning models. In the status quo, settlement rates are low and case durations are long. Plaintiffs seem overconfident about their chances of winning their lawsuit and have limited knowledge of the law, which private lawyers potentially exploit by pursuing cases when workers would be better off with a settlement. The results show that providing personalized outcome predictions significantly increases settlement rates, largely from cases that the plaintiff would have lost. The information treatment improves welfare for the average worker, as measured by the net present value of payouts and an improved ability to pay bills. There are also administrative benefits to the court, as the increased settlement rate reduces the case backlog and the administrative costs associated with litigation.

Importantly, these effects are only realized when the worker was present to receive the information directly, suggesting lawyers did not convey information about the opportunity to settle to their clients. The results show that "the incentives of the lawyers, who have information advantages, do not always align with those of their clients, even though private lawyers almost always receive a share of the award collected by their plaintiffs" (Sadka, Seira, and Woodruff, Working paper 2020). The results highlight how agency issues contribute to inefficiencies in institutional function and suggest that testing interventions that help citizens assess the quality of lawyers is an important area for future research.

Open questions

- What factors affect the observed variation in the level of in-group bias among judges across contexts?
- What policies and incentives can reduce agency issues in the lawyer-plaintiff relationship? Relatedly, what interventions can help citizens assess the quality of lawyers and the legal representation they receive?
- Beyond judges and lawyers, what is the role of staff and organizational support in producing timely outcomes, and how can support functions be improved?
- What is the impact of judicial quality on economic activity?

3. D. The quality of enforcement
3. D. i. The role of police and state capacity

The quality of enforcement is tightly related to the functioning of the police, judiciary, inspector and auditor cadres, as well as regulatory authorities. The evidence reviewed in the White Paper focused on the role of incentives, accountability, selection, monitoring, and information. The evidence suggests that only a subset of interventions work at improving policing, and how interventions affect the power of implementing bureaucrats is key to whether they are effective (Banerjee et al., 2021).

Two new studies, part of a broader cluster of RCTs on community policing,[3] study the efficacy of community policing interventions in different contexts. Community policing has been proposed as a way to build trust and increase collaboration between police and citizens by instituting foot patrols, town hall meetings, door to door visits, problem-oriented policing, and platforms for citizen feedback. The idea is that citizen cooperation is key to the efficacy of the police force, and therefore improving relations between citizens and police will reduce crime and violence.

[3] These studies were part of EGAP's Metaketa on community policing, which included six randomized controlled trials on community policing in Brazil, Colombia, Liberia, Pakistan, the Philippines, and Uganda. The six RCTs used common interventions and measurement techniques to facilitate comparison across studies. A meta-analysis of the six studies is described in Blair et al., (revise and resubmit, Science).

Blair, Grossman, and Wilke (Working paper 2021) tested this idea with a large-scale field experiment that randomized whether police stations in rural Uganda received a community policing intervention. Before the intervention, trust in police was low and police often engaged in acts of malfeasance and petty corruption. The community policing intervention was designed by the Ugandan Police Force, in collaboration with civil society organizations, and included officer training, town hall meetings, door to door visits, and the formation of community-watch teams. Using both administrative crime data and survey data, the results show that community policing had little to no effect on key outcomes such as perceptions of police, prevalence of crime, and citizens' sense of security. Likewise, the program did not improve police officers' empathy or sense of accountability to citizens. The null results likely arise from low compliance with the program, which in turn likely stemmed from a combination of resource constraints, frequent rotation of police officers, and a lack of reliable mechanisms for monitoring and sanctioning officers. "While 92% of treatment communities held at least one town hall meeting, only 69% held two, and only 34% held four or more," and reported door to door visits were even lower. Officers were expected to exert more effort to implement the program but did not receive any additional resources to do so. The program did, however, increase the frequency of interactions between civilians and police and citizen knowledge of the law, including suggestive evidence of an increase in crime reporting. The increased interactions may have also led to more opportunities for bribe seeking, as results suggest the program increased reported bribe payments. However, it is unclear whether the program increased citizens' recognition of informal fee requests as bribes or the actual incidence of payments.

