

RESEARCH STATEMENT PROGRESS AND PLANS REPORT

Rema Hanna
September 2011

1.1. Overview

My field of research is development economics. Within this field, my research has focused on two streams: service provision and the environment.

In the first stream, I have focused on understanding how governments can better serve the poor. Despite the recent liberalization policies observed in many developing countries, governments continue to be the primary provider of essential public services relating to health, education, safety nets, and so forth. However, “too often, services fail poor people—in access, in quantity, in quality” (World Development Report, 2004). Using economic theory as a framework, I empirically test hypotheses on what is driving these failures and try to determine solutions to these constraints.

For example, access is often limited due to challenges in targeting. As the poor typically live in remote areas or are excluded from local institutions, even locating them is tough. The fact that they are characterized both by high income volatility and informal labor markets also complicates the processes of verifying income status. In my research, I have examined how the characteristics of different targeting methods select in different types of individuals. For example, I showed that even though data-driven approaches to learn income may be better at identifying the poor than community-based targeting, the difference in the methods are economically small and are due to unexpected mechanisms: Elite capture is not a problem, but rather communities target on factors that are not captured by current income, such as factors that predict *earnings capacity*. In other work, I have shown that targeting by minority status (as in the case of affirmative action programs) successfully redistributes resources from richer to poorer households.

Improving access will not accomplish much if both the quantity and quality of the available services is low. Corruption and absenteeism are often cited as the biggest impediments to the delivery of services in the developing world, and so I have tried to explore the implications of both. For example, I studied driving licenses to understand whether corruption has efficiency costs. Corruption does help speed the process of obtaining goods, but it does so at a large social cost: bad drivers on the road. More generally, we show that bureaucrats are quite willing to break rules that have a social cost, but are unwilling to break rules that, by their nature, are easier for their superiors to monitor. Therefore, corruption can result in a misallocation of services, imposing large social costs.

One form of corruption is absenteeism: bureaucrats continue to collect a paycheck, and yet systematically fail to show up to work. This has real long-run implications: women do not show up to the hospital for birth or to vaccinate their child if the doctor is rarely there, children have an abridged school year, and so forth. Through a field experiment, we show that monitoring, coupled with high powered incentives, can improve teacher attendance, which in turn drastically improved child learning. Using structural techniques, we show that this effect is driven almost entirely by the incentives, rather than just pure monitoring.

In my second stream, I have focused on environment regulation and policy. Governments face considerable challenges both in determining the optimal tradeoff between environmental quality and growth, as well as understanding how to implement environmental regulations in low governance regimes. Regulation will disproportionately affect the poor: The most polluted cities in the world are located in developing nations, with the poor disproportionately living in the most polluted neighborhoods.

In my research, I have explored the implications of high levels of pollution. In addition to studying the effect of pollution on health, I have also explored how pollution affects economic well-being. Specifically, my work has shown that poor air quality exerts a large effect on work hours, which has important implications. First, given that the poor often disproportionately live in highly polluted neighborhoods, government inaction in the environmental space may have particularly large economic impacts for poverty. Second, understanding the elasticity between pollution and air quality is essential for designing optimal environmental taxes.

The question naturally becomes that given the limited institutions and governance structures observed in many developing countries, can governments do something about this? I explored this question in the context of India, a country that rates low on most measures of governance and additionally has a GDP per capita that is below the level that the Kuznets Curve would predict environmental regulation to “work.” Environmental regulations have appeared to contribute to the fall in air pollution across Indian cities, but the regulations have no observable effect on water pollution. Suggestive evidence implies that differences in citizen/political support may be responsible for the differences in these observed effects of the regulation.

1.2. Providing Public Services to the Poor

Within this stream, my work has centered around two key issues. First, are the right people getting access to services? Second, how do you effectively provide services to those individuals?

1.2.1. Targeting

A key challenge in providing social services for the poor is figuring out who is actually poor. The most obvious way to do so is to conduct an income census. However, this is often very expensive, particularly in developing countries. As the poor typically live in remote areas or are excluded from local institutions, even locating them is tough. The fact that they are characterized both by high income volatility and informal labor markets (where paper records do not exist) also complicates the processes of verifying income status. Finally, weak governance institutions may further dilute the ability of governments to effectively conduct the census, as the process may be captured by local elites or subject to high levels of corruption by the administering bureaucrats. Consequently, governments often try to find alternative methodologies to identify the poor.

One such methodology is to target traditionally disadvantaged groups (low caste groups in India, African-Americans in the United States, etc.) through affirmative action programs. However, these programs remain very controversial for two reasons. First, there is little evidence as to whether targeting by a group characteristic results in transfers to the poor. Many claim that affirmative action only benefits rich individuals from the traditionally disadvantaged group. Thus, if the marginal admit into the program of the “disadvantaged” group is richer than the marginal admit of the “advantaged” group, affirmative action programs may actually be regressive in nature. Second, there is often little evidence of whether these programs actually help the targeted group, which would have implications for the efficiency of many such programs.

