Essays at the Intersection of Environmental and Development Economics

Essays at the Intersection of Environmental and Development Economics

A dissertation presented

by

Janhavi Nilekani

to

The Department of Public Policy

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Public Policy

Harvard University

Cambridge, Massachusetts

April 2018

Dissertation Advisors:
Professor Rohini Pande
Professor Rema Hanna

Author:
Janhavi Nilekani

Essays at the Intersection of Environmental and Development Economics

Abstract

Air pollution levels in many developing countries are critically high and hazardous to human health. This dissertation studies interventions and policies to reduce air pollution in India. The first chapter focuses on greenhouse gas and local air pollution emissions from road transport. I evaluate whether a randomized introduction of a training program or financial incentives can improve fuel efficiency among public sector bus drivers. The training program increased fuel efficiency in the short term for four months and had no effect thereafter. The incentives scheme increased fuel efficiency for a twelve month period. The second chapter analyzes how personality traits relate to attendance, fuel efficiency, changes in incentive structure, and human resource decisions among public sector bus drivers. I find that personality traits do not predict attendance, rarely predict baseline fuel efficiency, and do not predict responses to randomized training or incentives. Personality traits predict a number of human resources decisions, and the organization's drivers are unusually pro-social and honest. The third chapter focuses on particulate matter emissions from industrial plants. Using granular plant-level data, we pair an engineering model of emissions with an economic framework in which plants minimize costs subject to regulatory constraints. We use this to estimate outcomes under potential pollution control policies. We find that abatement costs are modest and far exceeded by health benefits, and that market based instruments significantly reduce costs compared to command and control regulations.

Contents

List of Tables

List of Figures

Acknowledgments

I am grateful for mentorship, support, and help from numerous people throughout my PhD career.

I am indebted to my advisors Rohini Pande, Rema Hanna, and Rob Stavins for their invaluable guidance. I particularly thank Rohini for extensive mentorship in many roles and for encouraging me to always strive towards excellence, Rema for years of encouragement, optimism, and answers to intricate questions, and Rob for always encouraging and demonstrating ways to have a policy-relevant career. I also thank Nicole Tateosian for many years of cheerful support.

Many friends have made my time at Harvard special, and it is impossible to name them all. For companionship and advice on coursework, field experiments, dissertations, and life problems, I especially thank Nilesh Fernando, Yusuf Neggers, Elizabeth Mettetal, Tomoko Harigaya, Juan Pablo Chauvin, Martin Abel, and Sarika Gupta.

For Chapters 1 and 2, I also received numerous helpful suggestions from participants of the Harvard Development Economics Lunch and Development Economics Tea, particularly Asim Khwaja and Michael Kremer, the Harvard Environmental Economics Lunch and Environmental Economics Seminar, fellow Public Policy graduate students, and anonymous reviewers for the JPAL Governance Initiative and Foundations for Human Behavior grants. I thank the Karnataka State Road Transport Corporation for their support and cooperation, particularly B. Mukkanna and Sumalatha, J-PAL SA for excellent field research support, especially Gowramma and Vasudeva Naik, and Vatsal Khandelwal for exceptional research assistance. I am grateful for funding from the Weiss Family Fund, Lab for Economic Applications and Policy, JPAL Governance Initiative, Harvard Sustainability Science Program, and the Pershing Square Venture Fund for Research on the Foundations of Human Behavior. Part of this research was completed when I was a fellow at the Harvard Sustainability Science Program.

For Chapter 3, my co-author Santosh Harish and I are deeply grateful to Michael Greenstone, Rohini Pande, Nicholas Ryan, and Anant Sudarshan for sharing the data used for this paper, and the Central Pollution Control Board of India and Shakti Sustainable Energy Foundation for funding the data collection. Anant Sudarshan, Rohini Pande, Rema Hanna, Rob Stavins, Richard Zeckhauser, and participants of the Harvard Environmental Economics lunch, Sustainability

Science Program, and Harvard Doctoral Research Seminar provided many helpful suggestions. We are indebted to the J-PAL South Asia Emissions Trading Scheme team for their work on this project, in particular Abhilasha Purwar and Shreyas Gadgin Matha.

My family has supported me tremendously throughout my PhD. Ameeta, Sushil, and Richa gave Cambridge the warmth of family, and Nihar cheered me up. I especially thank my parents Rohini and Nandan, and my parents-in-law Rita and Sanjay, for all their help since my son Tanush was born.

Above all, I thank my husband Shray, without whose support, accommodations, and sacrifices this dissertation would not have been possible.

To Shray

Introduction

In my dissertation, I study a major environmental challenge in developing countries: air pollution. I focus on India, which has some of the worst air quality in the world. The first chapter deals with greenhouse gas and local air pollutant emissions from the road transportation sector, and evaluates interventions to improve fuel efficiency among public sector bus drivers. The second chapter builds on the first, and explores how personality traits relate to job performance among public sector bus drivers. The third chapter focuses on particulate matter pollution from the industrial sector, and estimates pollution abatement costs for industrial plants.

In Chapter 1, "Driving Down Demand for Diesel: Does a Bus Driver Training and Incentive Program Increase Fuel Efficiency?", I study interventions to reduce air pollution from road transport. Road transport is a major source of both greenhouse gas emissions and local air pollution. Large vehicles cause disproportionate damages. Research and policy prioritize improving vehicle quality rather than changing driver behavior, even though driving techniques substantially affect fuel consumption and emissions. I conducted a field experiment in Karnataka, India, randomly assigning public sector bus drivers to two interventions: a training program on safe and fuel efficient driving, and a financial incentives scheme for achieving fuel efficiency targets. The training program increased fuel efficiency in the short term for four months and had no effect thereafter. The incentives scheme increased fuel efficiency for a twelve month period. I find no evidence of any complementarities between training and incentives. Training increased fuel efficiency by a marginally significant 0.0186 kilometers per liter for four months, which saved 0.19% of baseline fuel consumption over twelve months, and had a return on investment of Rs. 3.12 per rupee spent. Incentives increased fuel efficiency by a statistically significant 0.0168 kilometers per liter for twelve months, which saved 0.35% of baseline fuel consumption, and had a return on investment of Rs. 4.22 per rupee spent. Along with the high return on investment from fuel savings, the interventions generated positive externalities from reduced vehicle emissions.

In Chapter 2, "Personality Traits and Job Performance: An Analysis of Public Sector Bus Drivers in India", I contribute to a growing literature in development economics on how personality traits relate to job performance. Research indicates that personality traits of government employees are an important determinant of job performance, and that public service delivery can be improved by taking these traits into account. I analyze how personality traits relate to attendance, job performance, human resource decisions, and changes in incentive structure among public sector bus drivers in Karnataka, India. In a baseline survey of 1522 drivers, I implement experimental measures of dishonesty, pro-socialness, time preferences, and risk aversion. I combine the baseline survey with extensive administrative data on attendance and fuel efficiency, as well as a follow-up survey tracking human resource decisions such as suspensions, transfers, and duty reassignments. I find that personality traits do not predict attendance, and risk aversion, time preferences, and dishonesty do not predict fuel efficiency. Pro-socialness is negatively correlated with baseline fuel efficiency. I do not find personality-related heterogeneous treatment effects from a randomized introduction of training and financial incentives. My results suggest that the organization effectively uses recruitment, hiring, and retention policies to select drivers in a way that among retained drivers, personality traits do not predict job performance. I find that personality traits predict a number of human resources decisions. Dishonest drivers are more likely to be transferred to other branches or reassigned to other duties, pro-social drivers are less likely to be on authorized leave, and risk-averse drivers are more likely to be in attendance, less likely to be disciplined or absent, and more likely to be on authorized leave. Drivers who remain in the organization's workforce are unusually pro-social and honest. I hypothesize that the large wage premium in this setting allows the organization to be highly selective on multiple dimensions during recruitment and retention.

Chapter 3, "Trapping Dust: How Much Would Pollution Control Cost Indian Industrial Plants?", is co-authored with Santosh Harish. We focus on particulate matter pollution from the industrial sector. India has critically high levels of fine particulate matter pollution. A rich literature demonstrates large life expectancy and health benefits to reducing ambient particulate matter concentrations, but comparatively little research has been done on estimated abatement (reduction) costs in India. We empirically estimate the particulate matter abatement costs for small and medium Indian industrial plants if status quo and alternate pollution control policies were rigorously enforced. Our estimates are based on a rich plant-level dataset on emissions, air pollution control devices, and operating practices from 311 industrial plants in Surat, Gujarat. Using this data, we build an engineering model to estimate particulate matter abatement and abatement costs for each plant for possible combinations of air pollution control technology and air pollution control operating practices. We pair the engineering model with an economic framework in which each plant chooses technology and operating practices to minimize costs subject to regulatory constraints. Using this framework, we estimate aggregate abatement and costs under constraints and incentives posed by different pollution control policies. We estimate that fully enforcing the current command and control regulation, an emissions concentration performance standard, would lead to a 66% reduction in average annual emissions, at a moderate average abatement cost of Rs. 36,150 per year ($556). Switching to a market based instrument, such as cap and trade or an emissions tax, would reduce costs by 39%. The health benefits of abatement greatly exceed the abatement costs. The primary hurdle to pollution control is weak monitoring and enforcement of regulations. We conclude that reforms to monitoring and enforcement are crucial, and that a reform bundle should include a shift to market based instruments.

Chapter 1

Driving Down Demand for Diesel: Does a Bus Driver Training and Incentive Program Increase Fuel Efficiency?

1.1 Introduction

India's local air pollution is a rapidly escalating public health crisis. As per the Global Burden of Disease's 2015 analysis, exposure to ambient particulate matter ($PM_{2.5}$) is responsible for 4.2 million deaths globally and 1.1 million deaths in India. Of the 11 most populated countries[1], India's population-weighted particulate matter concentration is the second highest, only better than Bangladesh's and significantly worse than China's. Among these countries, India also has the worst ozone pollution. Pollution levels in India have increased dramatically between 2010 and 2015 (HEI 2017). One study estimated that improving India's air quality to meet national particulate matter standards would increase life expectancy by 3.2 years (Greenstone et al. 2015). India is also the fourth largest greenhouse gas emitter in the world, behind China, the USA, and

1. Specifically, the 10 most populated countries and the European Union.

the European Union[2], though per capita emissions are below the global average[3].

Buses and trucks are important causes of India's air pollution crisis. In India, transportation is a significant source of both greenhouse gas emissions and local air pollutant emissions (Guttikunda and Mohan 2014; Kandlikar and Ramachandran 2000). Estimates suggest that in Indian cities, transport causes around 30-50% of the ambient particulate matter pollution (Guttikunda and Mohan 2014), and around 13-57% of greenhouse gas emissions (Ramachandra, Aithal, and Sreejith 2015). Within India's transport sector, heavy duty vehicles (buses and trucks) have an outsized impact on emissions and are the largest contributor for local air pollutants (Guttikunda and Mohan 2014). State governments operate 10% of all buses, which form the backbone of public transport provision in India and collectively serve 70 million passengers a day[4].

Policies addressing transportation pollution in India have focused on improving vehicle and fuel infrastructure, with minimal attention to changing driver behavior. Key policies include the ongoing Bharat standards for vehicle emissions and fuel quality, and successful past efforts using catalytic converters (Greenstone and Hanna 2014), compressed natural gas (Narain and Krupnick 2007), public transport (Goel and Gupta 2015), and so on. Improving infrastructure, vehicles, and fuel is a necessary component of pollution control. However, though driver behavior is neglected in Indian policy, it also affects vehicular pollution through choices about vehicle miles traveled, when to drive, whether to carpool, and how to drive.

I focus on one aspect of driver behavior - driving style and technique. I partnered with the Karnataka State Road Transport Corporation (KSRTC) to implement and evaluate two driver-focused interventions to improve fuel efficiency and reduce greenhouse gas and local air pollutant emissions. KSRTC is the 5th largest public sector bus service provider in India

2. WRI, CAIT Climate Data Explorer. 2015. Washington, DC: World Resources Institute. Available online at: http://cait.wri.org

3. Ge, Mengpin, Johannes Friedrich, and Thomas Damassa. "6 Graphs Explain the World's Top 10 Emitters." *World Resources Institute* (blog), November 25, 2014. http://www.wri.org//blog/2014/11/6-graphs-explain-world's-top-10-emitters.