In Medellín, Colombia, Arias, Hanson, Kronick, and Slough (Working paper 2021) test the impact of town hall meetings on police-community relations. In comparison to the Ugandan context, Colombian police have substantial resources and have implemented community policing strategies since the 1990s, including foot patrols, problem-oriented policing, and mechanisms for citizen feedback. Despite significant reforms and improvements since the 1990s, public perceptions of the police have not changed and trust in police is lower than their relative performance would predict. A large randomized experiment assessed whether town hall meetings, cross-randomized with information on police responsibilities and activities, could close this gap. Based on administrative crime data and survey data, the results suggest that town hall meetings generally improved citizens' perceptions of the police. However, they had no significant effects on crime, crime reporting, police abuse, or officer beliefs. The null results could be due to low statistical power, spillovers to control neighborhoods due to frequent officer transfers, or suggest that town hall meetings had limited additional impact given the existing community policing program.

These results fit with a broader pattern among RCTs studying community policing that generally find no significant effects of community policing interventions on citizen-police relations or crime (Blair et al., revise and resubmit, Science). Taken together, these papers highlight the importance of considering multiple binding constraints when designing interventions of incremental reforms. In contexts where multiple constraints bind

simultaneously, interventions that alleviate one constraint are unlikely to be effective if they do not also alleviate the other binding constraints.

Mirroring the idea that citizen cooperation is key to the efficacy of the police force, citizen loyalty also appears to be key in sustaining informal governance structures, such as gang rule. The conventional wisdom that crime flourishes where the state is weak implies that crime should reduce when the state increases its presence and provision of services. Blattman, Duncan, Lessing, and Tobon (Working paper 2021)[4] use novel qualitative and quantitative data on gangs in Medellín, Colombia, combined with experimental and quasi-experimental variation, to test this conventional wisdom. Based on large-scale, systematic interviews with gangs, they find that the conventional view overlooks the indirect incentives gangs have to rule, including fostering civilian loyalty, which helps keep police out, and protecting other business lines, such as retail drug sales. While a standard model of duopolistic competition predicts that an increase in state services reduces gang services, a model that incorporates these indirect incentives predicts that gangs can strategically respond to an increase in state services by increasing their own provision of services.

The researchers then test these predictions empirically using both quasi-experimental and experimental variation. Specifically, they randomly selected 40 out of 80 neighborhoods to receive intensified city governance for 2 years, including a full-time liaison to facilitate the provision of state services. To examine long-run effects, researchers also examined quasi-exogenous variation created by the introduction of new neighborhood borders in 1987, which generated discontinuities in access to government services. Results from both the experiment and quasi-experiment show that increased provision of state services did not reduce criminal protection and may have even increased gang rule in the longer run. These results are consistent with gangs strategically responding to increased state services, as well as the difficulties in changing perceptions of state capacity more generally. Taken together, the results suggest that states need to target the main motives for gang rule if they want to effectively reduce organized crime. This could include reducing gang revenues, such as lowering the profitability of the illicit drug trade, and changing social norms around civilian loyalty to gangs. The results also illustrate the advantages of using mixed research methods, such as combining systematic primary qualitative data with quasi-experimental and experimental methods, particularly when studying informal or illegal activities.

3. D. ii. The role of audits

Audits also play an important role in enforcement. The evidence reviewed in the White Paper shows that audits regulated behavior by firms (Duflo et al., 2018) and who pays inspectors (regulator versus regulated firm) matters in terms of what they report (Duflo et al., 2013).

[4] This research was supported by EDI Research Area 4, but is included in this discussion as a randomized evaluation supported by EDI on related topics.

However, this raises a new question of how audits compare to other enforcement mechanisms in terms of efficacy.

Berkouwer, Hsu, Miguel, and Wolfram (Working paper 2021) shed light on this question by comparing the effects of more stringent donor conditionality to third-party audits on the quality of construction in public infrastructure projects. In Kenya's last mile connectivity project, which seeks to achieve universal household electricity access, construction sites received funding from either the World Bank (WB) or the African Development Bank (ADB). Importantly, the WB and ADB use different contracting methods which result in relatively more stringent conditions for WB-funded projects. Complementing this naturally occurring variation, the researchers announced independent engineering audits for a randomly selected subset of construction sites. The results, while preliminary, suggest that both the relatively stringent conditions of the World Bank and the independent audits improved construction outcomes by similar magnitudes. However, the "World Bank's restrictions caused significant construction delays of several months on average." While sites funded by the World Bank experienced significant improvements in the quality of network configuration, there were no improvements in overall construction quality, installation quality, reliability, or safety. This project is ongoing and future analysis will try to disentangle the mechanisms driving these effects.