With Sendhil Mullainathan and Marianne Bertrand, I examined these issues in the context of the university system in India (the results are discussed in a paper entitled “**Affirmative Action in India: Evidence from Engineering College Admissions in India**” that is published in the *Journal of Public Economics*). India is home of one of the largest affirmative action programs in university admissions: in fact, about 50 percent of slots in universities are reserved for disadvantaged groups.

To study this question, we collected primary data. We obtained admissions data from the engineering entrance exam for 1996, which contained the entrance exam score, caste group, and parents’ contact information. As admissions are based solely on the exam score and caste group, we could determine who from the lower caste groups obtained admissions due to affirmative action (as well as who lost a university seat in the upper caste group due to affirmative action). Between 2004 and 2006, we conducted a survey of the applicants who scored around the admissions cutoff for each major caste group to learn about their family background at the time of admissions and to learn about their educational and employment outcomes around 8-10 years after being affected by affirmative action policies.

We had two key findings. First, caste-based targeting transfers resources to the poor. While those in the lower caste groups who benefited from affirmative action are richer than the average lower caste family in the state, they are indeed *poorer* than those from the upper caste groups whom they replaced in the university system. Our second finding suggests that the lower caste groups targeted by affirmative action benefit from the program in the form of higher wages after graduation, but that the gain is not as large as the loss incurred by the upper caste groups that lose a slot.

This paper not only has implications for the economics literature, it also has real implications for policy. Starting in 2005, affirmative action again rose to the forefront of the policy debate in India as proponents tried to increase their level and scope. In 2007, the Supreme Court placed a stay of order on the parliament bill for the expansion, citing that there was little understanding of whether affirmative action actually transferred resources to the poor and whether it actually worked. In response to this policy climate, in 2008, we were asked to write an op-ed for *Mint* (the Wall Street Journal’s venture in India) to provide insights from our research into these questions.

I have continued my work on targeting with a series of projects in Indonesia. The first explores the differences between proxy-means testing (PMT) and community based targeting approaches. This work was joint with Vivi Alatas, Abhijit Banerjee, Ben Olken, and Julia Tobias, and is described in a paper entitled “**Targeting the Poor: Evidence from a Field Experiment in Indonesia**” that has been accepted at the *American Economic Review*.

We designed a field experiment to explore the trade-offs between the two methods. Theoretically, community members might have better information about poverty, which would cause the community methods to perform better than the PMT. For example, by focusing on the permanent component of consumption, the PMT may miss out on recent shocks. In contrast, community methods may observe these shocks just by observing how their neighbors live. However, the community method has a risk 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, due to elite capture.

We find that proxy-means testing performed better at targeting the poor, based on the common \$2 per day per capita consumption metric, although the effect size was not large. For example, for a typically-sized transfer program in Indonesia, our simulations suggest that the different targeting methods would not yield significantly different effects on reducing the poverty rate in Indonesia.

However, community methods were better at identifying the very poor (those who are below \$1 per day) and community members had significantly higher satisfaction with the beneficiary list and the process when they participated in targeting.

We then tested several different dimensions as to why these processes led to the selection of different individuals. First, to test for elite capture, we randomly divided the community villages so that, in half of these villages, everyone in the community was invited to participate in the targeting meeting, whereas in the other half, only the “elites” were invited. We also gathered data in the baseline survey on which households were related to the local elites. We find no evidence of elite capture: the targeting error rates were the same regardless of whether only the elites attended the meeting and we find no evidence that households that are related to the elites are more likely to receive funds in the community treatments relative to the PMT.

Second, to examine the role of effort, we tested whether the effectiveness of community targeting differed between households that were (randomly) considered first and those (randomly) considered last (when fatigue may have set in). Effort matters: at the start, community targeting performs better than the PMT, but it worsens as the meeting proceeds.

Third, to examine the role of preferences and information, we studied the performance of the methods against the alternative metrics of poverty that we collected in our baseline survey. We find that the community treatments moved the targeting outcomes away from a ranking based purely on per capita consumption and towards the rankings that one would obtain by polling different classes of villagers or by asking villagers to rate themselves. We, then, further investigated how the community’s perception of well-being differs from consumption, and find that they place greater weight on factors that predict *earnings capacity* than would be implied by consumption.

Community based targeting will only be successful if the communities actually have information about one another. Therefore, we further exploit the detailed data that we collected on social networks within each village, individual community members’ beliefs about the income status of others in their village, and the actual income status to test theories from the networks literature on knowledge transmission. This dataset presents a unique opportunity to test network theories, given the large number of networks that we collected as part of this study (640 villages). The preliminary results indicate that knowledge transmits through the network as predicted by theory (higher degree individuals are more likely to be able to correctly rank the income status of their neighbors, those who are more connected themselves are more likely to be ranked correctly by others, etc.). This project is joint with Vivi Alatas, Abhijit Banerjee, Arun Chandrasekhar, and Ben Olken and is entitled “**Identifying the Poor: Social Networks and the Aggregation of Information.**”

The results of the first targeting study in Indonesia led to many further questions. First, given individuals may have better information on poverty status than a data driven approach, could self-targeting mechanisms be designed to identify the poor? Second, would the results from the first experiment (particularly those on elite capture) remain the same even if there was more money at stake? Third, can community information be utilized to update the beneficiary list? In doing so, could the community do the—always controversial—task of removing households from the list that are no longer poor?

