

# Report on Progress and Plans, January 2015

Rema Hanna

---

## INTRODUCTION

Despite a multitude of opinions and debates on exactly how to achieve this, many, including my colleagues at the Harvard Kennedy School, share an idealism that government can be improved to not only promote economic growth, but to also help ensure that the poor and disadvantaged can participate in increased prosperity. This idea has always motivated my research and teaching, as I use the tools from economics to provide rigorous, but practical insights on how to improve public sector institutions in developing countries.

Specifically, I have exploited experimental and quasi-experimental methods—coupled with a careful consideration of the underlying theoretical mechanisms and qualitative field insights—to contribute to our understanding of **how to improve the access to and quality of public services for the poor in developing countries**. Within this theme, I have focused on two streams. In the first, I explore questions relating to the **barriers that citizens face in accessing services**:

- How do we **target the poor** in low-income countries, where informal markets and an underdeveloped tax system result in limited data with which to verify income status? I show that including communities in targeting can help harness local information for the selection of beneficiaries in targeted transfer programs, especially in areas with stronger social networks. Self-selection techniques can also improve targeting over more common automatic enrollment systems. In follow-up work, I illustrate that greater transparency, as well as an increase in competition over who implements a transfer program, can reduce leakage in targeted programs.
- What's the **cost of corruption**? Studying driver's licenses, I establish that corruption does not just "grease the wheels," but that it also has real social costs in terms of misallocating licenses to the wrong people. In other work, I show that fraudulent absenteeism is a particularly costly form of corruption: reducing teacher absenteeism has a large impact on learning and health-care worker absenteeism reduces baby birth weight.
- Are low quality service outcomes driven by **selection or the incentives of bureaucrats**? Through two large-scale field experiments, I illustrate that improved monitoring and financial incentives can potentially improve outcomes (e.g., learning and health). In doing so, I also document the challenges to implementing these types of programs in real government settings. Nonetheless, selection also plays a role: I find that college students who cheat in laboratory tasks prefer public sector jobs. In short, we need to think both about how current incentive structures and screening mechanisms drive who enters the bureaucracy and how they behave once there.
- What are the **costs of discrimination and how can government programs mitigate them**? Using experimental methods, I illustrate that teachers engage in statistical discrimination against minorities. However, I also show that government-mandated affirmative action policies can mitigate historical injustices: the policies target low-income households and yield returns in the form of higher wages.

In the second, I offer insight into questions regarding the **delivery of environmental services** (e.g., clean air and water) in low income countries:

- Can countries with low institutional capacity effectively provide **public goods**, such as **environmental quality**? In a series of work, I quantify the costs of bad air quality, both on infant mortality and illness-related worker absenteeism. Using evidence from India, I then show that developing countries can successfully implement environmental regulation if there is citizen demand.
- Does **individual behavior affect the realized benefits from the distribution** of environmental products or agricultural extension? I show that households do not experience better health as a result of an improved stove (even though the same stove reduces emissions in the lab), simply because they do not take up, regularly use, or maintain the stoves. Another study uses agricultural extension as a context for demonstrating one explanation for low technology adoption: individuals cannot learn the benefits of a technique or product if they do not notice its importance or attend to the aspects of use that are important in order to reap benefits.

My work has been recognized broadly, both in the top academic journals (e.g., *American Economic Review*, *Quarterly Journal of Economics*, and *Journal of Political Economy*) and in major media outlets such as the *New York Times*, the *Economist*, the *Washington Post*, the *Times of India*, and the *Jakarta Post*. I regularly interact and work with policy-practitioners through conferences, executive education, and informal meetings. Importantly, in many of my projects, I work directly with government partners to ensure that my projects not only provide general insights into theoretical questions, but are also directly relevant to the kinds of questions that policy-makers are grappling with. As such, two of my projects (community targeting and transparency) informed the scale-up of reforms to Indonesia’s targeted transfer programs, helping to improve access for 15.5 million poor households across the country.

Moving forward, I hope to continue to contribute high quality, policy-relevant research to help further our understanding of how public services can be enhanced to help the poor in developing countries—and to use this knowledge to help train the next generation of public servants and public policy researchers.

## I. RESEARCH CONTRIBUTIONS

My research agenda is centered on understanding how to improve the design and delivery of government programs, with a special emphasis on *targeted social protection programs* and *environmental regulation*. Within this space, I aim to produce careful, empirical evidence on key questions. For ease of exposition, I first discuss my work on general service provision in Section A, and then focus on questions relating specifically to environmental issues in Section B.

### A. PROVIDING PUBLIC SERVICES TO THE POOR

As I discuss in “**Corruption**” (joint with Abhijit Banerjee and Sendhil Mullainathan, **Handbook of Organizational Economics**), the delivery of public services often comes down to a dance between the central government and local officials: the national government designs a program around its goals. Local officials are then tasked with implementing the program rules. In doing so, they often have quite a bit of discretion: some of this is intentional, since it is hard to contract on every aspect of the program. Moreover, you may want discretion, as it can bring in local pertinent

information over how programs should be allocated and run. However, some of this is de-facto, since it is hard to monitor whether the local officials implement the tasks that they are given.