In a related theme, Barnwal and Ryan (ongoing) are conducting a randomized experiment to study the impacts of introducing village-level electrification camps and ex-post audits of the electrification work completed by private contractors to improve the performance of a universal electrification program in Jharkhand, India. Private contracts may not share the government's goals of fairness and service quality. They may even increase the hassle for households in the name of red tape and ask for bribes. Indeed, pilot data suggests private contractors leave many consumers unconnected and often only partially install connections, even in villages that are declared to be "fully electrified." The researchers plan to study the distributional impacts of these interventions. This project is ongoing, and results are forthcoming.

Open questions

- What complementary reforms are needed for community policing to be effective?
- How can civilian perceptions and attitudes be shifted to combat gang rule?
- How does donor conditionality affect other aspects of performance, such as environmental protection and community trust? How does this compare to other enforcement tactics?
- How much do improvements in enforcement generate displacement in wrongdoing?
- Is there an interactive effect of incentives to monitors and the monitored?
- What are the distributional implications of weak or heterogeneous enforcement?

4. Cross-cutting themes

In reviewing the new evidence generated from EDI's RA3 portfolio of funded studies, a few key cross-cutting themes emerge.

The studies in the portfolio studied information constraints across many different contexts and interventions. One key lesson is that it is not only what information is available, but how that information is aggregated and analyzed, that matters for improving service delivery. The cognitive, time, and capacity costs associated with generating, accessing, and analyzing information can be prohibitive. Technological innovations that aggregate information (such as mobile apps or new data platforms), or analyze existing data (such as Machine Learning algorithms), can lead to important gains in tax collection (Mittal, Reich, and Mahajan, SIGCAS 2018), service delivery (Dodge et al., Working paper 2021), and legal outcomes (Sadka, Seira, and Woodruff, Working paper 2020). However, as the results from Bachas et al. (Working paper 2021) highlight, tradeoffs exist between technological innovations and bureaucrat autonomy and transparency. Technology is unlikely to lead to improvements without the buy-in of the personnel it is intended to help, not to speak of the personnel that the technology is meant to monitor and sanction. Better understanding of the supply and demand factors that affect the type of information available is also an important avenue for future research.

A second theme that emerges across the information interventions studied is that who receives information matters when agents face different incentives or when relationships are characterized by power imbalances. Sadka, Seira, and Woodruff (Working paper 2020) highlight the important agency issues that are present in the plaintiff-lawyer relationship, as information only improved legal outcomes when the plaintiff received information directly. Relatedly, Casey, Kamara, and Meriggi (AER 2021) illustrate that politicians voluntarily take information on voter preferences into account when their preferences are aligned and election risks are minimal. Dodge et al. (Working paper 2021) show that in contexts where bureaucrats share the same incentives, the provision of improved information at different levels of the bureaucracy act as substitutes. Taken together, the results suggest that who receives information, and the incentives they face, matter for how that information is used.

Many of the studies reviewed also highlight the importance of considering whether multiple binding constraints exist for a given policy goal. Deserranno (r) al. (Working paper 2021) show that relaxing both demand-side and supply-side constraints to households accessing health services had significantly larger effects than relaxing either constraint in isolation. Two studies on community policing (Blair, Grossman, and Wilke, Working paper 2021; Arias et al., Working paper 2021) show that incremental reforms are unlikely to reduce crime or improve citizen's sense of security without complementary changes, perhaps those that reduce resource constraints or the frequent transfers of officers. Taken together these studies highlight the importance of identifying the complementary interventions necessary to address all binding constraints. When limits are set by informal institutions and culture, it may be particularly difficult to stage effective complementary interventions. Therefore, understanding whether other constraints are created by formal or informal institutions is relevant. If simultaneous interventions are necessary, modeling and identifying interaction effects is also paramount.

Finally, many of the reviewed studies highlight how the welfare impacts of a policy as it is written *de jure* can differ substantially from how it is implemented *de facto*. Berkouwer et al.

(Working paper, 2021) show that weak or heterogenous enforcement can distort the intended welfare benefits of a policy. Cogneau et al. (Working paper 2021) illustrate that low compliance with the property tax in Senegal leads to a tax profile that is significantly more regressive than what is intended under perfect compliance. Similarly, the evidence from Field and Vyborny (Working paper 2020) suggest that expanding legal protection of women's rights does not automatically lead to improvements in women's outcomes without complementary interventions to improve compliance among implementers.