4. Calculation: Public sector buses are about 150,000 out of a total bus population of 1.6 million buses. Citation: "About Us" webpage, Association of State Road Transport Undertakings. Available at http://www.asrtu.org/about-asrtu/. Accessed 14th August 2017.

and serves 2.6 million passengers a day[5]. The bus driver's performance significantly affects fuel consumption, which is a large share of costs and thus affects the quantity of bus services, vehicular air pollution, and safety.

I used a randomized controlled trial to evaluate two interventions: 1) an existing training program for 'Safe and Fuel Efficient Driving' (SAFED) techniques, and 2) a pilot financial incentives scheme for achieving fuel efficiency targets. I recruited 1,522 drivers from 34 KSRTC bus depots (branches) and randomly assigned them to four groups: C - Control, T1 - Training Only, T2 - Incentives Only, and T3 - Training + Incentives.

Drivers in groups T1 and T3 attended either a 2 or 3 day intensive training program in 'Safe and Fuel Efficient Driving' techniques such as maintaining moderate speeds, alert driving, and 'eco-driving' practices that are established to promote both safety and fuel efficiency (Barkenbus 2010; Young, Birrell, and Stanton 2011). Drivers in groups T2 and T3 were eligible for financial bonuses of Rs. 500 for 3 months if they achieved a depot-specific monthly fuel efficiency target. These incentives were small, about 1-2% of the driver's salary, as KSRTC wanted to ensure that financial incentives would not crowd out intrinsic motivation and ensure the program would be affordable at scale.

To estimate the impact of these interventions on fuel efficiency, I collated a detailed panel dataset using KSRTC administrative records, with 14 months of pre-intervention data and 12 months of post-intervention data. The dataset has shift-level measures of kilometers traveled (KM), diesel consumption (liters), kilometers per liter (KMPL), vehicle used, route traveled, and other variables. I use a difference in difference empirical strategy with driver and month-year fixed effects to estimate the intent-to-treat effects of training, incentives, and the interaction. I also analyze time patterns and whether there are heterogeneous treatment effects.

Overall, I find the training program increased fuel efficiency in the short term for four months and had no effect thereafter. The incentives scheme increased fuel efficiency for a twelve month period. I find no evidence of an interaction effect. The training intervention had a larger heterogeneous effect on drivers who had previously been trained, while the

5. "History of KSRTC" webpage, Karnataka State Road Transport Corporation. Available at http://www.ksrtc.in/pages/history.html. Accessed on 15th August 2017.

incentives intervention had a larger heterogeneous effect on relatively high-performing drivers with above-median baseline KMPL.

More specifically, I find that over the entire 12 months post-intervention, the SAFED training program had a minimal and statistically insignificant impact on fuel efficiency, which increased by 0.009 KMPL. The incentives scheme increased KMPL by 0.0168, statistically significant with controls. I also exploit the high-frequency panel data to explore monthly and experiment phase treatment effects. For the training program, the overall 12-month 0.009 KMPL increase is composed of a marginally significant KMPL increase of 0.0186 for the 4 months that training sessions were ongoing, followed by no observable effect in later periods. The treatment effect for the incentives scheme was largest during implementation, a KMPL increase of 0.0264, and persisted at a smaller magnitude post-intervention. Drivers who had previously been trained and were attending a repeat training showed a relatively larger gain from training of a marginally significant 0.051 KMPL. High performing drivers with above-median baseline KMPL showed a relatively larger gain from financial incentives of a statistically significant 0.038 KMPL.

I find that the interventions primarily had an effect through the mechanisms of increasing effort from drivers and increasing the salience of fuel efficiency, while I do not find evidence that the interventions changed ability or long term habits. I propose a conceptual framework in which the training and incentives interventions have four main channels: increasing ability by direct investment in human capital through training, increasing effort by resolving principal-agent problems through performance pay, increasing salience of fuel efficiency through reminders embedded in the interventions, and improving habits through training and multiple months of incentives. I compare theoretical predictions based on the conceptual framework to the results for the overall intervention treatment effects, the time patterns of the treatment effects, and the heterogeneity of treatment effects, and conclude that the effort and salience channels are important, while the ability and habit channels are not.

The interventions generated positive externalities through reduced vehicular pollution and potential safety co-benefits. Moreover, both interventions were cost-effective. During the field experiment, I estimate the return per rupee spent on training was 3.12 rupees, and the return per rupee spent on incentives was 4.22 rupees. The training program saved about 0.19% of baseline fuel consumption, yielding a net savings of about $13,230. The incentives scheme saved about 0.35% of baseline fuel consumption, yielding a net savings of $27,410. At scale, I estimate the return per rupee on training would be 2.04 and the return per rupee on incentives would be 1.46. A complete cost-benefit analysis would also include the positive externalities from Safe and Fuel Efficient Driving, in particular safety co-benefits and reduced emissions. As these positive externalities on road injury and vehicular pollution are estimated to be large, I expect a complete cost benefit analysis would show significant net welfare benefits.

This paper makes two key contributions. First, I demonstrate that driver-focused interventions can play a role in improving fuel efficiency and reducing emissions, alongside other policies like improving vehicle and fuel standards. The three month long incentives intervention reduced fuel consumption by 0.35% over one year. Fuel savings may perhaps have been larger if the program had continued. In comparison, Allcott (2011) finds that household-focused energy conservation programs reduced residential energy consumption by about 2% during implementation. Allcott and Rogers (2014) find that the effect of residential energy conservation interventions are quite persistent when the program is discontinued, in contrast to many interventions on exercise, smoking, and other behaviors. I similarly find that fuel conservation persists for several months after discontinuation of the incentives scheme, which I attribute to the heightened salience of fuel efficiency.

The overall finding that driver-focused interventions can have a high rate of return and also reduce emissions is related to similar findings that investing in workers in the Indian private sector can have large returns. For instance, Adhvaryu, Kala, and Nyshadham (2016) find a 250% return on investing in soft skills for female garment factory workers. In the Indian public sector, policy and resources are repeatedly directed towards physical infrastructure rather than human resources and changing human behavior. Toilets are constructed in government campaigns and go unused (Coffey and Spears 2017). Government schools and health centers cover the country but are staffed by absentee teachers and nurses (Chaudhury et al. 2006; Banerjee, Glennerster, and Duflo 2008; Duflo, Hanna, and Ryan 2012). Across sectors in India, investment is needed not just in physical capital infrastructure, but also in the accompanying human capital, processes, and systems. I find that such investments in people can potentially yield high returns.

Second, this paper also contributes to the literature on improving the quality of government services. Existing experiments in developing countries have focused on health and education in particular and also services like policing, justice, and tax collection (Finan, Olken, and Pande 2017). Public transport provision, a key function of the state, has not received much attention, though lower income citizens depend on public buses to access job opportunities and meet other needs.

I find that human capital investment for service providers through training had minimal effect in this setting, though training programs have been successful among health workers in India (Das, Chowdhury, et al. 2016), and a body of case studies and small trials in developed countries have shown positive effects of 'eco-driving' training programs on both safety and fuel consumption (see Young, Birrell, and Stanton (2011) for an overview). Using the conceptual framework, I find that the marginally significant short-term impact of SAFED training was primarily due to the increased salience of fuel efficiency and increased scrutiny from managers post-training, rather than an increase in the driver's ability.

In contrast, I find that a small performance-based financial incentive improved fuel efficiency. In experimental studies in developing countries, outcome-based financial incentives for teachers and health care workers have typically shown positive results (see Finan, Olken, and Pande 2017 for an overview), including among Indian public service providers (Muralidharan and Sundararaman 2011; Singh and Masters 2017). However, the bonuses in my experiment were small compared to much of the current literature. For instance, Muralidharan and Sundararaman (2011) evaluate incentives of about 3% of teacher's salaries, 2-3 times the size of incentives in my setting. Additionally, KSRTC's wariness that financial incentives can undermine intrinsic motivation is a plausible and well-documented concern (Bénabou and Tirole 2006). At these incentive magnitudes I do not observe perverse effects from crowding out of intrinsic motivation in the overall sample.

Taken together, these results suggest drivers have the skill set to improve their driving but invest efforts in doing so only with incentives. This is similar to Das, Holla, Mohpal, et al.'s (2016) finding that the same doctors perform better in private practices with market incentives than they do when practicing in Indian rural public health centers. I also find evidence that fuel efficiency improves when fuel efficient driving is temporarily made salient, which again implies that drivers have the skills but do not always exert effort.

This dissertation chapter is organized as follows. Section 1.2 provides context on how large vehicles are a major source of air pollution in India, on how driving techniques affect fuel efficiency and vehicle emissions, on the Karnataka State Road Transport Corporation, on the two interventions, and on the conceptual framework. Section 1.3 describes the experimental design, including the study area, recruitment, randomization, timeline, data, and empirical strategy. Section 1.4 presents the results, including on take-up, compliance,and attrition, the overall impact of the interventions on fuel efficiency, the time patterns of the treatment effects, an analysis of heterogeneous treatment effects, a comparison of the results and the conceptual framework predictions, and robustness checks. Section 1.5 estimates the cost-effectiveness of the interventions and Section 1.6 concludes.

1.2 Context

1.2.1 Large Vehicles are a Major Source of Pollution In India

Both globally and in India, motorized road transport is a substantial contributor to local air pollution and greenhouse gas emissions. Within the transport sector in India, heavy duty vehicles (buses and trucks) have an outsized impact on emissions and are the largest contributor for local air pollutants. India's local air pollution is severe.

Globally, exposure to particulate matter (PM2.5) is the fifth highest ranking risk factor for death, as per the Global Burden of Disease's 2015 analysis, and other pollutants like ozone are separate additional risk factors. India's air quality is among the worst in the world for multiple pollutants like ozone and particulate matter, and has been deteriorating over the last decade. In 2015, about 25% of the 4.2 million global deaths attributable to particulate matter were in India, second only to China, and India had the largest number of ozone-attributable deaths in 2015 (HEI 2017).

Motorized road transport is responsible for a large number of deaths through road injuries and vehicular pollution. The Global Burden of Disease Study indicates that motorized road transport was the sixth-leading global cause of death in 2010, representing 2.9% of deaths from all causes. Vehicular local air pollution alone caused 184,000 deaths world-wide, with another 1.3 million deaths from road injury (The World Bank and IHME 2014). The transport sector is also a major source of greenhouse gas emissions - about 14% globally in 2010 (IPCC 2014). In India too, transportation is a significant source of both greenhouse gas emissions and local air pollutant emissions (Guttikunda and Mohan 2014; Kandlikar and Ramachandran 2000). One recent study estimated that in urban India, the transportation sector is the largest contributor to greenhouse gas emissions, around 13-57% of total emissions, depending on city (Ramachandra, Aithal, and Sreejith 2015). Evidence suggests that the transport sector contributes 30-50% of the ambient particulate matter pollution in cities (Guttikunda and Mohan 2014).

Buses and trucks cause disproportionate damage to India's local air pollution crisis. In the Indian transport sector, buses and trucks are about 1% and 4.4% respectively of all vehicles, yet are together the largest contributor to local air pollutant emissions (Guttikunda and Mohan 2014).

1.2.2 Driving Style Influences Fuel Efficiency and Vehicular Air Pollution

Driving style is an important influence on fuel consumption, greenhouse gas emissions, and local air pollutant emissions, as indicated by a substantial engineering literature.

Driving techniques that minimize fuel consumption and greenhouse gas emissions, and maximize fuel efficiency, are known as 'eco-driving' and involve practices like accelerating moderately, anticipating traffic and thus avoiding sudden starts and stops, maintaining a steady driving pace, avoiding speeding, and avoiding idling (Barkenbus 2010). Intuitively, driving on highways inherently involves these techniques and partly explains why highway driving is more fuel efficient than city driving or why using automatic cruise control saves fuel. Eco-driving techniques also have the co-benefit of reducing tailpipe emissions (Ericsson 2001; André and Rapone 2009; De Vlieger 1997) which affect the ambient air quality for pollutants like NOx, ozone, and particulate matter.