In this strain of research, I unpack different questions regarding the nature of this relationship between the central government, local officials, and citizens:

- **How does local discretion affect the ultimate allocation of services?** I explore the tension between having a formal rule and allowing for local discretion within the context of targeting programs: can you elicit local information to improve targeting, or will you simply run an increased risk of capture by local officials and elites? Similarly, in allowing for discretion, do you run the risk that a bureaucrat's private tastes dominate his decisions? For example, would you observe discrimination against minorities?
- **Does the nature of the rules affect how bureaucrats respond to them?** Do they break "bad" rules, either because they are benevolent or because of career concern (e.g., trying to minimize citizens' complaints)? Or do they break "good" rules, generating a social cost?
- **Does increased monitoring, coupled with incentives, ensure that bureaucrats enforce the rules that are on the books?** I explore "top-down" monitoring systems, paying attention to whether monitoring along a particular dimensions leads bureaucrats to shirk along other non-monitored dimensions that we may also care about. Alternatively, I explore "bottom-up" methods: I ask whether it is efficient for citizens to monitor, either by providing them with more transparency about program rules or by allowing them choice over who administers the program (e.g., local officials or private contractors).
- **Given the incentives in place, who chooses to become a bureaucrat?** Are the relative returns to corruption higher in the public or private sector? How does screening for ability affect who ends up in bureaucracy?

#### a. HOW TO TARGET RESOURCES TO THE POOR?

A key challenge in directing transfers towards the poor is figuring out who is *actually* poor. The most obvious method is an income census. However, this is often prohibitively expensive and logistically challenging. The poor typically live in remote areas and may be excluded from local institutions, so just locating them is tough. The fact that they are characterized both by high income volatility and informal labor markets further thwarts the process. Finally, weak governance may further dilute these efforts, since the process may be captured by local elites or subject to high levels of corruption by the local, administering bureaucrats. Thus, governments often try to find alternative methods that do not require collecting income data.

One method is to just create a simple rule, finding a few characteristics that are easy to monitor and targeting based on those. The most classic example is targeting a group identity (low-caste in India, women or African-Americans in the US, etc.) through affirmative action. However, these programs are universally controversial. Little evidence exists as to whether targeting by a group characteristic results in transfers to the poor. Moreover, even with progressive targeting, do these programs actually help the targeted group?

With Sendhil Mullainathan and Marianne Bertrand, I examined these issues ("**Affirmative Action in India: Evidence from Engineering College Admissions in India**," *Journal of Public*

**Economics**). Collecting primary data from college graduates in India, we find that caste-based targeting transfers resources to the poor: low-caste individuals are *poorer* than those in the upper caste whom they replace at university. We then show that the low-caste groups targeted by affirmative action ultimately achieve higher wages, but that these gains are smaller than the losses incurred by upper-caste individuals who lose a slot. Note that this paper was particularly timely: in 2007, the Supreme Court placed a stay of order on the Indian Parliament's bill to expand affirmative action, citing little evidence of its impacts. In response to the debate, we wrote an op-ed for *Mint* (the *Wall Street Journal's* India venture).

In more recent work, I directly tested the impact of allowing local discretion into the targeting process: In "**Targeting the Poor: Evidence from a Field Experiment in Indonesia**" (*American Economic Review*; joint with Vivi Alatas, Abhijit Banerjee, Ben Olken, and Julia Tobias), I experimentally compare proxy-means testing (PMT)—a formulaic approach—and community-based targeting. Theoretically, villagers might have better information about who is poor than the government, so community targeting may be more "accurate" than the PMT. However, the risk of community involvement is that targeting decisions may be based on factors beyond poverty as defined by the government, either through genuine disagreements about what "poverty" means, or more troubling, through elite capture.

Proxy-means testing performed better at targeting households below Indonesia's poverty line, but the difference was small. For a typically-sized transfer program in Indonesia, our simulations suggest that the two methods would not yield significantly different effects on reducing the poverty rate. However, the community method was better at identifying the very poorest (below \$1 per day) and generated significantly higher satisfaction levels with the beneficiary list, which can make it easier to implement controversial targeted programs.

We then explored when these processes lead to the selection of different types of individuals. First, using experimental variation in who was invited for the meeting, we show that elite capture is absent in this setting even when the local officials have full discretion over targeting. Second, using random variation in the order that households were considered at the meeting, community targeting performs better than the PMT at the start of the meeting, but it worsens as the meeting proceeds. This suggests that if one can elicit community information in a less time-consuming way, the community method may outperform the PMT.

Finally, examining the role of preferences, we show that the community method moved the targeting outcomes away from a ranking based on consumption and towards a ranking that one would obtain by polling citizens. In short, the community has different preferences than the government: they prioritize factors that predict *earnings capacity* rather than just consumption.

Community-based targeting will only be successful if communities have good information about their members. In a follow-up paper (joint with Vivi Alatas, Abhijit Banerjee, Arun Chandrasekhar, and Ben Olken; entitled "**Identifying the Poor: Social Networks and the Aggregation of Information**"), we further exploit the detailed data that we collected on social networks, individual beliefs about the income status of their neighbors and actual income status. This dataset also presented a unique opportunity to test network theories, given the large number of networks (640 villages). This paper is R&R at the *American Economic Review*.

We first estimate a model of information diffusion using within-village network variation and use this to simulate the cross-village relationship between information diffusion and network

characteristics. Our model is useful since it allows us to make predictions as to how different types of network characteristics—each conditional on one another—should affect diffusion. We then show that, both in the model and empirically, more “connected” villages spread information about each other’s income better. Importantly, we use this structure to show that networks with better diffusive properties differentially benefit from community targeting, implying that selection may be improved by varying targeting methods based on the underlying context.

Two key questions arose from the previous study: Can you elicit an individual’s information about their income status through self-targeting methods? Would the community targeting results (i.e., the findings on elite capture) hold up in an even higher stakes environment?