To answer these questions, we worked directly with the Indonesian government to set up a field experiment. The project tested these questions in the context of targeting for the actual expansion of PKH (Indonesia’s conditional cash transfer program) to new regions. As such, this project is a

collaborative effort with many different partners: 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). We are currently in the endline phase of this project and look forward to data analysis.

From a policy point of view, it is important to note that both targeting experiments in Indonesia were designed to inform the current overhaul of the government's targeting policies. The current methods of targeting result in high error rates (which are consistent with many programs around the world). For example, one of the current President's cornerstone programs was to provide an unconditional cash transfer (about US \$10 per month) to about 19.2 million poor households during periods of economic crisis. The World Bank estimated that about 47 percent of the poor were excluded from the program. As such, the government has been motivated to improve their targeting methodologies. Therefore, we designed and implemented the experiments in close collaboration with several government bodies to not only be of academic interest, but to also provide evidence on the specific types of targeting methods that the Indonesian government was considering for the development of a unified targeting system for Indonesia starting in 2011.

Finally, we are currently analyzing data from a field experiment in India that sought to understand whether different types of individuals are selected given the structure of the subsidized loan, as well as whether the different types of individuals selected results in different uses of the funds. This project is being conducted with Sendhil Mullainathan and Laura Trucco and is tentatively entitled "**Credit Scoring.**" We set up a micro-lender with eight branches across Tumkar District, Karnataka, India. In the randomly selected "loan villages," we randomly selected individuals to be offered a \$100 loan, and then randomly varied both the interest rate and the ease of repayment (whether debt can be restructured if the individual has a bad, documented shock). In a remaining set of villages, a randomly selected group of individuals were offered a grant, by a local NGO, of equal size to the loans. About half the individuals offered a loan took it up, with a repayment rate of about 50 percent; almost all of these offered the grant took it up. The field work is complete. We are currently analyzing the data to learn whether the loan structure affects the characteristics of individuals who select into the loan and whether this affects repayment/use of the funds. In addition, we are testing whether spending from a grant differs from a loan, and whether this is due to the type of person selected in by a loan.

1.2.2. Discrimination

Discrimination may hinder access to essential public services, including health and education. If teachers treat minority children negatively, they may be more likely to drop out of school. If bureaucrats treat minority entrepreneurs differently when trying to obtain registration licenses, they may be less likely to grow businesses.

In the affirmative action survey, I was particularly troubled by the college admissions test scores for the different caste categories. For example, the most disadvantaged group (the Scheduled Tribe group) had relatively few people sitting for the entrance exam. Most everyone who took the exam was admitted because too few people took the exam relative to the number of slots available, and the test scores of this group were far below the general category. For the 1996 engineering entrance exam, the minimum score of acceptance from the advantaged group was 524 out of 900, while it was 82 from the Scheduled Tribe group. This opened up a lot of questions as to why those from minority groups look so different from those of the majority group. Is it the typical ills of poverty (few educational resources in schools, children having to work, parents with low education levels not having the information to make choices about their child's education, etc.), or did the

children additionally face discrimination by teachers at the school level that discouraged achievement?

In 2006, Leigh Linden and I began to explore these questions in the field setting. This project is reported in an article entitled “**Discrimination in Grading**,” which has been accepted at the *American Economic Journal: Economic Policy*. It is very challenging to empirically test whether children from minority groups perform worse due to discrimination or due to other characteristics associated with poverty. To attempt to do so, we designed a lab experiment in the field to investigate discrimination in grading. In a small-scale city in India, we implemented an exam competition in which we recruited children to compete for a large financial prize. We then recruited local teachers and provided each teacher with a set of exams to grade. We randomly assigned the child “characteristics” to the cover sheets of the individual exams that were to be graded by the teachers in order to ensure that there would be no systematic relationship between the characteristics observed by the teachers and the quality of the exams. Therefore, any effect of the randomized characteristics on test scores can be attributed to discrimination. We found that the teachers give exams that are assigned to “lower-caste” scores that are about 0.03 to 0.08 standard deviations lower than exam scores that are assigned to “high caste.” These differences, however, are practically very small: given the observed test scores distribution, a reduction in score of this magnitude would only slightly change a students’ rank in the distribution.

In designing the study, we collected additional data so that we can not only test for the presence of discrimination, but also shed light on the mechanisms through which discrimination would operate. Specifically, we looked at three pertinent theoretical questions. First, is this taste-based or statistical discrimination? Second, does the subjectivity of the testing instrument increase the scope for discrimination? Third, do the graders exhibit in-group bias, i.e. is there positive discrimination against your own group members? We find that the data appear consistent with statistical discrimination: the teachers tend to discriminate more against children who are graded early in the evaluation process, suggesting that graders utilize demographic characteristics when the testing instrument or grade distribution are more uncertain. We find no evidence that the subjectivity of the test matters. We also do not find evidence of in-group bias; in fact, we observe the opposite, with discrimination against the low-caste children being driven by low-caste teachers.