Along with driving style, fuel efficiency also depends on other factors including road and traffic conditions, the vehicle, and day of the week. Figure 1.1 shows the overall distribution of fuel efficiency in kilometers per liter (KMPL) in the dataset of KSRTC shift-level data.

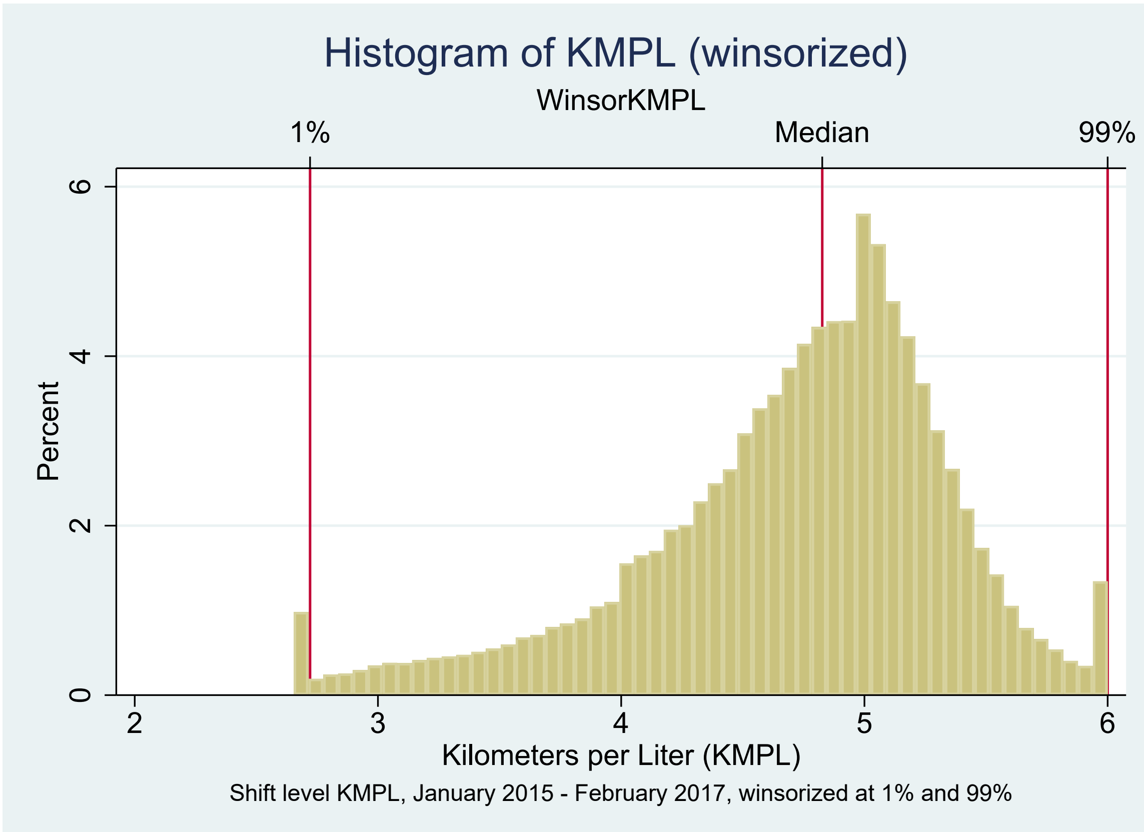

Figure 1.1: Distribution of Kilometers Per Liter

1.2.3 Background on Karnataka State Road Transport Corporation and it's Drivers

I partnered with the Karnataka State Road Transport Corporation (KSRTC), the 5th largest public sector bus service provider in India which serves 2.6 million passengers a day[6]. Table 1.1 contains descriptive statistics on KSRTC driver characteristics from the baseline survey.

Public transport provision in India is predominantly through buses. State governments operate 62 transport organizations that provide intercity and within-city bus services and collectively serve 70 million passengers a day[7]. Another 22 million passengers a day are served

6. "History of KSRTC" webpage, Karnataka State Road Transport Corporation. Available at http://www.ksrtc.in/pages/history.html. Accessed on 15th August 2017.

7. Citation: "About Us" webpage, Association of State Road Transport Undertakings. Available at http://www.asrtu.org/about-asrtu/. Accessed 14th August 2017.

by the nationally operated Indian Railways[8].

KSRTC operates passenger buses for intercity travel and employs over 20,000 bus drivers. It has 14 regional divisions which together contain 79 bus depots. A bus depot is a bus garage where buses are refueled, stored, and maintained, and which also contain administrative offices, driver rest areas, and other facilities. KSRTC is one of four public transport agencies of the Government of Karnataka, one of India's southern states. It is responsible for the southern section of Karnataka.

KSRTC personnel practices for drivers have many features typical of the public sector in developing countries [9] (Finan, Olken, and Pande 2017). Currently, hiring is based on a testing process for applicants who meet certain qualifications, followed by a probation period. However, hiring practices used to be more patronage based and many older drivers were hired through that system. KSRTC drivers are unionized, with a fairly powerful union with frequent strikes[10]. Promotion is based on tenure, and wage increments are negotiated by the union and are about 8-10% a year with an additional annual increment after 15 years of service. Some depots have small performance-based bonuses such as annual prizes for safety. A portion of the driver's remuneration is attendance based, which may reduce the absenteeism typical in Indian public sector settings (Chaudhury et al. 2006)[11]. Though management can suspend drivers, they have minimal ability to fire them. KSRTC drivers are very well compensated, with total remuneration ranging from Rs. 30,000-50,000 per month. New entrants get Rs. 20,000-25,000 a month. In comparison, the salary range for privately employed car drivers in Bangalore is Rs. 10,000-20,000.

8. The Indian Railways report 8,107 million passengers a year, about 22 million a day. Citation: Statistical Summary, Indian Railways, 2015-2016. Available at http://www.indianrailways.gov.in/railwayboard/uploads/directorate/stat_econ/IRSP_2015-16/Summary%20Sheet_Eng_pdf_pdf.pdf. Accessed 14th August 2017.

9. B. Mukkanna. Chief Mechanical Engineer, KSRTC. Personal interviews, 2014-2015.

10. About 5-6 strikes over the 1.5 years of project implementation.

11. KSRTC drivers get a small share of the revenue generated on trips they drive, so their remuneration increases with the number of days worked in a month.

1.2.4 Two Interventions - Training and Financial Incentives

There is substantial variation in fuel efficiency performance among KSRTC bus drivers, which implies there is scope for improvement. Drivers may vary on knowledge, ability, effort, and other dimensions. I evaluate two interventions to improve driver performance: an existing training program for safe and fuel efficient driving, and a new pilot financial incentives scheme for achieving fuel efficiency targets.

First, I implemented an intensive training program that is often conducted by KSRTC. The training on Safe and Fuel Efficient Driving (SAFED) techniques included theory and practical lessons with individualized feedback. The content focused on maintaining moderate speeds, alert driving, and eco-driving techniques including steady speeds, minimum harsh accelerations and minimum gear changes. Aside from these technical lessons, trainers advocated general safety practices such as being well rested, avoiding heavy meals and other sleep-inducing activities before night shifts, and so on. Drivers received motivational messaging about safe driving.

Two variations of the training program were implemented in the experiment. The first was a three day program conducted by instructors from the Petroleum Conservation Research Association (PCRA), an agency operated by the Government of India ('PCRA' training sessions). The second was a two day program taught by Mr. M.D. Haneef ('Haneef' training sessions), an instructor associated with the Andhra Pradesh State Road Transport Corporation (KSRTC's parallel organization in the erstwhile state of Andhra Pradesh). PCRA sessions had batches of about 20 drivers while Haneef sessions had batches of about 35-50 drivers. Overall, about 78% of drivers attended Haneef sessions.

Second, as part of the field experiment, I introduced a financial incentives scheme on a pilot basis for three months. All depots have monthly fuel efficiency targets set by KSRTC. In this intervention, drivers whose average fuel efficiency was above the depot-specific target in the month received a financial bonus of Rs.500 (1-2% of their salary). Treatment group drivers were eligible for financial incentives for three months. Prior to this experiment, some depots did provide bonuses to drivers who performed well on fuel efficiency metrics. However, performance based bonuses are not consistently part of KSRTC's driver remuneration. The incentive was fixed at Rs. 500 because KSRTC believed that the cost of a larger incentive would be too high if the intervention was scaled up. KSRTC management was also wary that a larger incentive would encourage a 'money-minded' culture and adversely affect existing expectations that since "a KSRTC career is a secured job for life, drivers should be grateful and should drive well to support the organization"[12].

1.2.5 Conceptual Framework

The training and incentives interventions could change driver performance through four main channels: ability, effort, salience, and habits. I investigate the importance of these channels by exploring the overall intervention treatment effects, the time patterns of the treatment effects, and the heterogeneity of treatment effects.

The training and incentives interventions have four main channels: increasing ability by direct investment in human capital through training, increasing effort by resolving principal-agent problems through performance pay, increasing salience of fuel efficiency through reminders embedded in the interventions, and improving habits through training and multiple months of incentives. First, the ability channel. Some low-performing drivers may simply lack knowledge on how to drive fuel-efficiently. Improving worker ability through direct training in the workplace has been effective in several settings (Adhvaryu, Kala, and Nyshadham 2016; Das, Chowdhury, et al. 2016; Walker et al. 2016). The SAFED training program tests whether direct human capital investment in bus drivers can improve performance. In this intensive training program, bus drivers were explicitly taught best practices for driving. Second, the effort channel. Drivers may not necessarily exert effort to drive fuel-efficiently, since the rewards for doing so accrue only to the employer KSRTC. Another method to improve worker performance is to align the incentives of the principal, in this case KSRTC, and the agent, in this case the bus driver. Thus, I evaluate performance pay in this setting by piloting a financial incentives scheme for achieving fuel efficiency targets. Third, the salience channel. Knowledgeable drivers who exert effort may still not maximize fuel efficiency if they are attending to other dimensions of driving. The interventions can increase fuel efficiency by increasing its salience. The training and incentives interventions both created several reminders or cues related to fuel efficiency, which itself might

12. B. Mukkanna. Chief Mechanical Engineer, KSRTC. Personal interviews, 2014-2015.

improve performance (Kahneman 2003). Fourth, the habit channel. The interventions could cause a behavioral change in driving habits, and the short-term interventions could thus lead to a persistent long term effect (Allcott and Rogers 2014).

Using this conceptual framework of four channels, I make theoretical predictions on the overall results. First, the ability channel. If this is key, then the training intervention should have a significant effect overall. Second, the effort channel. If this is important, then the incentives intervention should have a significant overall effect - a positive result for the incentives intervention implies drivers have the ability already but don't always deploy it. Third, the salience channel. Both interventions increased the fuel efficiency salience, but the incentives intervention had many more embedded cues. If the impact of incentives is greater than the impact of training, it suggests salience may be relevant - though it is difficult to separate out the effort and salience channels. Finally, habit. If habit is important, I expect a noticeable interaction effect, as the combined intervention of training plus incentives would be a greater shock to preexisting habits.

Using this conceptual framework of four channels, I also make theoretical predictions on the time patterns of the treatment effects. First, the ability channel. If lack of knowledge and weak ability is a primary cause of poor driver performance, and if it can be addressed through training, I expect an immediate level jump in performance post training. I expect the drivers to display a higher performance level which either stays constant, or increases over time as drivers become more adept at implementing new driving techniques. Second, the effort channel. Drivers were told that the incentives scheme was a three month pilot program. If the effort channel is critical, then drivers should only expend extra effort for the three months that they are eligible for financial incentives, after which I predict their performance would return to baseline. Third, the salience channel. For the training intervention, if training is effective primarily because it makes optimal driving practices salient temporarily, then I expect an immediate short run effect on driver performance which then fades and has no long run effect. For the incentives intervention, if the salience channel is important, then the treatment effect should persist post-eligibility as long as drivers receive frequent reminders of the program. Drivers were in regular contact with the field staff for about 6 months from the start of eligibility, so for 3 months post-eligibility[13]. This was due to slow implementation of the incentives scheme and slow disbursal of the incentives. Fourth, the habit channel. The short term interventions could have long term, persistent effects by changing habits. If the habit channel is important, I expect some persistence of the treatment effect after all implementation activities were complete, though this treatment effect might slowly taper off over time.