To study these questions, as well as inform Indonesia’s new unified targeting system, we worked closely with the government to design an experiment within the context of their conditional cash transfer program, PKH. Given that PKH is a national program, this project has been a true collaborative effort with the Census Bureau, the Finance Ministry, the Department of Family Welfare, an NGO, the World Bank, and the targeting taskforce within the Poverty Monitoring Unit in the Vice President’s office (TNP2K). In addition to its policy ramifications, the experiment generated two academic papers (joint with Vivi Alatas, Abhijit Banerjee, Ben Olken, Matt Wai-poi, and Ririn Purnamasari).

The first is entitled “**Self-Targeting: Evidence from a Field Experiment in Indonesia**” (*Journal of Political Economy*). We lay out a simple model to explore an individual’s decision to apply to a transfer program given one’s income, the cost of applying, and the probability of being selected during the means test. In doing so, we lay out the conditions in which increasing application costs in a self-targeting system can actually worsen selection and the conditions in which it may improve it. We then test the theory by randomizing whether the PKH expansion used self-selection versus an automatic enrollment system to target beneficiaries.

We first explore who applies under the self-selection system. We show that people select in based on two margins. First, their application decision is based on characteristics that the government could observe, saving the government the cost of having to screen them. Second, they also apply based on characteristics that are unobservable to the government, which saves the government considerably more: it reduces the chance that a rich person incorrectly passes the proxy-means test since they would not come to be interviewed in the first place.

We then compare how self-targeting “works” against the status quo policy, an automatic enrollment over a pre-selected list. Self-targeting reduces error and it does so at a lower cost. In fact, due to the selection on unobservable characteristics, self-targeting leads to a poorer group of beneficiaries, even when compared with a full-census, automatic enrollment system. While it is cheaper overall, a worry in self-targeting is that it shifts the costs of targeting from the government to households (who now need to travel to get there, wait, etc.). Though both experimental evidence and model simulations, we show that it only takes a small “cost” to induce targeting—thus, one can induce self-selection without imposing large, distorting costs on households.

The second paper is “**Does Elite Capture Matter? Local Elites and Targeted Welfare Programs in Indonesia.**” Systematically investigating Indonesia’s four major targeted transfer programs—which allow for varying levels of local discretion in the allocations of benefits—we show that there is some evidence of elite capture in the targeting, but the magnitude is very small. It is also

limited to formal elites: informal elites and their relatives are *less* likely to access the programs, perhaps to maintain their roles as village decision-makers.

We then experimentally varied the extent of elite influence—by varying who was invited to help select beneficiaries—on PKH targeting. This replicated my AER paper on community targeting, but in a high stakes environment (\$130/year for six years). We confirm that the targeting error from elite capture is small. Most of the error is driven by administrative challenges stemming from data collection and entry, as well as noise in the prediction formulas. This has motivated us to work with computer scientists to discern whether we can utilize cutting-edge computational techniques to improve the formulas’ predictive power; preliminary work shows that conditional on a given set of data, one can improve the formulas marginally, but not dramatically so.

## b. DO WE OBSERVE DISCRIMINATION?

If bureaucrats have discretion over processes, one concern is that their personal tastes for certain types of groups may prevail. This can have real consequences: if teachers treat minority children badly, they may become de-motivated and leave school. If minority entrepreneurs are treated differently when they apply for government loans, their businesses may stall.

Testing for discrimination is challenging, since it requires contexts where we can observe whether individuals who are “identical” aside from the group characteristic in question are treated differently. Leigh Linden and I found such a setting in exam grading (**“Discrimination in Grading,” *American Economic Journal: Economic Policy***). We recruited children to compete on an exam for a large financial prize. We then recruited teachers—both public and private—and provided each with a set of exams to grade. We randomly assigned child “characteristics” to each exam’s cover sheets in order to ensure that there would be no systematic relationship between the characteristics observed by the teachers and the exam quality. Therefore, any effect of a “characteristic” on test scores can be attributed to discrimination.

We find that teachers discriminate more against lower-caste children, but that this effect is practically small and would only slightly change a student rank in the test score distribution. On average, we do not observe discrimination by gender. We find that subjectivity of the test questions—which allow for more discretion through partial points—do not drive the results.

Importantly, the data appear more consistent with statistical discrimination than taste-based: randomly ordering the exams, we show that teachers discriminate considerably more against children who are graded earlier, implying that graders utilize demographic characteristics when the testing instrument or grade distribution are more uncertain. This implies that policies aimed at making graders more confident in the testing techniques may, perhaps, reduce the dependence of grading on child characteristics.

## c. GOVERNANCE AND THE NATURE OF RULES

An important question debated in the literature has been whether corruption simply gets around bad rules, thereby increasing efficiency (“grease the wheels hypothesis”), or whether it exerts a large social cost through the misallocation of social services? We examined these questions in the context of driving licenses. This is a unique setting because one can easily define and measure whether a misallocation has occurred, i.e., whether bad drivers have gotten licenses (**“Obtaining a Driving License in India: An Experimental Approach to Studying Corruption,”** joint with

Marianne Bertrand, Simeon Djankov, and Sendhil Mullainathan; *Quarterly Journal of Economics*).