5. Conclusions

The evidence reviewed in this paper shows that we have made significant progress in our understanding of the role of institutions, broadly defined, in development and growth. The studies also highlight the complexity and nuance that define these relationships and the need for further research. It is clear that this is an active area that will continue to push the frontier of our understanding of economic development over time.

A couple of patterns of how evidence gaps have evolved over time emerge from this review. Some areas have progressed significantly with the use of large field experiments, particularly in topics related to improving state capacity and the quality of enforcement. In other areas, there continues to be a notable lack of experimental evidence, such as the role of political norms and culture and the design of state procedures and the legal system. These persistent evidence gaps make clear the importance of renewing and expanding efforts to encourage governments to include data-driven decision making and embedded experimentation as an integral part of the policymaking process.

However, these questions by their nature can be challenging to study with field experiments, highlighting the importance of complementary research methods. While the majority of research funded by EDI RA3 and reviewed in this paper were RCTs, many studies employed hybrid research methods, combining experimental variation with naturally occurring variation, integrating experimental approaches with systematic qualitative methods, and producing valuable insights from data generated in the process of developing a RCT. Notably, the majority of studies were conducted in partnership with governments, studying interventions often implemented at scale. While naturally occurring experiments allow researchers to answer causal questions, the needed variation does not always arise organically. When there is not enough naturally occurring variation, researchers turn to RCTs to create the necessary exogenous variation to study causal relationships. However, the artificial nature of RCTs can limit their external validity. What is clear from the studies reviewed here is that researchers do not need to choose between the two.

Indeed, the RA3 EDI initiative was designed to promote what we would call "naturalistic field experiments." These interventions arise from partnerships between researchers and policymakers, are often implemented at scale or at least in the very context in which a full scale program will ultimately operate, and combine the internal validity of RCTs with the external validity of observational studies. In this sense, EDI's RA3 has funded the next generation of

empirical studies on institutions and development. It is clear from this review that this innovative approach has produced valuable insights on hitherto difficult to study topics. At the same time, this is clearly only the beginning. To continue to make progress, we need to increase and expand collaboration between researchers and government and strengthen the commitment from governments, policymakers, and implementers to evidence-based policy.

References

**** Indicates research supported by EDI RA3*

Aker, Jenny, Paul Collier, and Pedro Vicente. 2017. "Is information power? Using mobile phones and free newspapers during an election in Mozambique," *The Review of Economics and Statistics,* 99 (2): 185-200.

Ali, Daniel, Klaus Deininger, and Markus Goldstein. 2014. "Environmental and gender impacts of land tenure regularization in Africa: Pilot evidence from Rwanda," *Journal of Development Economics,* 110(c): 262-275.

*** Arias, Eric, Rebecca Hanson, Dorothy Kronick, and Tara Slough. 2021. "Does Community Policing Work? Evidence from Medellin," chapter for edited volume, forthcoming.

*** Ash, Elliott, Sam Asher, Aditi Bhowmick, Sandeep Bhupatiraju, Daniel Chen, Tanaya Devi, Christoph Goessmann, Paul Novosad, and Bilal Siddiqi. 2021. "In-group bias in the Indian judiciary: Evidence from 5 million criminal cases," Working paper, August 2021.

Ashraf, Nava, Oriana Bandiera, Edward Davenport, and Scott Lee. 2020. "Losing prosociality in the quest for talent? Sorting, selection, and productivity in the delivery of public services," *American Economic Review,* 110(5): 1355-94.

Azulai, Michel, Imran Rasul, Daniel Rogger, and Martin Williams. 2020. "Can training improve organizational culture? Experimental evidence from Ghana's civil service," Working paper, October 2020.

*** Bachas, Pierre, Anne Brockmeyer, Alipio Ferreira, and Bassirou Sarr. 2021. "How to target enforcement at scale? Evidence from tax audits in Senegal," Working Paper, July 2021.

Banerjee, Abhijit, Raghabendra Chattopadhyay, Esther Duflo, Daniel Keniston, and Nina Singh. 2021. "Improving police performance in Rajasthan, India: Experimental evidence on incentives, managerial autonomy, and training," *American Economic Journal: Economic Policy,* 13(1):36-66.

Banerjee, Abhijit, Esther Duflo, and Rachel Glennerster. 2008. "Putting a band-aid on a corpse: Incentives for nurses in the Indian public health care system," *Journal of the European Economic Association,* 6(2-3): 487-500.