Finally, I also make theoretical predictions on treatment effect heterogeneity based on this conceptual framework of four channels. First, if the ability channel is important, I expect the training program to have a larger effect on drivers with less training or less ability at baseline. I explore two measures of previous training and two accident-based proxies for driving ability. I test whether the training program has heterogeneous effects on drivers who were formally taught driving, drivers who had previously attended KSRTC training, drivers who had never had a major accident, and drivers who had never had a minor accident. The baseline survey data indicates that drivers have little formal training. Only 4.20 % of drivers learned how to drive a bus by attending driving school. The majority, 71.81 %, were trained through an apprenticeship system, where they were employed by truck drivers who provided on-the-job training. Another 20.11 % were taught by friends and family. However, most drivers, 85.68 %, but not all, had received some training before at KSRTC. Notably, 16.71 % had been involved in major accidents while driving for KSRTC. The baseline survey defined a major accident as “one in which there were any serious injuries or fatalities" and a minor accident as “one in which there were no serious injuries or fatalities, but only vehicle damage or mild injuries". 48.62 % of drivers had been in minor accidents, which are perhaps inevitable in India’s chaotic traffic.

Second, if the effort channel is important, I predict that the financial incentives would have a greater effect on drivers whose baseline fuel efficiency was close to the target and who hence faced low marginal costs of incremental effort. I also hypothesize that drivers with patient time preferences would have less personal costs of driving slowly and not speeding, and may have a larger response to both training and incentives. Finally, I hypothesize that the financial incentives would have a larger effect on dishonest drivers who exhibited a greater propensity to cheat for financial reward. To measure dishonesty, I implemented a dice task, closely following

13. See Section 1.3.2 for details.

Hanna and Wang (2017). I measure time preferences through a series of hypothetical time discounting questions, closely following Ashraf, Karlan, and Yin (2006). I also create a measure of drivers whose baseline KMPL was above or below the depot-specific median. Monthly targets are set at the depot level. Thus, drivers with low, below-median baseline KMPL would have to improve dramatically to achieve the target, whereas drivers with high baseline KMPL would achieve the target relatively easily. The further a driver is from the target at baseline, the more marginal effort he would have to expend to achieve the target, or the target might be out of reach altogether.

Third and fourth, the salience and habit channels. If the habit channel is important, I predict that younger drivers may find it easier to change their driving habits, and explore heterogeneity by age. If the salience channel is important, I predict greater treatment effects for drivers who are intrinsically motivated to drive well and for whom a salient cue activates existing motivation. I look at two behavioral measures, risk aversion and pro-socialness. I hypothesize that risk averse drivers have a larger intrinsic motivation to improve their personal safety, and pro social drivers possibly felt a greater responsibility to drive safely. To elicit risk aversion, I used an ordered lottery selection task, following Eckel and Grossman (2002, 2008). To measure pro-socialness, I conducted a standard dictator game where drivers could divide Rs. 30 between themselves and a charity of their choice out of 7 options. The charity choices and task instructions followed Hanna and Wang (2017).

1.3 Experimental Design

1.3.1 Study Area, Driver Recruitment, Randomization, and Randomization Verification

Out of Karnataka State Road Transport Corporation (KSRTC)'s 79 bus depots, I selected 34 bus depots to participate in the experiment. I recruited 1,522 drivers from these depots and randomly assigned them to four groups: C - Control, T1 - Training Only, T2 - Incentives Only, and T3 - Training + Incentives. I use baseline survey data to verify the randomization.

The 34 participating depots are in 7 administrative divisions in the regions of Bangalore, Mysore, Mandya, Ramanagara, Puttur, and Mangalore. Some KSRTC depots use a bus on board diagnostics (OBD) system that provides extensive administrative data on driver behavior. I selected all 30 such depots and 4 additional depots based on proximity. Appendix Figure A.1 marks the approximate locations of the depots. All the depots are in towns or cities, so buses depart from and return to urban locations, though they may stop at rural bus stops en route.

After selecting the 34 participating depots, I recruited drivers and simultaneously conducted a baseline survey in September-October 2015. In each depot, I aimed to recruit 45 drivers. KSRTC provided lists of drivers with relatively weak fuel efficiency performance in April, May, and June 2015. I targeted these low and medium performing drivers, but also recruited from outside these lists. Drivers were screened to eliminate those who had attended KSRTC training in the previous six months. In total, 1,522 drivers were recruited.

I randomly assigned these 1,522 drivers to one of four treatment arms: Control (N=381), T1 (Training Only, N=380), T2 (Incentives Only, N=381), and T3 (Training + Incentives, N=380). The randomization was stratified on three variables: bus depot (34 depots/categories), age (2 categories: drivers 40 or younger, and drivers older than 40), and whether the driver had previously attended training (2 categories: had attended training and had not attended training).

I use the 2015 baseline survey to check the randomization. Table 1.1 demonstrates that the four groups are overall balanced on baseline survey characteristics. Of the 39 differences in Section II, only two are statistically significant at the 10 % level, somewhat less than predicted by chance. To test if the baseline characteristics jointly predict treatment group assignment, I conduct three regressions comparing each treatment group separately to the control group. Tests for the joint significance of baseline characteristics yield p-values of 0.81, 0.64, and 0.23 for groups T1, T2, and T3 respectively.

Table 1.1: Balance on Baseline Characteristics

	Section I: Means					Section II: Differences		
	(1) C	(2) T1	(3) T2	(4) T3		(1) Training	(2) Incentives	(3) Tr. × Inc.
Age (years)	38.77 (8.444)	38.81 (8.654)	38.94 (8.311)	38.99 (8.442)	.	0.326 (0.319)	0.253 (0.315)	-0.388 (0.445)
Education (years)	10.22 (1.920)	10.25 (1.960)	10.31 (1.813)	10.29 (1.990)	.	-0.00170 (0.119)	0.0818 (0.117)	-0.00831 (0.169)
KSRTC tenure (years)	10.02 (8.059)	10.14 (8.550)	9.903 (7.896)	9.979 (8.060)	.	0.345 (0.369)	-0.0523 (0.360)	-0.369 (0.508)
Job satisfaction (score)	4.598 (0.801)	4.488 (0.892)	4.614 (0.747)	4.470 (0.891)	.	-0.103* (0.0619)	0.0318 (0.0564)	-0.0379 (0.0859)
Risk aversion (score)	3.037 (1.425)	3.021 (1.510)	2.997 (1.461)	3.166 (1.475)	.	0.00168 (0.106)	-0.0234 (0.104)	0.173 (0.152)
Pro-socialness (Rs.)	25.01 (6.765)	24.58 (7.763)	24.28 (7.826)	24.53 (7.716)	.	-0.470 (0.536)	-0.713 (0.528)	0.787 (0.773)
Consistently patient	0.582 (0.494)	0.577 (0.495)	0.608 (0.489)	0.552 (0.498)	.	-0.00343 (0.0357)	0.0238 (0.0358)	-0.0498 (0.0506)
Dice points > 99p	0.0789 (0.270)	0.0923 (0.290)	0.101 (0.301)	0.0802 (0.272)	.	0.0119 (0.0206)	0.0156 (0.0207)	-0.0302 (0.0295)
Dice points > 95p	0.176 (0.382)	0.161 (0.368)	0.161 (0.368)	0.203 (0.403)	.	-0.0189 (0.0269)	-0.0262 (0.0270)	0.0682* (0.0388)
Attended driving school	0.0394 (0.195)	0.0368 (0.189)	0.0341 (0.182)	0.0579 (0.234)	.	-0.000603 (0.0138)	-0.00154 (0.0136)	0.0249 (0.0207)
Had major accident	0.147 (0.355)	0.156 (0.363)	0.181 (0.386)	0.185 (0.389)	.	0.00591 (0.0259)	0.0341 (0.0271)	-0.00474 (0.0383)
Had minor accident	0.470 (0.500)	0.507 (0.501)	0.462 (0.499)	0.507 (0.501)	.	0.0431 (0.0362)	-0.00393 (0.0360)	0.00342 (0.0512)
Attended KSRTC training	0.856 (0.352)	0.858 (0.350)	0.858 (0.349)	0.855 (0.352)	.	0.00225 (0.0254)	0.00262 (0.0254)	-0.00526 (0.0360)
Number of drivers	381	380	381	380	.			
.					.			
Joint F-Test Statistic	.	0.650	0.820	1.257	.			
P-Value	.	0.812	0.639	0.234	.			

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Section I: reported coefficients are means for groups C(Control), T1(Training Only), T2(Incentives Only), and T3(Training+Incentives). Standard deviations are in parentheses. Section II: reported coefficients are from regressions of each baseline characteristic on Training, Incentives, Training × Incentives (Tr × Inc), and strata fixed effects. Robust standard errors are in parentheses. For the stratification variable 'Attended KSRTC training' strata fixed effects are not included. The joint hypothesis tests are from separate regressions of each treatment group compared to the control group. The treatment indicator is regressed on all baseline characteristics. I test the joint hypothesis that all coefficients are equal to zero. Job satisfaction scores range from 1 (very dissatisfied) to 5 (very satisfied). Risk aversion scores range from 1(riskiest lottery selected) to 5(safest lottery selected). Pro-socialness (Rs.) indicates the amount donated (between Rs. 0-30) in a pro-socialness game. Consistently patient indicates the respondent was patient in all time preference questions. Dice points > 99p and > 95p are indicators that the respondent's total dice points was above the 99th and 95th percentiles respectively of the theoretical distribution.

1.3.2 Timeline

The baseline survey took place in September-October 2015. Training took place in March, April, and June 2016. The incentives scheme was implemented between April-December 2016. The two interventions were implemented in three batches. There was a joint timeline, such that for T3 drivers who were assigned to both training and incentives, incentive eligibility always began in the month immediately after training. Appendix Figure A.2 depicts the project timeline.

The training sessions were conducted in March, April, and June 2016. 18 Haneef sessions and 11 PCRA sessions were part of the experiment[14], and about 78% of drivers were trained in Haneef sessions. Managers at the bus depots selected the treatment dates for particular drivers based on driver availability and other factors. For the initial session, I requested depot managers to select from the complete list of drivers randomly assigned to training. For subsequent sessions, I requested depot managers to select from the list of drivers who were randomly assigned to training and were not already trained. For each specific training session, I asked depots to select a roughly equal number of drivers from treatment groups T1 and T3.

The incentives intervention also had three batches, consisting of drivers who were eligible in April-May-June 2016, May-June-July 2016, and July-August-September 2016. T3 drivers whose training took place in March, April, and June 2016 were assigned to incentive eligibility in the months April-May-June, May-June-July, and July-August-September respectively. I randomly assigned T2 drivers to the three batches such that for each batch, within each strata, the number of T2 and T3 drivers was roughly equal. As each incentive batch started, field staff contacted drivers in that batch via phone calls to inform them that they had been randomly assigned to incentive eligibility, and informing them of the specific months they were eligible. Of the drivers not available via phone, some were instead informed in person and a few remained unreachable.

The disbursement of incentives took place on a gradual basis in the months after incentive eligibility. For drivers eligible in April, KSRTC estimated their April average fuel efficiency in May. All April-eligible drivers were contacted in May, June, or July and requested to sign a receipt acknowledging their April average fuel efficiency. If a driver's April average exceeded

14. Training sessions were simultaneously ongoing for non-participating drivers and depots. A handful of participating drivers were trained in those sessions.

the depot-specific April fuel efficiency target, he received Rs. 500. Drivers who didn't exceed the target were also requested to sign receipts. All drivers were reminded of how many more months they were eligible.