We showed that if one wants to receive a license in the minimally acceptable amount of time, one can do so, providing evidence that corruption may grease the wheels. But this has a social cost: many who obtained a license did not know how to drive. In fact, knowing how to drive was not correlated with being able to pass the licensing test.<sup>1</sup>

We then asked the question: what determines whether a rule is broken? In an audit experiment, also described in the paper, actors were sent to the licensing offices and asked if it was feasible to obtain a license (and the price) if one needed to break a particular rule. We find that rules that, by their nature, are difficult-to-verify are cheap to break: for example, it is hard to observe whether a bureaucrat actually administered a driving test after the fact, and so it is relatively easy to bribe your way around this rule. In contrast, the residency rule requires submitting paperwork on where you live, which is relatively easy to monitor. Therefore, while you can still break this rule, it costs more to do so.<sup>2</sup> Thus, corruption had little to do with circumventing inefficient rules and was simply based on the likelihood that the bureaucrat would be caught.

We also wrote a policy version of this paper (“**Corruption in Driving Licenses in Delhi**”; *Economic and Political Weekly*). This version also provided new analyses. In a licensing office that happens to be in a high-profile government area and is thus naturally monitored, we observed very little corruption. However, relatively few licenses are given out. Individuals who live in this area get licenses at other offices where it is easier to bypass rules. This speaks to two key points. First, just because we observe little corruption does not mean that we have efficient service delivery—illustrating why it is important to study real service outcomes, rather than just corruption levels. Second, this shows the dangers of *unsystematic* monitoring: in studying corruption, it is important to look at the different ways in which it can simply be re-routed.

#### d. MONITORING THE RULES

The driving license paper illustrated that bureaucrats are responsive to how inherently easy it is to get caught breaking a rule. This leads to the question of whether finding a mechanism to increase the probability of being caught can reduce bad behaviors. What should we monitor? Who should do the monitoring? Does monitoring alone work, or do incentives matter?

We first explore these questions in “**Incentives Work: Getting Teachers to Come to School**” (joint with Esther Duflo and Stephen Ryan; *American Economic Review*). Often, teachers are incentivized based on their added value in terms of student test scores (the outcome that we care about). However, the worry is that teachers will teach to the test and so real learning will not occur. We take a different approach: we evaluate whether contracting on teacher inputs—specifically, attendance—ultimately improves test scores.

Worldwide, teachers are often fraudulently absent: in our setting, they attended work 60% of the time. Monitoring is difficult, since many schools are very remote. Thus, we devised a monitoring solution: we introduced tamper-proof cameras to a random subset of NGO-run schools and paid

---

<sup>1</sup> Everyone in the study was offered free driving lessons after its completion.

<sup>2</sup> This paper was one of the few to document “agents,” illegal brokers who facilitate the process of obtaining services in many countries, thus providing insight into the industrial organization of corruption.

the teachers based on the number of valid class photos they took each month. Control group teachers continued to receive their normal fixed salary.

Teacher attendance in the treatment schools increased by almost 20 percentage points relative to the control schools. Treatment teachers did not appear to be slacking in any dimensions other than attendance, and *ultimately, test scores starkly rose*.

What drives these results? Was it the increased monitoring, or the incentives? We tried to disentangle the two effects by taking advantage of a discontinuity in the incentive structure: if treatment teachers have not attended a certain number of days in a given month, they will then be paid the same base salary no matter what. Thus, they would not be incentivized on the last day of the month, but the incentive would “kick in” again on the first day. We exploited the discontinuity to estimate the teachers’ marginal utility of money and find that the incentive itself is responsible for most of the observed program effect. We then used the model to estimate the cost-minimizing incentive scheme.

Methodologically, this paper was innovative because it was one of the first in development economics to combine a randomized experiment to measure reduced form impacts with a structural analysis to understand the mechanisms through which we found an effect. In their Annual Review of Economics Paper on “Structural Estimation and Policy Evaluation in Developing Countries,” Todd and Wolpin describe it as one of the five key examples of using discrete choice dynamic programming methods for policy evaluation. I subsequently combined experimental and structural techniques in a number of my papers (including the self-targeting work and the networks paper).

After the experiment, the NGO scaled up the cameras program to all its schools, changing the incentive scheme based on the optimal design predicted by our model. As of March 2011, their teachers exhibited an 80 percent attendance rate. Using yearly test data, I have compared the test scores in control schools before and after they received the cameras program against treatment schools that had always had the program (controlling for fixed effects of schools, time, and trends). In doing so, I find an effect on test scores when the program is introduced to the control schools, providing further evidence for its success (unpublished results by the author).

The National Rural Health Mission (NRHM) of Karnataka, India, expressed interest in adopting a similar program in the *government-run* primary health centers in Karnataka. They planned to pilot fingerprinting devices, which are attached to a cell phone that uploads the attendance data to the NRHM’s control room, in about 140 primary health centers (which served a total of about 2.5 million households). The idea was to use this attendance data to better enforce the leave policies on the books. Iqbal Dhaliwal and I helped design the pilot of the program as an experiment (“**Making a Deal with the Devil: Experimental Evidence on Bureaucratic Reform in India**,” under submission). The evaluation component was one of the first grants under USAID’s DIV Innovation Fund; in his remarks at the FICCI-USAID Innovation Forum in India in 2011, Dr. Rajiv Shah, the administrator of USAID, talked about the study when highlighting USAID’s work.

The program led to a 15 percent increase in the attendance of health care workers relative to the control. The effect was entirely due to lower-level civil servants (e.g. nurses and pharmacists), with doctors exhibiting no change in behavior. However, deliveries conducted by doctors increased by 16 percent in the treatment area, with more deliveries in larger hospitals. Some of this may have been due to nurses with better attendance sending women with high-risk

pregnancies to better hospitals. However, we also find evidence that the treatment staff may have directed women to costly private practices and also prevented the women from claiming their state entitlements (cash and in-kind payments) for institutional delivery, which the staff could then siphon off. In short, better attendance monitoring reduced one particular form of leakage (fraudulent absenteeism), but it may have exacerbated other forms in response.