Barnwal, Prabhat, and Nicholas Ryan. Ongoing. "Monitoring the performance of agents in public service delivery," Project ongoing, results forthcoming.

Basinga, Paulin, Paul Gertler, Agnes Binagwaho, Agnes Soucat, Jennifer Sturdy, and Christel Vermeersch. 2011. "Effect on maternal and child health services in Rwanda of payment to

primary health-care providers for performance: An impact evaluation," *The Lancet*, 377(9775): 1421-28.

Beath, Andrew, Fotini Christia, Georgy Egorov, and Ruben Enikolopov. 2016. "Electoral rules and political selection: Theory and evidence from a field experiment in Afghanistan," *The Review of Economic Studies,* 83(3): 932-968.

Beaman, Lori, Raghabendra Chattopadhyay, Esther Duflo, Rohini Pande, and Petia Topalova. 2009. "Powerful women: Does exposure reduce bias?" *The Quarterly Journal of Economics*, 124 (4): 1497-1540.

Berdejó, Carlos, and Noam Yuchtman. 2013. "Crime, punishment, and politics: An analysis of political cycles in criminal sentencing," *Review of Economics and Statistics,* 95(3):741-56.

*** Berkouwer, Susanna, Eric Hsu, Edward Miguel, and Catherine Wolfram. 2021. "Promoting accountability in public projects: Donors, audits, and rural electrification," Working paper, March 2021.

Besley, Timothy, and Anne Case. 1995. "Incumbent behavior: Vote-seeking, tax-setting, and yardstick competition," *American Economic Review,* 85(1):25-45.

Bidwell, Kelly, Katherine Casey, and Rachel Glennerster. 2020. "Debates: Voting and expenditure responses to political communication," *Journal of Political Economy,* 128(8):2880-2924.

*** Blair, Robert, Guy Grossman, and Anna Wilke. 2021. "Can community policing improve police-community relations in an electoral authoritarian regime? Experimental evidence from Uganda," Working paper, January 2021.

*** Blair, Graeme, Jeremy Weinstein, Fotini Christia, Eric Arias, Emile Badran, Robert Blair, Ali Cheema, Ahsan Farooqui, Thiemo Fetzer, Guy Grossman, Dotan Haim, Zulfiqar Hameed, Rebecca Hanson, Ali Hasanain, Dorothy Kronick, Benjamin Morse, Robert Muggah, Fatiq Nadeem, Lily Tsai, Matthew Nanes, Tara Slough, Nico Ravanilla, Jacob Shapiro, Barbara Silva, Pedro Souza, and Anna Wilke. 2021. "Does community policing build trust in police and reduce crime? Evidence from six coordinated field experiments in the global south," *Science*, revise and resubmit.

Blattman, Christopher, Gustavo Duncan, Benjamin Lessing, and Santiago Tobón. 2021. "Gang rule: Understanding and countering criminal governance," NBER Working Paper 28458, February 2021.

Blattman, Christopher, Alexandra Hartman, and Robert Blair. 2014. "How to promote order and property rights under weak rule of law? An experiment in changing dispute resolution behavior through community education," *American Political Science Review,* 108(1): 100-120.

Blumenthal, Marsha, Charles Christian, and Joel Slemrod. 2001. "Do normative appeals affect tax compliance? Evidence from a controlled experiment in Minnesota," *National Tax Journal*, 54(1): 125-38.

Cantoni, Davide, Yuyu Chen, David Yang, Noam Yuchtman, and Jane Zhang. 2017. "Curriculum and ideology," *Journal of Political Economy*, 125 (2): 338 - 392.

*** Casey, Katherine, Abou Bakarr Kamara, and Niccolo Meriggi. 2021. "An experiment in candidate selection," *American Economic Review,* 111(5):1575-1612.

Castro, Lucio and Carlos Scartascini. 2015. "Tax compliance and enforcement in the Pampas: Evidence from a field experiment," *Journal of Economic Behavior & Organization,* 116:65-82.

Chattopadhyay, Raghabendra, and Esther Duflo. 2004. "Women as policy makers: Evidence from a randomized policy experiment in India," *Econometrica,* 72(5):1409-43.

*** Cogneau, Denis, Marc Gurgand, Justine Knebelmann, Victor Pouliquen, and Bassirou Sarr. 2021. "Measure the potential for property taxation in Dakar, Senegal," Working paper, February 2021.