In addition to contributing to the academic economics literature on discrimination, this paper can provide insights into policy design. First, while the effect of discrimination on the grading of the exams is low, these results suggest that if discrimination exists in the subtle grading of an exam, other more blatant forms of discrimination may exist in the educational system as well. Second, understanding the channels through which discrimination operates can help inform the design of future anti-discrimination policies. For example, given that the graders appear to statistically discriminate, policies aimed at making graders more confident in the testing techniques may, perhaps, reduce the dependence on child characteristics while grading.

We followed up this paper with an op-ed on our findings for VOX, a website for policy analysis and commentary based on academic research.

1.2.3. Governance

Corruption may affect both the access and quality of services. For many years, however, a theoretical debate has raged in the economics literature on the efficiency of this corruption: Does corruption simply get around bad rules, thereby increasing efficiency (“grease the wheels hypothesis”)? Or does it exert a large social cost through the misallocation of social services? We examined these questions in the context of driving licenses in India. This process offers a very

unique setting to study these questions because one can easily define and measure whether a misallocation occurred, i.e. whether bad drivers get licenses. The project is described in a paper entitled “**Obtaining a Driving License in India: An Experimental Approach to Studying Corruption.**” This work is joint with Marianne Bertrand, Simeon Djankov, and Sendhil Mullainathan, and is published in the *Quarterly Journal of Economics*.

For the study, about 800 individuals were recruited and followed through the process of getting a license. About a third were offered incentives to get a license in the legally allowed minimum time (“the willingness to pay” group), a third were given driver’s lessons to ensure that they knew how to drive (“social deservingness group”), and the remaining third obtained the license with no interference from the research group. We found that those who wanted to obtain a license fast were able to do so, providing evidence that corruption may grease the wheels. However, it did come at a social cost: many who obtained a license did not know how to drive. Interestingly, for those who tried to go through the official process, driving skills appeared to be somewhat uncorrelated with getting a license, suggesting that the bureaucrats may have generated red tape to encourage corruption.

A follow-up experiment, also described in the paper, provided further insight into how the nature of the rule affects whether corruption flourishes. Actors were sent to the licensing offices and were asked if it was feasible to obtain a license (and the price of doing so) given that they had to break a rule. We find that rules that, by their very nature, are difficult to verify are easy (and relatively cheap) to break: for example, it is hard to observe after the fact if a bureaucrat actually administered a driving test, 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, and it is relatively easy for someone to check whether the person actually lives within the district. Therefore, while you can still break this rule, it requires faking documents to submit to the licensing office. As such, rules like these that are easier to monitor are harder to break, and when it is possible, it is much more costly.¹ Thus, the existence of corruption had little to do with getting around inefficient rules, and instead was simply a matter of how likely it was the bureaucrat would get caught.

To reach policy makers, we also wrote a policy version of this paper entitled “**Corruption in Driving Licenses in Delhi**” for *Economic and Political Weekly*, a widely read journal in India. This paper not only summarized the QJE paper, but also provided new analyses. Specifically, we provided evidence of the dangers of unsystematic monitoring: in the high profile licensing office that is highly monitored, we observed very little corruption. However, there are also relatively few licenses given out. Individuals who want to get around the rules go to offices where corruption is rampant. Therefore, if an anti-corruption program is only placed in a few locations, an evaluation of the program may give flawed results: corruption may appear to decline in the monitored locations, but in reality, it may simply be re-routed.

One particular form of corruption is bureaucratic absenteeism: government employees fail to complete their duties while still collecting a paycheck. I examined this in the context of schooling in a paper entitled “**Incentives Work: Getting Teachers to Come to School,**” which is joint with Esther Duflo and Stephen Ryan and has been accepted at the *American Economic Review*.

¹ As a side note, this paper was one of the few papers at that time to document the existence and process of “agents,” illegal brokers who facilitate the process of obtaining services in many developing countries, thus providing additional insight into the industrial organization of corruption.

In rural Udaipur, India (the setting of this study), teacher attendance was roughly constant at 60 percent for many years. Part of this problem was driven by the high costs of monitoring teachers, as many of the schools were located in remote, difficult to reach locations. Applying a principal-agent framework, we used randomized evaluation techniques to test whether increased monitoring, coupled with high-powered incentives, could reduce absenteeism. The key was to solve the monitoring challenge, and after piloting several mechanisms, we settled on cameras. Working in Seva Mandir's (a local NGO) non-formal schools, we introduced tamper-proof cameras in 60 of their randomly selected schools. In these schools, teachers were paid based on how many valid photos were submitted each month. Those in the control group continued to receive their normal fixed salary.

Teacher attendance in the treatment schools rose by almost 20 percentage points relative to the control schools. The big fear in incentive-contracts is multi-tasking, where individuals focus their attention on receiving the incentive and reduce their effort along other dimensions. In this case, our fears were unfounded: Child learning levels rose by about 17 standard deviations and teachers did not appear to be slacking upon any dimensions that we were able to observe.