Appendix Table A.4 shows that the three treatment groups T1, T2, and T3 are balanced in how many drivers are in each batch. Overall, 45% of drivers are in Batch 1, 38% in Batch 2, and 18% in Batch 3. The table also shows that the training groups T1 and T3 are reasonably balanced in the number of drivers who attended Haneef and PCRA training sessions.

1.3.3 Data

I collected three sets of data through a baseline survey, close monitoring of the interventions, and KSRTC administrative data. I combined these to create a 26 month panel dataset from January 2015-February 2017 for all shifts for all participating drivers.

Baseline data: First, the 2015 baseline survey collected background data on the driver's education, previous training experiences, accident history, job satisfaction, baseline knowledge of SAFED theory, and measures of risk aversion, time inconsistency, pro-socialness, and dishonesty.

Implementation data: Second, I collected significant details on the implementation of the interventions, such as the dates a driver attended training, whether he attended PCRA or Haneef training, the date he received his financial incentive, and so forth.

Administrative data: Third, I gathered extensive KSRTC administrative data on fuel efficiency. I collected shift-level data on kilometers traveled (KM), diesel consumption (liters), vehicle used, route traveled, number of drivers on the shift, date, and other shift-level variables. I use this to estimate shift level fuel efficiency, i.e. kilometers per liter (KMPL) (calculated as kilometers traveled in KM/diesel consumed in liters). For some shifts, shift-level KMPL targets assigned by KSRTC are available. For some shifts, depot-level monthly KMPL targets are collected. The fuel efficiency data starts between 2011 and 2014, depending on depot, and ends in February 2017.

I combine these three sources to create a 26 month panel dataset at the RCT driver-shift level. The dataset has uniform start dates and uses all 385,310 unique shifts between January 1 2015 to February 20 2017.

1.3.4 Empirical Strategy

To analyze the impact of training, incentives, and the interaction on fuel efficiency, I estimate the intent-to-treat (ITT) effects using a difference in difference model.

In Equation 1.1, I estimate ITT effects for the entire twelve months post-intervention. For all drivers, I consider the post-intervention period to be March 2016 onwards, when the training intervention started.

$$\begin{aligned} Y_{imd} &= \beta_1 Training_d \times Post_i + \beta_2 Incentives_d \times Post_i \\ &\quad + \beta_3 Training_d \times Incentives_d \times Post_i + \alpha_d + \delta_m + Controls_{imd} + \epsilon_{imd} \end{aligned} \tag{1.1}$$

Where Y_{imd} is kilometers per liter (KMPL) for shift i in month-year m for driver d. $Training_d$ is an indicator variable for groups T1 and T3, $Incentives_d$ is an indicator for groups T2 and T3, $Post_i$ is an indicator for any shift i on or after March 1, 2016, α_d controls for driver fixed effects, δ_m controls for month-year fixed effects, and ϵ_{imd} is the error term. The robust standard errors are clustered at the driver level.

As drivers were randomly assigned to treatment groups, the β's measure the unbiased intent-to-treat (ITT) effect of the interventions. This is the average difference between treatment and control groups in the post-intervention period (March 2016-February 2017), relative to the average difference in the pre-intervention period (January 2015-February 2016).

I next exploit the high frequency panel dataset to estimate how the treatment effect varies over time. In Equation 1.2 I estimate monthly treatment effects.

$$\begin{aligned} Y_{imd} &= \sum_{m=Sep15}^{m=Feb17} \beta_{1,m}(Training_d \times I_m) + \sum_{m=Sep15}^{m=Feb17} \beta_{2,m}(Incentives_d \times I_m) \\ &\quad + \sum_{m=Sep15}^{m=Feb17} \beta_{3,m}(Training_d \times Incentives_d \times I_m) + \alpha_d + \delta_m + Controls_{imd} + \epsilon_{imd} \end{aligned} \tag{1.2}$$

Where I_m is a set of indicator variables for month-year m and the rest is as in Equation 1.1. I include the six months prior to the intervention as a placebo test, to check for pre-trends. The $\beta_{1,m}$s, $\beta_{2,m}$s, and $\beta_{3,m}$s measure the impact of training, incentives, and the interaction respectively, for the 6 months leading up to the interventions (September 2015-February 2016), and the period after the interventions began (March 2016-February 2017). The β's are the average difference between treatment and control groups in month-year m, relative to the average difference in the reference period of January-August 2015.

In Equation 1.3, I pool together months based on experiment phases. I present experiment phase treatment effects as follows, where I_p indicate the four experiment phases March-June 2016, July-September 2016, October-December 2016 and January-February 2017. The reference period is January 2015-February 2016 and the rest is as in Equation 1.2.

$$\begin{aligned} Y_{imd} = & \sum_{p=1}^{p=4} \beta_{1,p}(Training_d \times I_p) + \sum_{p=1}^{p=4} \beta_{2,p}(Incentives_d \times I_p) \\ & + \sum_{p=1}^{p=4} \beta_{3,p}(Training_d \times Incentives_d \times I_p) + \alpha_d + \delta_m + Controls_{imd} + \epsilon_{imd} \end{aligned} \tag{1.3}$$

I control for vehicle, route, and day of week fixed effects. While the control variables significantly improve precision, it is possible that driver assignment to route and vehicle was affected by the interventions. Thus, for all equations, I present results including and excluding controls $Controls_{imd}$.

1.4 Results

1.4.1 Take-Up, Compliance, and Attrition

Random assignment to treatment is a strong and statistically significant predictor of take-up. In all treatment arms, there are high levels of overall compliance with assigned treatment, with especially high compliance in the control groups. I present a detailed breakdown of causes of non compliance for treatment groups T2 and T3. There is an overall attrition rate of 6% in the 26 month fuel efficiency panel dataset, balanced across the four treatment groups. I verify that the post-attrition subsample remains balanced on baseline survey characteristics and baseline fuel efficiency, and that the results on take-up and compliance in the post-attrition subsample are qualitatively unchanged.

Table 1.2: Take-up, Compliance, and Attrition

	Section I: Means					Section II: Differences		
	(1) C	(2) T1	(3) T2	(4) T3		(1) Training	(2) Incentives	(3) Tr. × Inc.
Panel A: Take-Up					.			
Took up training	0.00525	0.937	0.0131	0.955	.	0.931***	0.00947	0.0130
	(0.0724)	(0.244)	(0.114)	(0.207)		(0.0133)	(0.00721)	(0.0179)
Took up incentives	0	0	0.916	0.921	.	-0.000509	0.920***	0.00427
	(0)	(0)	(0.278)	(0.270)		(0.00430)	(0.0137)	(0.0193)
Panel B: Compliance					.			
Complied with training	0.995	0.937	0.987	0.955	.	-0.0576***	-0.00777	0.0295*
	(0.0724)	(0.244)	(0.114)	(0.207)		(0.0133)	(0.00721)	(0.0179)
Complied with incentives	1	1	0.916	0.921	.	-0.000509	-0.0796***	0.00427
	(0)	(0)	(0.278)	(0.270)		(0.00430)	(0.0137)	(0.0193)
Panel C: Attrition (Data)					.			
Missing 0 months	0.580	0.550	0.541	0.582	.	-0.0223	-0.0336	0.0630
	(0.494)	(0.498)	(0.499)	(0.494)		(0.0338)	(0.0334)	(0.0475)
Missing 3+ months	0.281	0.297	0.302	0.282	.	0.0111	0.0150	-0.0299
	(0.450)	(0.458)	(0.460)	(0.450)		(0.0307)	(0.0314)	(0.0437)
Missing all 26 months	0.0551	0.0395	0.0709	0.0500	.	-0.0186	0.0105	0.00228
	(0.229)	(0.195)	(0.257)	(0.218)		(0.0118)	(0.0135)	(0.0173)
.					.			
Number of missing months	3.840	3.953	4.520	3.911	.	-0.0131	0.534	-0.569
	(7.222)	(7.054)	(7.932)	(7.319)		(0.447)	(0.473)	(0.650)
.					.			
Attrition for analysis	0.0577	0.0447	0.0761	0.0579	.	-0.0147	0.0140	0.000211
	(0.234)	(0.207)	(0.266)	(0.234)		(0.0127)	(0.0143)	(0.0186)
Number of drivers	381	380	381	380	.			

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Section I: reported coefficients are means for groups C(Control), T1(Training Only), T2(Incentives Only), and T3(Training+Incentives). Standard deviations are in parentheses. Section II: reported coefficients are from regressions of each variable on Training, Incentives, Training × Incentives (Tr × Inc), and strata fixed effects. Robust standard errors are in parentheses. Compliance on training is calculated using attendance records from the RCT training sessions (all records available) and non-RCT training sessions for February-December 2016 (most but not all records available). I do not have attendance records for September 2015-January 2016 (pre-intervention) or 2017 (post-intervention). 4 of the 7 non-compliers in groups C and T2 were mistakenly trained in non-RCT training sessions. As attendance records are incomplete, it is possible that there was additional non-compliance in the non-RCT sessions. Compliance on the incentives intervention refers to partial or full compliance (i.e. complied for at least one of the three months). 'Attrition for analysis' indicates the driver is not included in the difference in difference specifications in Table 1.4. This group of 90 includes the 82 drivers with complete attrition and 8 drivers with insufficient observations for the empirical strategy. I use the *reghdfe* Stata estimator for multiple levels of fixed effects, in which singleton groups are dropped from the regression sample (Correia 2015, 2016). This table and others in this dissertation rely heavily on Ben Jann's *estout* Stata command (Jann 2005, 2007).

Using intervention implementation data, I find that for both interventions, random assignment to treatment is a strong and statistically significant predictor of take-up. The point estimates are 0.93 for training and 0.92 for incentives, significant at the 1% level. Panel A of Table 1.2 shows there was high take-up in the treatment groups, ranging between 92-96%. In contrast, less than 1 % of drivers (7 of 762) in the control groups took up training, and none took up the incentives intervention.

Using intervention implementation data, I also find high levels of overall compliance with assigned treatment arms, ranging from 91.6% to 100% in Panel B of Table 1.2. Compliance is defined as non-participation for the control groups, training attendance for the training groups, and being in compliance for at least 1 of the 3 eligible months for the incentives groups. There were statistically significant differences in compliance, which was 6-8% higher in the control groups (which had compliance rates close to 100%) than the treatment groups (which had compliance rates of 92-96% due to incomplete take-up). Compliance with the training intervention was somewhat higher (3%) in group T3 (Training + Incentives) than T1 (Training Only).

Appendix Table A.1 provides a detailed breakdown of the non-compliance and attrition for incentives groups T2 and T3. The non-compliance in these treatment groups is due to a mixture of attrition (resignation, transfers to other depots, suspensions, and so on) and 'true' non-compliance (scheduling conflicts when the training sessions were held, drivers on short-term disability leave, drivers assigned to do conductor duty, et cetera). Some of the non-compliance was due to implementation challenges, particularly drivers who achieved their target but did not receive their incentive.

Using KSRTC administrative data, I find an overall attrition rate of 6% in the 26 month fuel efficiency panel dataset, balanced across treatment groups, as shown in Panel C of Table 1.2. 'Attrition for analysis' indicates the driver has insufficient observations and is not in the subsample used for the main difference in difference specification in Table 1.4. Thus the post-attrition subsample for analysis consists of 1,432 drivers out of the initial 1,522, implying an overall attrition rate of 6%. Drivers with missing data for all 26 months are typically cases where baseline survey driver data could not be successfully matched to administrative records. Partial attrition for some of the 26 months was primarily due to resignations, transfers, conductor duty, and so on.