*** Cogneau, Denis, Marc Gurgand, Justine Knebelmann, Victor Pouliquen, and Bassirou Sarr. Ongoing. "Measuring the effect of technological improvements and compliance nudges on property taxes," Project ongoing, results forthcoming.

Cohen, Isabelle, Ernesto Dal Bó, Frederico Finan, Kizito Omala, and David Schönholzer. Ongoing. "Teacher rotation and student outcomes: Experimental evidence from Uganda," Project ongoing, results forthcoming.

Collier, Paul and Pedro Vicente. 2014. "Votes and violence: Evidence from a field experiment in Nigeria," *The Economic Journal*, 124 (574): F327-55.

Cruz, Cesi, Philip Keefer, and Julien Labonne. 2021. "Buying informed voters: New effects of information on voters and candidates," *The Economic Journal*, 131(635): 1005-1134.

Dal Bó, Ernesto, Frederico Finan, and Martin Rossi. 2013. "Strengthening state capabilities: The role of financial incentives in the call to public service," *Quarterly Journal of Economics*, 128(3): 1169-1218.

de Soto, Hernando. 2000. *The mystery of capital: Why capitalism triumphs in the west and fails everywhere else*. New York: Basic Books.

DellaVigna, Stefano, and Ethan Kaplan. 2007. "The Fox News effect: Media bias and voting," *Quarterly Journal of Economics,* 122(3): 1187-1234.

Deserranno, Erika. 2019. "Financial incentives as signals: Experimental evidence from the recruitment of village promoters in Uganda," *American Economic Journal: Applied Economics*, 11(1):277-317.

Deserranno, Erika, Philipp Kastrau, and Gianmarco León-Ciliotta. 2020. "Financial incentives in multi-layered organizations: Empirical evidence from the community health worker program in Sierra Leone," Working Paper, 2020.

*** Deserranno, Erika (r) Stefano Caria (r) Philipp Kastrau (r) Gianmarco León-Ciliotta. 2021. "Financial incentives in multi-layered organizations: An experiment in the public sector," Working paper, September 2021.

*** Deserranno, Erika, Philipp Kastrau, and Gianmarco León-Ciliotta. 2021. "Promotions and productivity: The role of meritocracy and pay progression in the public sector," UPF Economic Working Paper No. 1770, February 2021.

*** Dodge, Eric, Yusuf Neggers, Rohini Pande, and Charity Moore. 2021. "Updating the state: Information acquisition costs and public benefit delivery," Working paper, July 2021.

Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas 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, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan. 2018. "The value of regulatory discretion: Estimates from environmental inspections in India," *Econometrica*, 86:2123-2160.

Fellner, Gerlinde, Rupert Sausgruber, and Christian Traxler. 2013. "Testing enforcement strategies in the field: Threat, moral appeal, and social information," *Journal of the European Economic Association*, 11(3):634-60.

Ferraz, Claudio and Frederico Finan. 2008. "Exposing corrupt politicians: The effects of Brazil's publicly released audits on electoral outcomes," *Quarterly Journal of Economics,* 123(2):703-45.

Field, Erica. 2005. "Property rights and investment in urban slums," *Journal of the European Economic Association,* 3(2-3):279-90.

*** Field, Erica, and Kate Vyborny. 2020. "Information gaps and de jure legal rights: Evidence from Pakistan," Working paper, September 2020.

Frankel, Martin, Carroll Seron, Gregg Ryzin, and Jean Frankel. 2001. "Impact of legal counsel in outcomes for poor tenants in New York City's housing court: Results of a randomized experiment," *Law & Society Review*, 35(2):419-34.

Fujiwara, Thomas. 2015. "Voting technology, political responsiveness, and infant health: Evidence from Brazil," *Econometrica,* 83(2):423-64.

*** Ghatak, Maitreesh, Markus Goldstein, James Habyarimana, Joao Montalvao, Michael O'Sullivan, and Christopher Udry. Ongoing. "Relaxing credit constraints and tenure insecurity in imperfect markets," Project ongoing, results forthcoming.

Greiner, James, and Cassandra Pattanayak. 2012. "Randomized evaluation in legal assistance: What difference does representation (offer and actual use) make?" *Yale Law Journal*, 121(8):2118-2214.

Grossman, Guy, Oren Gazal-Ayal, Samuel Pimentel, and Jeremy Weinstein. 2016. "Descriptive representation and judicial outcomes in multi-ethnic societies," *American Journal of Political Science*, 60(1):44-69.