We were curious as to what was driving these quite striking results: was it increased monitoring, or was it the incentives? Therefore, we tried to disentangle the two effects by taking advantage of a discontinuity in the incentive structure. Seva Mandir did not want to fully penalize a teacher if they had a bad shock (illness, crop loss, etc.) that caused them to miss an entire month of work. Therefore, even if a teacher worked less than 10 days in a month, they still received 10 days of pay. Thus, intuitively, at the end of the month, someone who knows that they can no longer work more than 10 days that month no longer has an incentive to show up to work; the incentive "kicks in" again at the start of the next month. Empirically, we find that these teachers sharply increase their attendance starting day 1 of the next month, while teachers who had worked more than 10 days the previous month (and are thus always incentivized) have a roughly constant attendance rate regardless of the day of the month. We used this intuition to estimate the teachers' marginal utility of money, and find that teachers are very responsive to the financial incentive: using our preferred estimates, we find that the incentive itself is responsible for most of the observed effect of the program. We then used the model to estimate the cost-minimizing incentive scheme (based on size of the incentive, and threshold for earning the incentive).²

From a policy perspective, this paper has two take-home findings. First, contrary to what many claim, if there is demand from the top of the organization (district supervisors, principals, etc.), it is possible to introduce monitoring and incentives into schools. Positive incentives, in addition to penalties for poor attendance, drove the political acceptance of the program among teachers. Removing discretion from the system (salaries based entirely on photos) additionally increased the political acceptance as the program was perceived as tough, but fair. Second, there is a large debate over whether non-formal teachers (who are often both cheaper and more readily available in poor, remote areas) can have an impact on child learning. Given their limited education background and training, there is often a call to ban non-formal education centers, even if there are no good alternatives readily available in these areas. This study showed that non-formal teachers *can* increase learning, as increasing their attendance lead to substantial gains in test scores.

Seva Mandir subsequently scaled up the cameras program to all the schools they run. As of March 2011, many years after the original "cameras" experiment, teachers in their schools still experience

² Note that based on the optimal incentive scheme calculated, Seva Mandir subsequently lowered the threshold and increased the bonus.

about an 80 percent attendance rate. However, despite these dramatic and persistent results, few other governments or organizations have tried to adopt these methods. This is why I was particularly excited when the National Rural Health Mission (NRHM) of Karnataka, India, expressed interest in adopting a similar program in their 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 150 primary health centers. If the program is successful, they planned to scale it up to all primary health centers in the state, affecting the lives of over 52 million individuals.

Working with Iqbal Dhaliwal, we aimed both to inform NRHM's program based on absenteeism literature and then to also evaluate the program. We provided key feedback into the program design. For example, we showed that based on the cameras study, monitoring alone would not work, and that instead the monitoring would be best used as a mechanism to enforce the penalties of poor attendance that was already on the books. We then worked with the government to set up a randomized experiment with 150 health centers receiving the system and 150 remaining in the control. Data collection began in July 2010 and is currently ongoing.

1.3. Environmental and Agricultural Economics and Policy

The most polluted cities in the world are located in developing nations, with the poor disproportionately living in the most polluted neighborhoods. Moreover, the poor additionally face indoor air pollution, as many rely on traditional fuels for heating and cooking. Much of my work in this stream focuses on first understanding the costs of pollution on economic outcomes, and then on understanding how effective regulations are in reducing pollution in developing nations.

1.3.1. Understanding the Health and Labor Market Effects of Pollution

In March 2005, I started a research project exploring indoor air pollution (IAP) with Michael Greenstone and Esther Duflo, which is tentatively titled "**The Impacts of Indoor Air Pollution: Evidence from a Randomized Study in Orissa, India.**" Half of the world's population, and up to 95% in poor countries, continues to rely on solid fuels, including biomass fuels (wood, dung, agricultural residues) and coal, to meet their energy needs. We were motivated to understand whether improved cooking stoves were a cost-effective way to improve health, as well as use the exogenous change in pollution induced by the stoves to estimate the relationship between air quality and health and the relationship between air quality and labor productivity.

The field work is complete and we are currently analyzing the data. Preliminary results indicate that while the stoves appeared to have a small effect on indoor air quality when households first receive them, the effect quickly disappears after the first year. This is mostly due to the fact that the stoves break and there is little attempt to fix them, as well as the fact that few households use the stove "properly." Interestingly, the low observed demand for the stove differs greatly from the reported levels of high satisfaction with the stoves. Given the low effects of the stoves on reducing indoor air pollution, it is then somewhat unsurprising that we do not observe an effect of the stoves on improving health or reducing health care expenditures. We are currently analyzing the patterns of take-up, breakages and repairs: specifically, we are asking whether we can predict the characteristics of households that choose to take up the stoves, as well as those who continue to properly maintain and use the stoves.

We expect that an academic paper for the economics literature will arise from this project by Fall 2011. We also expect to then write a policy paper geared towards those who work in the development and health arena. In addition, we have already written two policy papers with insights from the previous literature (published in *Surveys and Perspectives Integrating*

Environment and Society) and findings from the baseline survey (published in *Economic and Political Weekly*, a widely read policy journal in India), as well as an op-ed on the research for the *Indian Express*.