On net, however, overall health improved dramatically, with a *26 percent decline in low-weight births*. The level of ante-natal visits was already quite high and was not affected by the treatment, but there is some evidence that the quality of ante-natal care (e.g., distribution of iron tablets) increased, which could account for the results. This suggests that even improving attendance among lower-level medical staff can have real impacts on health.

Importantly, the full potential program gains were not realized, since it was only partially implemented: the monitoring occurred, but the state government never used the better data to deduct the employees' leave balances. In part, this is because this requires a lot of effort. Moreover, given the growing private sector in cities and the remote locations of government hospitals, state officials claim that they must give the staff more leeway along dimensions other than salary ("work-life balance"). This view is somewhat merited: the treatment workers who were monitored more report significantly less satisfaction with their jobs, and the treatment hospitals attract fewer nurses, lab technicians, and pharmacists than the control ones. Over time (and without researcher intervention) the monitoring program also fell out of use and was cancelled.

In short, this project highlights that while there are potentially large gains from reducing leakages in government, top-down monitoring may be an impractical way to achieve them. Given this, I have explored "bottom-up" monitoring in more recent projects. Specifically, citizens may be more effective than central governments since they observe day-to-day what the local officials are doing. We worked with the Government of Indonesia to help them design an experiment to test the impacts of a transparency pilot program on leakage within their largest targeted transfer program, a subsidized rice program called "Raskin." In particular, we tested what happened when the Indonesian Government mailed identification cards, which provided information on both eligibility status and entitlements, to program beneficiaries ("**The Power of Transparency: Identification Cards and Food Subsidy Programs in Indonesia**," joint with Abhijit Banerjee, Jordan Kyle, Ben Olken, and Sudarno Sumarto; under submission).

The identification cards changed households' beliefs about their eligibility status and led to a *36% reduction in leakages*. One worry is that, due to targeting error, benevolent local leaders—who have much more local information than the central government—may have been reallocating rice from rich, eligible households to poor, ineligible households, and thus this intervention would undo such "fixes." We find no evidence that this is the case.

Experimentally varying different features of the program, we show a number of important dimensions of how transparency can work in practice. First, we show that publically revealing the beneficiary lists in addition to mailing the cards increases the efficacy of the cards; this treatment also increases protests, which suggests that the public information allows households to coordinate their protests against the local leader. Second, providing information to citizens can also provide a signal to the official that the central government is monitoring him more along this dimension. Through multiple pieces of evidence, we show that the observed effects are likely not driven by this channel, but are likely due to citizen empowerment.

Note that the cards program was scaled up to 15.5 million poor households for a number of different social programs in 2013.

In another project, also conducted with the Indonesian Government within the context of Raskin, we tested what happens when you not only provide information, but when you give villages the choice over who has the *right to distribute* a transfer program. Specifically, we compare the status-quo distribution process (local bureaucrat in charge) with a system where a competitive bidding process occurs to select the distributor. This study is joint with Abhijit Banerjee, Jordan Kyle, Ben Olken, and Sudarno Sumarto; we are currently preparing a manuscript.

Villages choose a new distributor when a very high co-pay is charged. Comparing the winning and losing bids, villages choose distributors who offer lower prices, as well as those with experience in business. Six months after the bidding process, we observe small reductions in the co-pay that citizens are charged. Interestingly, quality is not cut to reduce price, and if anything, citizens report an increase in rice quality.

Through two experimental variations, we show that competition drives these results. First, we show that price reductions only occur in villages (selected at random) in which we imposed a minimum number of bids. Thus, competition in the bidding led to lower prices. Second, since the bidding process requires that information announcing how Raskin operates be released to potential bidders and local elites, this form of transparency could drive the results. Thus, we included an experimental treatment arm that simply provided this information. The information treatment had no impact, ruling out that the results are driven by just information alone.

#### e. WHO SELECTS IN TO THE BUREAUCRACY

What types of people are attracted to government? Using a simple model, Shing-Yi Wang and I illustrate that selection into the bureaucracy depends on whether the relative returns to specific characteristics are higher in the public versus private sector, and that screening on a particular characteristic (e.g., ability) can result in more or less of a particular type of individual (i.e., dishonest types), depending on the correlation between the characteristics. We then recruited about 650 college students to collect data to test the model ("**Dishonesty and Selection into Public Service in India**").

Our key measure comes from a dice task: students throw a die 42 times and are paid based on the outcomes. While we cannot directly observe lying, we can statistically back out the likelihood that they have cheated. Cheaters are more likely to prefer public sector jobs. Screening on ability within the interested pool would not exacerbate, nor improve, this problem.

Importantly, we also administered the dice task to the staff nurses from my Karnataka project. We find that cheating nurses are more likely to be absent, implying that this task predicts real corruption. (I have also now tested local Indonesian officials, and find a correlation between cheaters and real measures of leakage in the rice aid program).

These results have implications for thinking about how one recruits and screens potential civil servants, which is a topic that I hope to further explore in future work. This paper is under submission, but has been widely reported on in the media.

## B. ENVIRONMENTAL POLICY IN DEVELOPING COUNTRIES

The WHO's list of the 20 most polluted cities in the world are all located in developing countries, and primarily in South Asia, with New Delhi claiming the top of the list. Poorer households may face the biggest brunt of this pollution, with weaker housing infrastructure, worse transport methods (i.e., motorbikes rather than cars), already-poor health, and so forth. Moreover, the poor are also exposed to indoor air pollution (IAP), as many rely on traditional fuels for heating and cooking: a 2014 review in the *Lancet* claims that indoor air pollution kills as many as 3.5 to 4 million individuals per year.