Table 1.3: Post-Attrition Subsample: Balance on Baseline Fuel Efficiency

	Section I: Means					Section II: Differences		
	(1) C	(2) T1	(3) T2	(4) T3		(1) Training	(2) Incentives	(3) Tr. × Inc.
KMPL Average for:					.			
Jan2015-Feb2016	4.651	4.685	4.671	4.666	.	0.0246	0.00783	-0.0322
	(0.578)	(0.535)	(0.537)	(0.547)		(0.0194)	(0.0197)	(0.0275)
	[357]	[360]	[351]	[356]				
September 2015	4.670	4.679	4.690	4.669	.	0.0126	-0.00407	-0.0261
	(0.597)	(0.570)	(0.544)	(0.573)		(0.0231)	(0.0219)	(0.0321)
	[348]	[346]	[332]	[343]				
October 2015	4.654	4.678	4.676	4.692	.	0.0210	0.0145	-0.0148
	(0.586)	(0.574)	(0.563)	(0.590)		(0.0219)	(0.0219)	(0.0319)
	[343]	[351]	[335]	[340]				
November 2015	4.661	4.674	4.671	4.677	.	0.0166	0.00812	-0.00724
	(0.589)	(0.554)	(0.537)	(0.570)		(0.0226)	(0.0232)	(0.0326)
	[341]	[342]	[328]	[340]				
December 2015	4.649	4.672	4.679	4.682	.	0.0148	0.0243	-0.0129
	(0.601)	(0.586)	(0.557)	(0.577)		(0.0238)	(0.0237)	(0.0335)
	[337]	[329]	[324]	[329]				
January 2016	4.707	4.715	4.689	4.720	.	0.00205	-0.00796	0.00902
	(0.593)	(0.605)	(0.547)	(0.600)		(0.0261)	(0.0246)	(0.0365)
	[337]	[333]	[321]	[335]				
February 2016	4.722	4.745	4.744	4.769	.	-0.000883	0.000602	0.0287
	(0.612)	(0.583)	(0.542)	(0.588)		(0.0235)	(0.0229)	(0.0334)
	[330]	[328]	[320]	[326]				
.					.			
Number of Drivers	359	363	352	358	.			
Joint F-Test Statistic	.	0.705	1.943	1.688	.			
P-Value	.	0.668	0.0607	0.109	.			

This table presents balance checks on baseline fuel efficiency for the post-attrition subsample used for analysis. Kilometers per liter (KMPL) averages presented are a simple average of KMPL in all shifts in the specified time period. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Section I: reported coefficients are means for groups C (Control), T1 (Training Only), T2 (Incentives Only), and T3 (Training+Incentives). Standard deviations are in parentheses. Observation counts are in square brackets. Section II: reported coefficients are from regressions of each variable on Training, Incentives, Training × Incentives (Tr × Inc), and strata fixed effects. Robust standard errors are in parentheses. The joint hypothesis tests are from separate regressions of each treatment group compared to the control group. The treatment indicator is regressed on all KMPL variables. I test the joint hypothesis that all coefficients are equal to zero.

I verify that the post-attrition subsample remains balanced on baseline survey characteristics and baseline fuel efficiency. Also, the results on take-up and compliance in the post-attrition subsample are qualitatively unchanged. In Table 1.3 I verify there is balance on baseline fuel efficiency (KMPL) variables in the post-attrition subsample. Appendix Table A.2 shows that the post-attrition subsample is balanced on baseline survey characteristics. Appendix Table A.3 demonstrates that take-up and compliance was qualitatively similar in the post-attrition subsample and the complete sample.

1.4.2 Twelve Month Treatment Effects of Training and Incentives on Fuel Efficiency

Over the 12 months of post-intervention data, I find the training program had no overall statistically significant effect on fuel efficiency. The incentives scheme did have a positive impact in the entire period. I find no evidence of an interaction effect.

Table 1.4 presents the results from Equation 1.1 for the entire 12 month post-intervention period March 2016-February 2017. Column 1 is the specification without controls, and Column 4 has full controls with depot $\times$ route, depot $\times$ vehicle, and day of week fixed effects.

I find that the training program had no overall statistically significant effect on fuel efficiency. The treatment effect for training is 0.00787 kilometers per liter (KMPL) in Column 1 and 0.00906 KMPL in Column 4. Though positive, these estimates are small and statistically insignificant.

In contrast, the incentives scheme did have a positive impact in the entire 12 month period. In Column 1, the estimated effect is 0.0190 KMPL and is statistically insignificant. Adding controls significantly improves precision. In Column 4, the point estimate of 0.0168 KMPL is significant at the 5% level. As discussed in Section 1.3.4, I cannot rule out the possibility that driver assignment to routes and vehicles was affected by the interventions, so it is possible the controls were affected by the treatment. However, the estimates from columns 1 and 4 are very similar.

There appears to be no overall interaction effect of the two interventions. The point estimates of 0.0176 KMPL in Column 1 and 0.001 KMPL in Column 4 are statistically insignificant. For the interaction effect alone, the point estimates differ substantially when controls are added. The estimates for the interaction effect are also much more imprecise than those for the main effects.

Table 1.4: Impact on Fuel Efficiency in the Twelve Month Post-Intervention Period

	(1) KMPL	(2) KMPL	(3) KMPL	(4) KMPL
Training × Post	0.00787	0.0192	0.00890	0.00906
	(0.0189)	(0.0124)	(0.0102)	(0.0102)
Incentives × Post	0.0190	0.0386***	0.0169**	0.0168**
	(0.0163)	(0.0105)	(0.00828)	(0.00829)
Training × Incentives × Post	0.0176	-0.0113	0.000771	0.000948
	(0.0255)	(0.0169)	(0.0135)	(0.0136)
Observations (shifts)	384241	381612	381375	381375
Number of clusters (drivers)	1432	1432	1432	1432
Driver fixed effects	Yes	Yes	Yes	Yes
Month × year fixed effects	Yes	Yes	Yes	Yes
Depot × route fixed effects		Yes	Yes	Yes
Depot × vehicle fixed effects			Yes	Yes
Day of week fixed effects				Yes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ Robust standard errors clustered by driver in parentheses. KMPL is winsorized at 1% and 99%. Post is an indicator for all shifts from March 1 2016 onwards.

1.4.3 Time Patterns of the Treatment Effects

Next, I examine the time pattern of the treatment effects. I analyze the monthly and experiment phase treatment effects for each intervention. I confirm that there are no treatment effects or pre-trends before the intervention begins. The training program had a positive effect in the four months of implementation and no effect thereafter. The incentive program had a positive effect during implementation which persists for some time past implementation and then gradually fades.

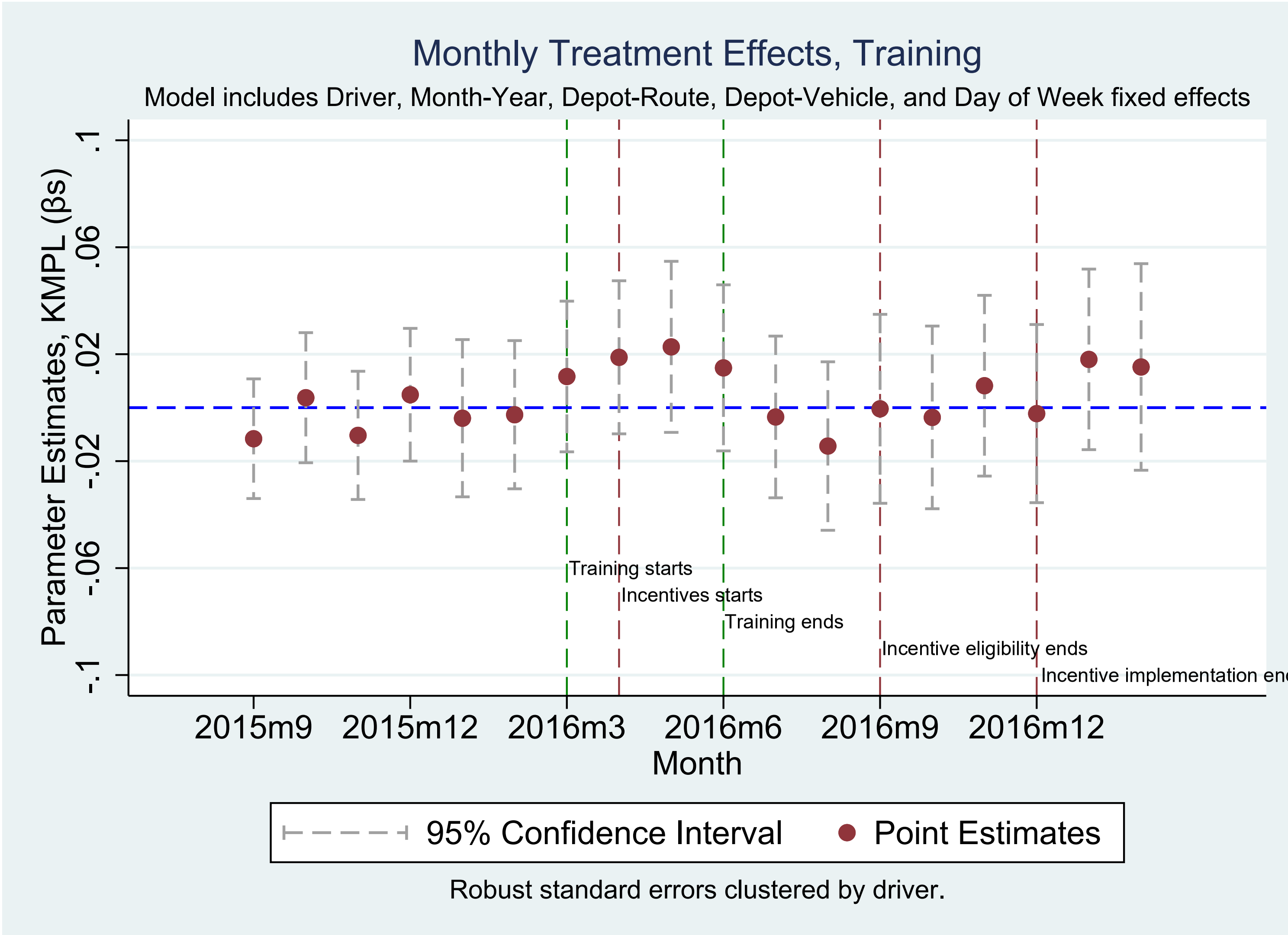

Figure 1.2: Monthly Treatment Effects, Training ($\beta_{1,m}s$)

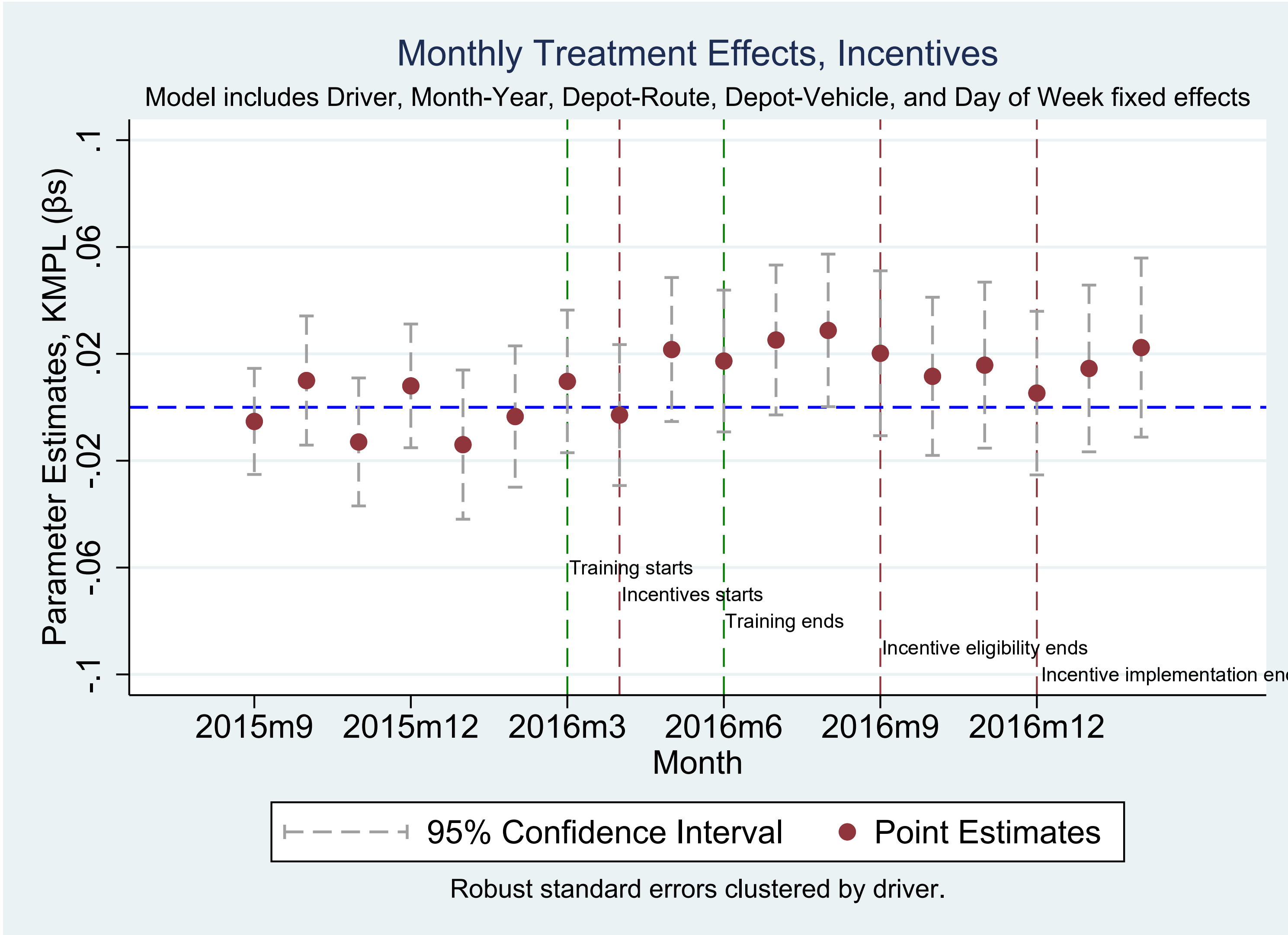

Figure 1.3: Monthly Treatment Effects, Incentives ($\beta_{2,m}s$)