When designing the stove experiment, I was very interested in understanding how pollution affected work outcomes. As we analyzed the data from the project's earlier surveys, it soon became obvious that the stoves only had a small effect on reducing air pollution and that the effect was not sustainable. As such, it would be difficult to measure the relationship between pollution and work outcomes. Since then, I have been very motivated to find another context to study this question, but it has been challenging since one would need to find a context where there has been an exogenous change in pollution that does not directly affect the labor market. Working with Paulina Oliva, we found such a context in Mexico City. In the early 1990s, a large refinery was closed in the center of Mexico City. In the 5 kilometers surrounding the refinery, there was an 8 percent fall in pollution levels. 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.

The findings are described in the paper entitled "**The Effect of Pollution on Labor Supply: Evidence from a Natural Experiment in Mexico City.**" We find that a one percent increase in sulfur dioxide results in a 0.61 percent decrease in the hours worked. 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 many claim are often hesitant to enter international climate change treaties (or are lax about enforcing them) due to fears that tougher regulations may lead to reductions in productivity. Second, an understanding of the effects of pollution reduction on hours worked is necessary to design optimal pollution taxation policies. Specifically, the fact that we observe a large and positive effect of pollution on labor supply implies that the optimal tax on pollution is therefore closer to (and potentially above) the marginal damage of air pollution.

Finally, in a new project joint with Paulina Oliva and Eva O. Arceo-Gomez, we have estimated the causal impact of pollution on infant and elderly mortality in Mexico City. We are now testing whether alleviating information and income constraints allows individuals to practice mitigating behaviors that reduce the effect of air pollution. First, taking advantage of a program that provided announcements when pollution levels hit a maximum point, we are using regression discontinuity design methods to test whether increased information can reduce observed effects of pollution on health. Second, taking advantage of the introduction of a transfer program to the elderly in some parts of Mexico City, we are using difference-in-difference techniques to understand if individuals are better able to practice mitigating behaviors as incomes rises.

1.3.2. Effects of Government Regulations on the Environment

Do rich countries ship their pollution to poor nations? To study this question, I asked what happens to a company's foreign investment in developing countries when the U.S. strengthens its environmental regulations. Specifically, I tested what happens to the foreign manufacturing of U.S. firms when their manufacturing plants face tougher environmental regulation due to the Clean Air Act Amendment (CAAA) Regulations. Under the CAAA, each year, counties whose air pollution concentrations exceed federal standards for a specific pollutant receive a nonattainment designation for that pollutant, while the remaining counties receive an attainment designation. Manufacturing plants that emit a criteria pollutant and are located in a county that is designated as

nonattainment are subject to relatively tougher regulatory oversight than emitting plants in attainment counties.

I obtained access to firm-level U.S. foreign investment level data from 1966 to 1999. The dataset included information on the parent firm in the U.S., as well as detailed data on foreign operations and sales (broken down by country of investment). I then manually matched this data to U.S. plant-level data from printed marketing yearbooks to determine the county and primary industry of each firm's U.S. plants. Combining 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 (i.e. which percentage of plants were both located in a nonattainment county and were in an industry that emitted the criteria pollutant that the county was in nonattainment for). Then, I tested whether having a greater percentage of manufacturing plants fall under tougher environmental regulation had an effect on outbound FDI, controlling for county-level trends, industry trends, and firm-specific fixed effects.

The findings are described 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 CAAA regulations caused multinationals to increase their foreign assets in polluting industries by 5.3 percent and their foreign output by 9 percent. 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 developed nations.

As part of this project, I collected plant-level data on environmental inspections of U.S. plants from the Environmental Protection Agency. I did not end up using the data for the previous project, but I thought it would be interesting to test whether the EPA's monitoring activities of plants actually affected plant emissions.³ This led to a secondary project entitled **"The Impact of Inspections on Plant-Level Air Emissions,"** which is joint with Paulina Oliva and published in the *Berkeley Electronic Press Journals*.

Theoretically, inspections might improve the environmental performance of manufacturing plants through two mechanisms. First, plants may remain in compliance with regulations if both the threat of an inspection and the cost of non-compliance are high. Second, an actual inspection may reduce plant emissions if it causes the plant to improve performance in order to avoid potential penalties. However, inspections might have little impact if the cost of non-compliance is set too low or the inspections fail to catch the worst violators. Moreover, firms may multi-task: if firms are inspected for air pollution, they may increase emissions through another medium to offset their costs.

We studied the relationship between plant-level inspections under the Clean Air Act and plant emissions in the years following an inspection. Specifically, we analyzed data from a 17-year panel dataset of plant-level inspections and emissions for about 17,200 manufacturing plants. The primary difficulty in establishing a causal relationship between inspections and emissions is that a

³ At the time we started this project I was also working on a project exploring the effect of monitoring and incentives on teacher attendance, and therefore, was very immersed in the theoretical and empirical literature on principal-agent models. Even though this was with U.S. data and my focus is development, I thought that this was an opportunity to test these models with this unusually rich data.

plant with higher emissions levels may be more likely to be targeted for an inspection. Thus, a cross-sectional approach may underestimate the impact of an inspection on emissions. Instead, we use an event study research design. This design allows us to exploit the variation in the timing of inspections to understand whether, after controlling for fixed plant-specific and time-varying characteristics, a plant reduces its emissions levels in the period immediately after an actual inspection. In addition, to understand how the threat of an inspection affected firm behavior, we predicted the inspection probability of each firm given observable characteristics, and tested whether this “threat” affected emissions.