Governments are often the biggest providers of environmental safety for poor households, from implementing regulations in order to curb ambient air pollution, to running sewage treatment plants, to designing household fuel policy. Thus, in this stream of research, I first ask: what are the costs of pollution in developing countries? I then ask what can be done: is regulation effective given weak institutional structures? What drives take-up of environmental products, given household perceptions, the costs of protecting oneself and others, and learning failures?

### a. UNDERSTANDING THE HEALTH & LABOR MARKET EFFECTS OF POLLUTION

There has been considerable attention paid to the health costs of pollution in developed countries. However, data limitations have traditionally made it hard to study this issue for developing ones. I aim to overcome these limitations and offer evidence on this imperative topic.

A key project in this area was to try to measure the effect of pollution on infant mortality. These results are described in **“Does the Effect of Pollution on Infant Mortality Differ between Developing and Developed Countries? Evidence from Mexico City,”** which is accepted at the *Economic Journal* and is joint with Paulina Oliva and Eva O. Arceo-Gomez.

The first challenge was how to distinguish pollution's impact from other factors, such as income or education. Turning to the sciences, this led us to use thermal inversions—a meteorological phenomenon where pollutions gets trapped in the lower atmosphere—as an instrument for pollution. To the best of our knowledge, this is the first paper in the environmental economics literature to use this strategy. We find large effects of both carbon monoxide (CO) and particulate matter (PM<sub>10</sub>) on infant deaths. We show that this is not simply due to harvesting, so these costs are real. Importantly, our estimates for PM<sub>10</sub> tend to be similar (or even smaller) than US estimates, while our findings on CO tend to be larger. We provide suggestive evidence that a non-linearity in the relationship between CO and health explains this difference.

In a second project, I explore the effects of *indoor* air pollution (IAP). Consisting of a six-year field project spanning 2500 households, this project has cumulated in a paper entitled **“Up in Smoke: The Long-Run Impacts of Improved Cooking Stoves.”** It is joint with Michael Greenstone and Esther Duflo and is R&R at the *Applied Economic Journal: Economic Policy*. There had been qualitative evidence that improved cooking stoves reduce IAP and improve health. Thus, we decided to implement a randomized trial of an improved stove program in order to quantify the relationship between pollution and health, as well as between pollution and labor productivity.

However, the experiment revealed that while the stoves were very effective in reducing IAP in laboratory tests, they did not have large, sustained impacts in practice. At first, the stoves reduced smoke exposure (as measured by a CO analyzer), although the effect was not as large as it could have been since many households chose not to adopt them or use them properly. After the first

year, the effect disappeared as households updated their beliefs on the technology—they used them less, the stoves broke and were not repaired, etc. We find no observable effect on health.

In short, this study had two broader implications for how we think about evaluating environmental products. First, while many studies focus on short-run effects (6 months to a year), we show the importance of understanding impacts after people have time to update their beliefs on products, as well as capture how quickly technologies degrade. Second, we illustrate the importance of evaluating products in field settings, as “normal” use may cause “real” field outcomes to look different than when they are tested in the laboratory.

This paper has been featured in the *Washington Post*, the *Boston Globe*, the *New York Times India Ink*, *Nature*, and other publications. More recently, Nicholas Kristof and Sheryl WuDunn highlighted it on NPR as an example of why evidence is needed for policy. In addition to the main paper, we have written two policy papers published in **Surveys and Perspectives Integrating Environment and Society** and in **Economic and Political Weekly**, as well as an op-ed for the *Indian Express*.

This project led to two follow-up ones. One of my initial goals of the stove experiment was to analyze the relationship between pollution and work outcomes, but the preliminary analysis soon convinced me that it would not be feasible in that context. Thus, I was motivated to find another context to look at this relationship.

Working with Paulina Oliva, we studied the relationship between pollution and work outcomes in Mexico (“**The Effect of Pollution on Labor Supply: Evidence from a Natural Experiment in Mexico City**,” *Journal of Public Economics*). In the early 1990s, a large refinery was closed in Mexico City. In its vicinity, there was a 20 percent decrease in pollution. We exploit the timing of the closure, as well as the distance of a household from the refinery, to estimate the effect of pollution on hours worked. We find that the decrease in pollution led to a 1.3 hour (or 3.5 percent) increase in work. We then show that the effects do not appear to be driven by labor demand shocks as a result of the closure, nor differential migration.

This work is important for several reasons. First, understanding whether environmental regulation can have positive effects on growth is important, particularly for developing countries, which may be hesitant to enforce regulations. Second, estimates of the labor supply effect are necessary to design optimal taxation policies: the fact that we observe a large and positive effect of pollution on labor supply implies that the optimal tax on pollution is closer to the marginal damage of air pollution. Third, the findings provide another reason for developing countries’ absenteeism problems: high pollution can drive health-related absenteeism.

Given that the stoves were a bit of a failure in improving health, I next asked: would households naturally use better fuel as they get richer? Paulina Oliva and I explore this in “**Moving up the Energy Ladder: The Effect of an Increase in Economic Well-being on the Fuel Consumption Choices of the Poor in India**.” This paper is forthcoming in the *American Economic Review Papers and Proceedings* (May 2015).