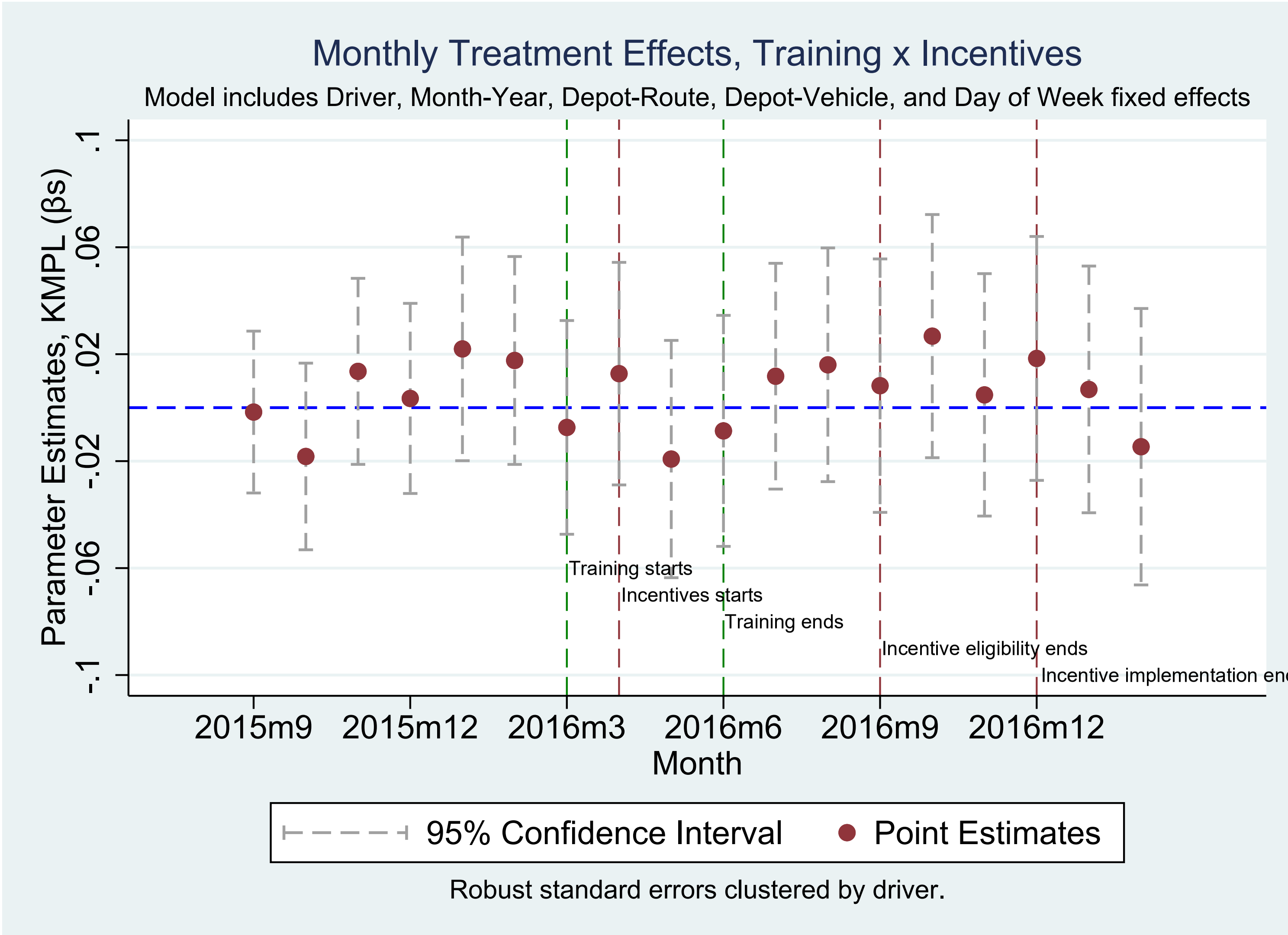

Figure 1.4: Monthly Treatment Effects, Training × Incentives ($\beta_{3,m}s$)

Monthly treatment effects:

First, I analyze the monthly treatment effects based on Equation 1.2. I confirm that there are no treatment effects or pre-trends before the intervention begins. The monthly treatment effects for training, incentives, and the interaction are graphed in Figures 1.2, 1.3, and 1.4 respectively, using the specification with full controls. Vertical lines in the figures mark the four experiment phases[15]. The full numerical results are presented in Appendix Table A.5.

The training program had a positive effect in the four months of implementation and no effect thereafter. Figure 1.2 plots time patterns for training. Pre-training, before the first vertical line at March 2016, there was no treatment effect. During the training implementation, between the first vertical line and the third vertical line at June 2016, there was a noticeable positive impact. This treatment effect quickly faded post implementation, for all months after the third vertical line. The pattern is similar in the specification without controls, Panel A of Appendix Figure A.3, though the estimates are significantly noisier and more volatile.

The incentive program had a positive effect during implementation which persisted for some time past implementation and then gradually faded. Figure 1.3 plots time patterns for the incentives intervention. April 2016, at the second vertical line, is the first month drivers were eligible for incentives. Incentives disbursal started in the following month, May 2016. From May 2016 until the incentive eligibility ended in September 2016, at the fourth vertical line, there was a noticeable positive treatment effect. The incentives effect persisted for some time post-intervention and gradually faded. The monthly treatment effects show broadly similar patterns with and without controls in Appendix Figure A.4.

There appears to be no statistically significant interaction effect in any phase of the experiment. In the specification without controls, the interaction is occasionally positive at fairly large magnitudes, but it is always within the confidence interval.

Experiment phase treatment effects:

Next, I estimate the magnitude of the treatment effects for each experiment phase. Equation 1.3 studies treatment effects for 4 different experiment phases. Complete results are presented in

15. Appendix Figures A.3, A.4, and A.5 reproduce these figures in Panel B, and compare them to the coefficients from the specification with no controls in Panel A.

Table 1.5. Phase 1 is March-June 2016, which was the implementation of the training program and the first three months of incentive eligibility. Phase 2 is July-September 2016, the final 3 months of incentive eligibility. Phase 3 is October-December 2016, when incentive eligibility was complete and implementation was still ongoing. Phase 4 is January-February 2017, after all intervention implementation activities were completed. Details on the experiment phases and timeline are described in Section 1.3.2 and graphed in Appendix Figure A.2.

During the Phase 1 training intervention implementation, from March-June 2016, the training treatment effect point estimates in Table 1.5 are 0.0190 KMPL in Column 1 without controls and 0.0186 KMPL in Column 4 with controls. With controls, the estimate is marginally significant (p-value of 0.076). For the remaining post-intervention Phases 2, 3, and 4, there is no training treatment effect.

I turn next to the incentives intervention treatment effects in Table 1.5. The incentives treatment effect peaked in Phase 2 in July-September 2016, during the final three months of incentive eligibility, when the point estimates are 0.031 KMPL in Column 1 without controls and 0.026 KMPL in Column 4 with controls. In both specifications, the incentives effect remained positive at a smaller magnitude in Phases 3 and 4 after the drivers were no longer eligible. In the specification without controls, the magnitude drops in Phase 3 and declines further in Phase 4, while the specification with controls suggests that the post-eligibility treatment effect is fairly constant in Phases 3 and 4.

The results overall suggest that the treatment effect persisted post eligibility as long as incentive implementation continued, but faded once implementation was complete. I implement a complementary event-time approach in Appendix A.2 in which this pattern is very noticeable. In the main approach in Table 1.5, the peak effect was in July-September 2016, when 69% of the observations are from the post-eligibility implementation phase[16].

16. For the 43% drivers in Batch 1, incentive eligibility was completed in June 2016, so all observations are from the post-eligibility implementation phase, and for the 39% of drivers in Batch 2, the final month of eligibility was July, so 2/3 of the observations are from the post-eligibility implementation phase.

Table 1.5: Impact on Fuel Efficiency in Four Experiment Phases

	(1) KMPL	(2) KMPL	(3) KMPL	(4) KMPL
Training $\times 1(Phase1: Training)$	0.0190	0.0268**	0.0184*	0.0186*
	(0.0191)	(0.0135)	(0.0105)	(0.0105)
Training $\times 1(Phase2: IncentiveEligibility)$	-0.0157	0.00983	-0.00500	-0.00488
	(0.0225)	(0.0156)	(0.0134)	(0.0134)
Training $\times 1(Phase3: IncentiveImplementation)$	0.0151	0.0152	0.00185	0.00207
	(0.0256)	(0.0171)	(0.0149)	(0.0149)
Training $\times 1(Phase4: Post - Implementation)$	0.00838	0.0209	0.0182	0.0184
	(0.0264)	(0.0191)	(0.0166)	(0.0166)
Incentives $\times 1(Phase1: Training)$	0.0139	0.0269**	0.0132	0.0130
	(0.0159)	(0.0112)	(0.00903)	(0.00904)
Incentives $\times 1(Phase2: IncentiveEligibility)$	0.0305	0.0531***	0.0264**	0.0264**
	(0.0213)	(0.0148)	(0.0119)	(0.0119)
Incentives $\times 1(Phase3: IncentiveImplementation)$	0.0196	0.0387**	0.0126	0.0124
	(0.0218)	(0.0161)	(0.0127)	(0.0127)
Incentives $\times 1(Phase4: Post - Implementation)$	0.0112	0.0488***	0.0195	0.0193
	(0.0242)	(0.0188)	(0.0141)	(0.0142)
Training $\times$ Incentives $\times 1(Phase1: Training)$	0.0174	-0.0111	-0.00945	-0.00933
	(0.0254)	(0.0185)	(0.0148)	(0.0149)
Training $\times$ Incentives $\times 1(Phase2: IncentiveEligibility)$	0.0336	-0.00813	0.00879	0.00875
	(0.0319)	(0.0219)	(0.0181)	(0.0181)
Training $\times$ Incentives $\times 1(Phase3: IncentiveImplementation)$	0.00758	-0.00743	0.0134	0.0137
	(0.0344)	(0.0234)	(0.0192)	(0.0192)
Training $\times$ Incentives $\times 1(Phase4: Post - Implementation)$	0.00715	-0.0268	-0.00516	-0.00465
	(0.0374)	(0.0268)	(0.0216)	(0.0216)
Observations (shifts)	384241	381612	381375	381375
Number of clusters (drivers)	1432	1432	1432	1432
Driver fixed effects	Yes	Yes	Yes	Yes
Month $\times$ year fixed effects	Yes	Yes	Yes	Yes
Depot $\times$ route fixed effects		Yes	Yes	Yes
Depot $\times$ vehicle fixed effects			Yes	Yes
Day of week fixed effects				Yes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered by driver in parentheses. KMPL is winsorized at 1% and 99%. Reported coefficients are interactions between the treatments and indicators for experiment phases p. The reference group is all shifts before March 1 2016. Phase 1 is March-June 2016, Phase 2 is July-September 2016, Phase 3 is October-December 2016, Phase 4 is January-February 2017. Details on the Phases are in Section 1.4.3.