Controlling for plant-level heterogeneity, we find that plants significantly reduce their air emissions in the four years following an actual inspection. The magnitude of the effect is quite large: an actual inspection reduces the air emissions of a plant by 15 percent. We find no evidence of multi-tasking, as other types of emissions remain unchanged by the inspections. We find no effect of the threat of inspection on plant emissions. Therefore, it appears that firms react to the recommendations that arise from an actual inspection (i.e. to avoid getting fined), rather than react to just the possibility that they could be inspected.

While these (and other studies) illustrated how environmental regulations may have real impacts on behavior in developed countries, it is not clear that similar types of regulations would have similar results in developing countries, given the weak institutional and governance structures often observed. Moreover, a fairly extensive theoretical and empirical literature on the Environmental Kuznets Curve suggests that given the income level of many developing countries, the demand for the enactment and enforcement of environmental policies may be weak as individuals would be unwilling to trade marginal increases to growth for marginal improvements in environmental quality.

Using data from India, we test whether key environmental regulations had an effect on air and water pollution, despite relatively low governance structures and a GDP per capita that is below the level that the Kuznets Curve would predict environmental regulation to “work.”⁴ This work is documented in a paper entitled “**Environmental Regulations, Air and Water Pollution, and Infant Mortality in India.**” The paper is joint with Michael Greenstone and is currently submitted.

Specifically, we study mandated catalytic converter policies and the Supreme Court mandated Action Plans for air pollution, as well as India’s cornerstone anti-water pollution policy (the National River Conservation Plan). Each of these policies had substantial variation across different cities across different years, allowing us to use difference-in-difference techniques to estimate the casual effect of each policy on pollution.

We find contrary to the Kuznets curve’s predictions, both air pollution policies, especially the mandated catalytic converters, contributed to the observed decline in air pollution across Indian cities during the last 25 years. In contrast, the water pollution policies had no statistically, nor economically significant effect on water quality. This naturally raised the question of why only one type of policy succeeded. While we cannot empirically determine this, the historical context of the policies provides a few insights. Specifically, the air pollution policies were instigated by a frustrated citizenry and an activist Supreme Court. In contrast, the water pollution policies were

⁴ It is worth pointing out that we exploit an incredibly rich panel dataset that includes monthly readings for several measures of air and water across numerous cities (about 140 for air and 420 for water pollution) for almost 25 years. While the recent years of data is available electronically, I preserved many years of older data by collecting print sources from libraries within India and electronically entering the data.

designed by a new (at the time) government ministry with little authority to implement the policies; instead, the authority to implement was shared by a number of different government bodies with no one agent taking responsibility for the program.

1.3.3. Agricultural Economics

Several years back, we started to evaluate a program geared at seaweed farmers in Indonesia that provided the farmers with information on best practices. The goal of the study was to initially understand why farmers were not taking up the advice on best practices, and what can be done to improve the program geared at farmer efficiency. In this preliminary work, we discovered that the advice was not right, or at the very least not right for all farmers, as the best practices appear to vary by plot characteristics. Therefore, we did not follow through with the randomized experiment. However, in the process of our work, we were particularly struck by the fact that farmers did not seem to experiment with the advice to figure out whether it was indeed correct. Moreover, there is natural variation in farming practices within a plot, and we were surprised that many farmers did not *notice* whether the advice was correct by just gauging the data that they amass from their own plots over time.

This led to a project that is entitled “**Seaweed Farming: Learning by Noticing: Theory and Experimental Evidence in Farming,**” which is joint with Sendhil Mullainathan and Josh Schwartzstein. In this paper, we argue that learning by noticing is an important feature of technology adoption and use. Attending to a practice is costly, and as a result individuals only notice practices that they currently believe to be important. The model has several implications. First, if the agent does not believe that a practice is important, he will fail to notice how changes in the practice affect the outcome. As such, it will make it hard for him to learn from information generated by his own actions, his neighbors’, or even his own participation in a trial. However, being presented with a summary of the information on how practices perform may result in learning, as the summary may highlight features of the data the agent has already seen but failed to notice.

We test the model's predictions through a survey of about 500 seaweed farmers and a field experiment with a subset of these farmers. While seaweed farming has several practices that may affect output, we focus on two: distance between initial pods and the size of these pods. Through the survey, we document that a failure to notice may affect practices. Farmers have very strong beliefs about distance between 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 there is therefore a wide range of observed pod sizes, even within a farmer’s own plot.

For the 120 farmers assigned to the treatment group, we ran an experiment where we varied both pod size and distance between pods to determine if their current practices were optimal. We went back several months later to understand if farmers had changed any practices as a result of the experiment, and then provided them with a summary of the results. Sometime after that, we went back again to understand if the information changed their farming practices. 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’s predictions, just participating in the trial had no effect on subsequent behavior, but when we presented the summarized data to farmers that they changed their behaviors.

1.4. Future Work

In future work, I plan to continue working within the field of development economics. In addition to my ongoing work, I have a clear set of questions that I hope to address within the next few years.