We use a simple model to first show that the effect of a capital shock on the purchases of dirty fuel is ambiguous, depending upon the degree of substitutability with clean fuel and the degree to which labor supply is affected by dirty fuel use. To test the model, we then exploit an experiment that gave households an asset to generate exogenous, long-run changes to households’ economic

status. We find that overall fuel consumption rose with higher income status, but in the same proportion as non-fuel expenditures. Treatment households were more likely to use electricity as their primary lighting source (rather than kerosene) as a result of the treatment, but also increased their overall kerosene use. On the other hand, we do not observe treatment households switching to a better-quality fuel source. As Wolfram, Shelef and Gerler (2012) discuss, the current estimates of fuel growth do not take into account changing demand from the poor (or even the near poor). This paper illustrates that this is indeed a problem if current policy planning is designed around these underestimates.

## b. EFFECTS OF ENVIRONMENTAL REGULATION

How can policy reduce pollution? In addition to the evaluation of Mexico's refinery closure above, I ask three key questions in this space: (1) Does tougher environmental regulation in rich countries induce them to ship their pollution abroad? (2) Do inspections reduce pollution? (3) Do regulations in poor countries have bite, given their weaker institutional structures?

First, I tested what happened to the foreign manufacturing of U.S. firms when their manufacturing plants faced tougher environmental regulation due to the Clean Air Act Amendment (CAAA) Regulations. I obtained access to firm-level U.S. foreign investment data at the Bureau of Economic Analysis. Next, I needed plant-level data in the US. I found a series of print marketing yearbooks and—somewhat ironically—outsourced their data-entry to India. I then manually matched both datasets using parent firm histories on all their subsidiary plants. Once I accumulated all these data, I calculated the percentage of plants that each parent firm owned in that year that were subject to the CAAA regulations in a given year. Then, I tested whether having a greater percentage of manufacturing plants fall under CAAA regulation had an effect on outbound FDI, controlling for county-level trends, industry trends, and firm-specific fixed effects.

The findings are in a paper entitled “**U.S. Environmental Regulation and FDI: Evidence from a Panel of U.S. Based Multinational Firms,**” which is published in the *American Economic Journal: Applied Economics*. I find evidence that the CAAA regulations increased FDI. The data suggest that the regulated firms took advantage of the under-utilization of existing foreign plants, rather than creating new plants in new countries. Thus, somewhat unsurprisingly, the ratio of FDI in developed countries to that of developing countries did not change when firms faced tougher U.S. regulation. In sum, regulation in the US shifted manufacturing abroad, but not disproportionately to developing nations.

As part of the previous paper, I collected plant-level data on the environmental inspections of U.S. plants through a Freedom of Information request. I ended up using the data for another project entitled “**The Impact of Inspections on Plant-Level Air Emissions**” (joint with Paulina Oliva; *Berkeley Electronic Press Journals*). We empirically test whether inspections have an effect on pollution levels using a 17-year panel dataset for about 17,200 manufacturing plants. We show that firms react to an actual inspection (i.e., to avoid further fines), rather than reacting to just the threat of inspection. We also show that they do not multi-task; i.e., they do not increase other types of emissions to compensate for being monitored for air pollution.

While this illustrates how environmental regulations may have real impacts in developed countries, it is not clear that regulation would have similar results in developing countries, given the weak institutional and governance structures often observed. Thus, Michael Greenstone and I

test whether environmental regulation reduced pollution in India (“**Environmental Regulations, Air and Water Pollution, and Infant Mortality in India,**” *American Economic Review*).

One challenge in such a project is data. Thus, I compiled 25 years of pollution readings across cities in India from a series of print books, performed an extensive historical mapping of India’s environmental policies, and collected infant mortality data from both print books and from each state. Combing all of these data, we find that the air pollution policies contributed to the observed decline in air pollution. In contrast, the water pollution policies had no observable impacts. The historical context of the policies provides insight into these observed differences: the air pollution policies were instigated by citizens and an activist Supreme Court. In contrast, no one took responsibility for the water pollution policies.

In short, the findings show that environmental regulation is feasible, even in low-income nations, if there is political will. The work garnered considerable attention, being written up in the New York Times India Ink, among others. Notably, Jairam Ramesh—a member of parliament and former member of the cabinet ministry who has held some of the top environmental positions in India—discussed the findings in his recent opinion piece in *Mint*.

### c. TAKE-UP OF ENVIRONMENTAL PRODUCTS

My stoves project is just one of many examples where individuals may not necessarily adopt a product or idea, even if it has potentially large returns. In 2005, we started a project to explore these barriers to adoption. We focused on agricultural extension advice, evaluating an IFC program geared at seaweed farmers in Indonesia. Initially, we were trying to understand what could be done to improve program take-up. However, we discovered in preliminary work that the advice was not right for all farmers, as the best practices appear to vary by plot characteristics. Therefore, we cancelled the bigger experiment.

However, we were struck by the fact that farmers did not seem to experiment to determine whether this advice was right. Moreover, there was so much natural variation in farming practices within a plot that we were surprised that farmers did not *notice* whether the advice was correct just from this variation. Thus, in “**Seaweed Farming: Learning by Noticing: Theory and Experimental Evidence in Farming**” (joint with Sendhil Mullainathan and Josh Schwartzstein; *Quarterly Journal of Economics*), we argue that learning by noticing is an important feature of technology adoption. If one does not believe that a practice is important, he will fail to notice how it affects outcomes. This makes it hard to learn from information generated by one’s own actions or that of others. However, being presented with a summary of the information may induce learning, as it highlights features of the data the individual has failed to notice.