Haggard, Stephan, Andrew MacIntrye, and Lydia Tiede. 2008. "The rule of law and economic development," *Annual Review of Political Science*, 11(1):205-34.

Kendall, Chad, Tommaso Nannicini, and Francesco Trebbi. 2015. "How do voters respond to information? Evidence from a randomized campaign," *American Economic Review*, 105(1):322-53.

Khan, Adnan, Asim Khwaja, and Benjamin Olken. 2016. "Tax farming redux: Experimental evidence on performance pay for tax collectors," *Quarterly Journal of Economics,* 131(1):219-271.

Kleven, Henrik, Martin Knudsen, Claus Kreiner, Søren Pedersen, and Emmanuel Saez. 2011. "Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark," *Econometrica,* 79(3):651-92.

*** Leaver, Clare, Owen Ozier, Pieter Serneels, and Andrew Zeitlin. 2021. "Recruitment, effort, and retention effects of performance contracts for civil servants: Experimental evidence from Rwandan primary Schools," *American Economic Review,* 111(7):2213-2246.

Lim, Claire. 2015. "Media influence on courts: Evidence from civil case adjudication," *American Law and Economics Review,* 17(1):87-126.

Lim, Claire, and James Snyder. 2015. "Is more information always better? Party cues and candidate quality in U.S. judicial elections," *Journal of Public Economics*, 128:107-23.

Lim, Claire, James Snyder, and David Strömberg. 2015. "The judge, the politician, and the press: Newspaper coverage and criminal sentencing across electoral systems," *American Economic Journal: Applied Economics*, *7(4):103-35.*

*** Luksic, Jan, and Shekhar Mittal. 2019. "Red tape? The revenue impact of the VAT filing thresholds," Working paper, May 2019.

McGraw, Kathleen, and John Scholz. 1991. "Appeals to civic virtue versus attention to self-interest: Effects on tax compliance," *Law & Society Review,* 25(3):471-98.

*** Mittal, Shekhar, Ofir Reich, and Aprajit Mahajan. 2018. "Who is bogus? Using one-sided labels to identify fraudulent firms from tax returns," *ACM SIGCAS Conference on Computing and Sustainable Societies (COMPASS)*, June 20-22, 2018, Menlo Park and San Jose, CA.

Mueller, Valerie, Lucy Billings, Tewodaj Mogues, Amber Peterman, and Ayala Wineman. 2018. "Filling the legal void? Impacts of a community-based legal aid program on women," *Oxford Development Studies*, 46(4):453-469.

North, Douglass. 1991. "Institutions," *The Journal of Economic Perspectives*, 5 (1): 97-112.
Olken, Benjamin. 2010. "Direct democracy and local public goods: Evidence from a field experiment in Indonesia," *American Political Science Review,* 104(2):243-267.

Pande, Rohini. 2003. "Can mandated political representation increase policy influence for disadvantaged minorities? Theory and evidence from India," *American Economic Review*, 93(4):1132-51.

Perez-Truglia, Ricardo, and Ugo Troiano. 2018. "Shaming tax delinquents," *Journal of Public Economics,* 167:120-137.

Perry, James, and Annie Hondeghem. 2008. "Building theory and empirical evidence about public service motivation," *International Public Management Journal*, 11(1):3-12.

Petrovsky, Nicolai. 2009. "Does public service motivation predict higher public service performance? A research synthesis," Working Paper, September 2009.

Pomeranz, Dina. 2015. "No taxation without information: Deterrence and self-enforcement in the Value Added Tax," *American Economic Review*, 105(8):2539-69.

*** Prillaman, Soledad. Ongoing. "Is knowledge power? Civics training, women's political representation, and local governance in India," Project ongoing, results forthcoming.

Rasul, Imran, and Daniel Rogger. 2018. "Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service," *The Economic Journal*, 128:413-446.

*** Sadka, Joyce, Enrique Seira, and Christopher Woodruff. 2020. "Information and bargaining through agents: Experimental evidence from Mexico's labor courts," NBER Working Paper 25137, October 2020.

Sandefur, Justin, and Bilal Siddiqi. 2015. "Delivering Justice to the Poor: Theory and Experimental Evidence from Liberia," Working paper, March 2015.

Vicente, Pedro. 2014. "Is vote buying effective? Evidence from a field experiment in West Africa," *The Economic Journal*, 124 (574): F356-F387.