Within the corruption literature, much of the work (including my own) has focused on testing principal-agent models: if you monitor bureaucrats more (or increase penalties for getting caught), what happens to corruption? Several influential studies in the literature have shown that the introduction of tougher monitoring and incentives can both reduce corruption and improve the quality of services if there is a commitment at the top to implement these systems. However, this political will often is not there: there may be no incentives for high level bureaucrats or politicians to monitor their staff if they themselves receive bribes to look the other way. Even if there is someone at the top who wants to make the system work, to monitor at a large scale usually requires hiring staff to become monitors, who themselves may become corrupted. If one can design mechanisms to ensure that the monitors are not corrupted, bureaucrats may then think of creative ways to get around the rules or increase the scope of corruption on unmonitored tasks. Therefore, in my future work, I would like to test alternative theories of corruption to think more broadly about designing solutions to improve service provision.

First, individuals may have different preferences or beliefs that may make corrupt behaviors more or less costly to them, conditional on the same set of incentives. Thus, the low quality of services may be a consequence of whether the government selection process results in those who are more likely to participate in corruption. Many case studies suggest that these differences in preference may have large consequences, as these often highlight how one relatively honest individual can transform an organization (see, for example, Klitgaard (1988)'s examples of the Hong Kong Police Force or Singapore Excise Department). In my own research, the monitoring systems have only worked if there is underlying support from someone in the organization who wants to enact change despite facing the same incentives as everyone else for business as usual.

In order to think about designing better selection mechanisms, we must first understand whether (and how) the current government institutions select corrupt individuals. Therefore, I am developing a project (joint with Shing-Yi Wang) in which we will collect primary data to empirically test whether government systems select individuals who have a propensity toward dishonesty. We plan to start the study in India, but hope to eventually extend the work to other countries in order to understand how observable differences in employment processes alter the type of person chosen into the bureaucracy.

Second, rather than focus on incentives to complete a task, I would like to focus on thinking through how the nature of the task that the bureaucrat is assigned may generate both corruption and inefficient behaviors. In many countries, bureaucrats have many tasks to complete and there are no systems in place to help prioritize tasks and manage their work.

Take, for example, a government field nurse in India. She is tasked to conduct sweeps to identify pregnant women, conduct three additional antenatal visits, and then conduct five visits to the household after the birth to complete the full round of vaccinations for the child. She records all of the information from each visit in large, onerous paper registers. As data is recorded in different locations in the paper registers, the nurse herself is often not sure where to go or what houses need a follow-up visit. In reality, no one knows if she actually shows up to work, as her supervisors do not necessarily know what villages she is supposed to be working in that day. Much of the data may be wrong (or more worrisome, faked) since she records the data when she is back at the hospital, rather than in the field at the time of visit. As many details need to be recorded in the books, with no prioritization towards more important information, the nurse gets quickly overloaded trying to enter it all; among all the recorded details, it becomes difficult to find the information needed to make real decisions, such as whether the woman has a high risk pregnancy and needs to be

referred to the hospital. Unsurprisingly, health outcomes are low: for example, only 44 percent of children aged 1-2 years old have received the basic package of immunizations (India National Health Survey, 2006).

Therefore, I have started work with the City of Hyderabad to test whether improving the work environment can reduce nurse absenteeism and improve health. Specifically, a team of doctors have prioritized and streamlined the data that needs to be collected in each home visit. Each nurse will then be provided with a cell phone where they can enter these data at the time of the visit. Each month, the system provides the nurse with a list of women that she needs to visit and what type of visit it is, so she can manage both her schedule and the types of supplies she needs for the visits. Her supervising doctor also receives this list to be able to more easily monitor the nurse if he or she chooses to. In sum, the system changes the process of how nurses go about their duties. I hope to test whether this results in improved service provision, *even* if incentives remain unchanged. In future work, it may be interesting to layer improved incentives over the system, by adding aspects such as GPS monitoring to the system to determine if she made each field visit.

Finally, I would like to understand how providing citizens with information and choice affects the level of service delivery. For the last few years, I have been very focused on understanding how to “find” the poor in order to provide social services to them. However, even if you find the right people, it is not clear that the transfers actually end up in their hands. For example, the World Bank estimates that 50 percent of those targeted in the BLT program (Indonesia’s unconditional cash transfer program) did not receive the full level of funds. Similarly, while about 1.5 billion dollars is spent each year on Raskin (Indonesia’s rice transfer program), the total amount procured by villagers is about half of the rice actually procured by the government. In some ways, these dismal facts are not all that surprising: there is so much misinformation about who is supposed to receive what, and so it is hard to demand your full transfer if you do not even know how much you were supposed to receive. Moreover, the transfers are often run through one intermediary (for example, a village bureaucrat). Thus, this places the applicant and the bureaucrat in a bilateral monopoly setting, which as Shleifer and Vishny (1994) point out may or may not be a plus.

I, along with Ben Olken and Abhijit Banerjee, have begun preliminary talks with the Indonesian government to learn whether we can test theories of citizen empowerment within a transfer program in Indonesia. Specifically, we would try to understand whether increased transparency of the amount and timing of the transfers would reduce leakage. Furthermore, we would also like to test whether introducing competition in the distribution points for the transfers reduces theft.