We then test the model’s predictions using primary data from Indonesian seaweed farmers. We first show that farmers have very strong beliefs about the distance between their pods, leading to very little variation on distance in practice. On the other hand, most farmers do not even know their own pod size and therefore we observe a wide range of pod sizes, even within a farmer’s own plot. We then conducted trials with the farmers where we varied both pod size and distance between pods to determine if their current practices were optimal. We found that on distance (which farmers already noticed), they were very close to the optimal, but that farmers could gain considerably if they changed their pod size. Consistent with the model, participating in the trial had no effect on subsequent behavior, but seeing the summarized data did.

This work has been highlighted in the 2015 World Bank *World Development Report* on “Productivity,” as well as in Science’s Editor’s Choice section in their June 13, 2014 issue.

## **II. FUTURE RESEARCH AGENDA**

Moving forward, I envision continuing and building upon my current research agenda—finding projects that both contribute to the way we think about how governments function, but also offer concrete policy recommendations.

### **A. PROVIDING PUBLIC SERVICES TO THE POOR**

I have a series of projects in early stages. First, I ask whether government mandates are effective if institutions to enforce it are weak. Indonesia is mandating that everyone has health insurance by 2019, but existing administrative and survey data suggest that take-up is low and that there is adverse selection. This topic was raised as a key policy concern during a brainstorming conference that I co-organized in Jakarta in 2013. With Abhijit Banerjee, Sudarno Sumarto, Ben Olken and Amy Finkelstein, I have begun to work on it, with approval for about a \$3.5 million grant from the Australian Government. We are currently working on logistics pilots to understand what can be done to improve enrollment. Depending on the pilots, we will decide whether to conduct a full study.

Second, I ask whether targeted transfer programs have labor supply effects, and whether how you target affects the magnitude of these effects. To study this question, Ben Olken, Abhijit Banerjee, and Gabriel Kreindler, and I have compiled data from RCTs of cash and in-kind transfer programs across the developing world. We will also study this through follow-up data on labor supply that we collected from the areas where households were selected into PKH through different targeting methodologies.

Third, there are two broad types of questions that I hope to explore next. First, I have often been struck by the idea that much of what we label as corruption or poor performance by bureaucrats stems from the idea that unreasonably high expectations are placed upon them, without the proper tools for success. I had set up a pilot study for government nurses in Hyderabad, India, that made it easier for them to organize their workload, under the idea of testing whether performance would improve if we made work simpler. However, preliminary field work made clear that this was not the right setting to study this question, so I am still seeking other opportunities. Similarly, I have been interested for quite a while in whether private sector competition could improve government performance. While I recently studied whether introducing choice into who distributes transfer programs had effects, I want to do more within this space.

Next, I have often found it fascinating that when people highlight cases where corruption was stamped out, it was often attributed to a single, honest person who forced change. In an earlier project, I showed that in India, a desire for public sector jobs is negatively correlated with experimental measures of honesty and pro-social behavior. Now, I am very much interested in understanding whether you can change how people view public service and whether this translates to different type of selection into the bureaucracy. To then end, Shing-Yi Wang and I have had talks with an NGO in India that does civics training for kids, with the idea of testing whether these types of programs change beliefs on public service and propensity to engage in corrupt activities (as measured by lab tests). We are still exploring whether this is the right setting to study this kind of question.

Finally, I am in the process of trying to develop new projects. In January 2015, I will attend a meeting in Jakarta with the Department of Family Welfare to brainstorm potential new projects in this space (this is part of an executive education class we are organizing for them). In addition, I have been working on a chapter on designing experiments for anti-poverty programs (commissioned for the Handbook of Field Experiments; joint with Dean Karlan) and this has given me the opportunity to take a step back and ponder the questions that I hope to explore.

## B. ENVIRONMENTAL POLICY

I am working on a chapter on the effects of climate change on children in developing countries for a special environmental issue of the Futures of Children (joint with Paulina Oliva).

Moreover, I have been actively raising money and doing background research for two potential projects. First, I would like to study the willingness to pay for reductions in pollution exposure to obtain a better measure of the benefits of environmental regulation, as well as to better understand the extent to which households engage in avoidance behaviors. To this end, Paulina Oliva, Carlos Muñoz, and Daniel Velez Lopez, and I have designed a field experiment that we believe can get at both of these issues.

To inform our experiment before we move forward, we first conducted a survey of 400 households in Mexico City on general beliefs about air pollution, specific knowledge of daily air pollution, and beliefs about the relationship between air pollution and health, and engagement in avoidance behavior (e.g. wearing masks, keeping children inside on high pollution days). The survey was complete in September 2014 and funded with grants from the Weiss Family Grant Program and the Centro Mario Molina. The preliminary analysis suggests that households believe air pollution levels are high and engage in quite a bit of avoidance behavior, but their specific knowledge about pollution is low. We are developing these results into a descriptive piece. Based on our survey findings, we are now refining our study design. We have submitted an application for a grant from the International Growth Centre and are currently exploring other potential grant opportunities for the full experiment.

The other project is one that explores the effect of changes in energy prices on electricity use and asset acquisition in Indonesia (joint with Michael Greenstone and Ben Olken). This is an important question, as many developing countries are grappling with what to do about costly energy subsidies and how to price and regulate electricity. Due to the nature of Indonesia's electricity distribution system, it is possible to vary the marginal costs that a given household will pay for electricity. Thus, we can experimentally test long-run changes in energy prices on household behaviors (i.e., we can randomly subsidize household electricity for X number of years). We won an International Growth Center Grant for the logistics pilot, have hired staff for the project, and are in talks with the Indonesian government for project approval. If the logistics pilot is successful, we plan to launch the full experiment next year.