1.4.4 Heterogeneity of the Treatment Effects

I analyze the heterogeneity of treatment effects based on the conceptual framework in Section 1.2.5. I investigate the relative importance of four channels through which the interventions could affect driver performance: ability, effort, salience, and habit. In contradiction to the conceptual framework prediction on the ability channel, I find that drivers who had previously been trained at KSRTC show larger gains from the repeat training than drivers who were trained for the first time. In line with the conceptual framework prediction on the effort channel, I find that drivers with an above-median baseline KMPL who needed to exert less effort to achieve the target responded significantly more to the incentives intervention. In Chapter 2 of this dissertation, I further explore the heterogeneity of treatment effects.

To investigate heterogeneity, for the qualitative variable of time preference, I conduct a subgroup analysis of Equation 1.1. For all other variables, I adapt Equation 1.1 and add an interaction with the driver characteristic as follows

$$\begin{aligned}
Y_{imd} &= \beta_1 Training_d \times Post_i + \beta_2 Incentives_d \times Post_i \\
&\quad + \beta_3 Training_d \times Incentives_d \times Post_i + \beta_4 Training_d \times Post_i \times Char_d \\
&\quad + \beta_5 Incentives_d \times Post_i \times Char_d + \beta_6 Training_d \times Incentives_d \times Post_i \times Char_d \\
&\quad + \beta_7 Post_i \times Char_d + \alpha_d + \delta_m + Controls_{imd} + \epsilon_{imd}
\end{aligned} \tag{1.4}$$

First, to evaluate the importance of the ability channel, I test the predictions that drivers with more previous training - either at KSRTC or at driving school- would show less gains from the SAFED program, and that drivers with more accidents - a proxy for lower baseline ability - would show higher gains from the SAFED program. Specifically, as detailed in Section 1.2.5, I test whether the training program had heterogeneous effects on drivers who were formally taught driving, drivers who had previously attended KSRTC training, drivers who had never had a major accident, and drivers who had never had a minor accident. Table 1.6 presents the results.

Table 1.6: Heterogeneity Analysis: Importance of the Ability Channel

	Attended KSRTC training		Attended driving school		Had major accident		Had minor accident	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post × Char	-0.0117	-0.0187	0.0209	-0.0688**	0.000716	-0.00426	0.00729	-0.00986
	(0.0352)	(0.0178)	(0.0671)	(0.0344)	(0.0342)	(0.0152)	(0.0238)	(0.0122)
Training × Post	-0.0538	-0.0352	0.00938	0.00519	0.00514	0.00831	0.0149	0.0101
	(0.0441)	(0.0252)	(0.0192)	(0.0103)	(0.0207)	(0.0111)	(0.0252)	(0.0126)
Incentives × Post	0.0799*	0.0359	0.0215	0.0145*	0.0167	0.00879	0.0219	0.0134
	(0.0463)	(0.0278)	(0.0166)	(0.00840)	(0.0179)	(0.00897)	(0.0220)	(0.0109)
Training × Incentives × Post	-0.0175	-0.00824	0.0136	0.00537	0.0252	0.00618	0.0284	0.00269
	(0.0670)	(0.0368)	(0.0261)	(0.0138)	(0.0281)	(0.0142)	(0.0336)	(0.0177)
Training × Post × Char	0.0697	0.0507*	-0.0609	0.149***	0.0210	0.00784	-0.0129	0.000247
	(0.0486)	(0.0275)	(0.115)	(0.0451)	(0.0511)	(0.0279)	(0.0376)	(0.0202)
Incentives × Post × Char	-0.0676	-0.0208	-0.0949	0.0609	0.0124	0.0413*	-0.00652	0.00750
	(0.0494)	(0.0293)	(0.0798)	(0.0381)	(0.0440)	(0.0219)	(0.0327)	(0.0169)
Training × Incentives × Post × Char	0.0390	0.00964	0.150	-0.154***	-0.0451	-0.0290	-0.0222	-0.00515
	(0.0724)	(0.0396)	(0.127)	(0.0552)	(0.0676)	(0.0402)	(0.0508)	(0.0271)
Observations (shifts)	384241	381375	384241	381375	383767	380908	383767	380908
Number of clusters (drivers)	1432	1432	1432	1432	1430	1430	1430	1430
Controls		Yes		Yes		Yes		Yes
.								
Characteristic Mean	0.859	0.859	0.0318	0.0318	0.174	0.174	0.485	0.485

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Parentheses contain robust standard errors clustered by driver. KMPL is winsorized at 1% and 99%. Post indicates shifts after March 1 2016. Char is the driver characteristic in the column title. All specifications include driver and month × year fixed effects. The control variables in even-numbered columns are depot × vehicle, depot × route, and day of week fixed effects. Characteristic mean is the mean of the characteristic in the post-attrition driver subsample used for analysis. Attended KSRTC training indicates the driver had previously attended KSRTC training at the time of the baseline survey. The variables on driving school and accident history are explained in detail in Section 1.2.5.

Overall, I find no heterogeneous effects of training based on whether a driver had attended driving school or based on accident history, and the results indicate that in contradiction to the prediction, drivers who had previously attended KSRTC training actually had larger gains from SAFED training - a difference of 0.07 KMPL (insignificant) without controls and a difference of 0.051 KMPL (marginally significant) with controls. Surprisingly, it seems that repeated/refresher training was more effective than an initial training. It is possible that multiple trainings are needed to change long-held driving habits. Another explanation could be that KSRTC management usually successfully identifies a minority of drivers who are unlikely to benefit from training and does not train them. Typically, KSRTC management selects drivers for training, whereas the RCT randomly trained half of all recruited drivers. For instance, in the post-attrition subsample, 47.84 % of previously-trained drivers had above-median baseline KMPL, while 57.87 % of never-trained drivers were above median. This implies that the 'never been trained at KSRTC drivers' may be a high ability subgroup who thus show minimal gains from SAFED training. There are no differential effects of training based on history of major or minor accidents. The results on the effect of having attended driving school are erratic and inconsistent - possibly due to the small numbers of this subsample, as only 3% of drivers in the post-attrition subsample have formal training before joining KSRTC, and the rest typically learned to drive from a mentor, friends, or family.

Second, to test the importance of the effort channel, I use the predictions based on the conceptual framework in Section 1.2.5. I investigate whether the incentives intervention had a larger effect on drivers who had less personal costs or greater personal benefits from achieving the target - namely, drivers who had a high baseline KMPL and could hence achieve the target with minimal effort, drivers who were patient and hence potentially bore less costs from driving slower, and drivers who cheated on a dishonesty test for financial reward and hence may place a higher value on the financial incentives. As described in Section 1.2.5, I use a measure of drivers whose baseline KMPL was above the depot-specific median. Since targets were at the depot-level, drivers who were already relatively high performing would have to invest less effort to achieve the target. For the behavioral measures of dishonesty and time preferences, I describe the measures and instructions briefly in Section 1.2.5, and in detail in Chapter 2 of this dissertation. Table 1.7 presents the results of the subgroup analysis for drivers with different time preferences, and Table 1.8 presents the results of the heterogeneity analysis.

I find no heterogeneous effects of the incentives intervention based on drivers' patience or dishonesty, but I do consistently find that the incentives intervention had a much larger impact on drivers with above-median baseline KMPL. The difference in the treatment effect is 0.074 KMPL without controls and 0.038 KMPL with controls, both of which are statistically significant. In fact, among the below-median drivers, there was no impact of the incentives intervention at all - the entire overall incentives effect was due to the above-median relatively high performing drivers. Below-median drivers may have anticipated the target was out of reach and thus did not respond to the incentives scheme. In line with the conceptual framework prediction on effort, above-median drivers who could achieve the target with less effort responded relatively strongly to the incentives scheme.

Table 1.7: Subgroup Analysis by Time Preference

	(1) All	(2) CoPa	(3) Hype	(4) CoIm	(5) PNIL	(6) Miss
Training × Post	0.00787	0.00766	0.0335	0.00761	-0.0688	0.0505
	(0.0189)	(0.0245)	(0.0333)	(0.0702)	(0.0657)	(0.281)
Incentives × Post	0.0190	0.0342	0.0297	-0.0304	-0.0204	-0.186
	(0.0163)	(0.0220)	(0.0351)	(0.0450)	(0.0468)	(0.163)
Tr. × Inc. × Post	0.0176	0.0221	-0.0388	0.0625	0.0535	0.167
	(0.0255)	(0.0339)	(0.0500)	(0.0812)	(0.0835)	(0.319)
Observations (shifts)	384241	220476	84318	48015	28319	3113
Clusters (drivers)	1432	824	316	174	108	10
Time Preference Category		CoPa	Hype	CoIm	PNIL	Miss

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Parentheses contain robust standard errors clustered by driver. KMPL is winsorized at 1% and 99%. Post is an indicator for all shifts from March 1 2016 onwards. Tr. × Inc. is an abbreviation for Training × Incentives. Columns 1-6 present results for all drivers followed by drivers in each time preference category CoPa, Hype, CoIm, PNIL, and Miss. CoPa is Consistently Patient, Hype is Hyperbolic, CoIm is Consistently Impatient, PNIL is Patient Now, Impatient Later, and Miss indicates Missing data. Further details on the categories are in Table 2.2. All specifications include driver and month × year fixed effects. The model with controls is not presented as the fixed effect controls cannot be appropriately estimated using driver subgroups.

Table 1.8: Heterogeneity Analysis: Importance of the Effort Channel

	Baseline KMPL > depot median		Consistently patient		Dice > 99p	
	(1)	(2)	(3)	(4)	(5)	(6)
Post × Char	-0.107***	-0.0356***	-0.00690	0.00568	0.0281	0.0244
	(0.0232)	(0.0123)	(0.0241)	(0.0123)	(0.0416)	(0.0245)
Training × Post	0.0345	0.0102	0.00846	0.0347**	0.0181	0.0141
	(0.0245)	(0.0141)	(0.0299)	(0.0162)	(0.0197)	(0.0106)
Incentives × Post	-0.0131	-0.000506	0.00186	0.0127	0.0214	0.0208**
	(0.0219)	(0.0110)	(0.0241)	(0.0130)	(0.0170)	(0.00862)
Training × Incentives × Post	0.0207	0.0260	0.00923	-0.0328	0.0117	-0.00717
	(0.0337)	(0.0184)	(0.0388)	(0.0213)	(0.0266)	(0.0142)
Training × Post × Char	-0.0444	0.000893	-0.000817	-0.0430**	-0.0958	-0.0461
	(0.0365)	(0.0205)	(0.0386)	(0.0208)	(0.0695)	(0.0401)
Incentives × Post × Char	0.0740**	0.0384**	0.0324	0.00413	-0.0268	-0.0370
	(0.0320)	(0.0168)	(0.0326)	(0.0169)	(0.0617)	(0.0323)
Training × Incentives × Post × Char	-0.00931	-0.0504*	0.0127	0.0595**	0.106	0.0913*
	(0.0498)	(0.0271)	(0.0515)	(0.0277)	(0.0941)	(0.0497)
Observations (shifts)	384099	381232	381128	378270	381437	378578
Number of clusters (drivers)	1424	1424	1422	1422	1423	1423
Controls		Yes		Yes		Yes
.						
Characteristic Mean	0.492	0.492	0.579	0.579	0.0878	0.0878

* * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Parentheses contain robust standard errors clustered by driver. KMPL is winsorized at 1% and 99%. Post indicates shifts after March 1 2016. Char is the driver characteristic specified in the column title. The characteristic mean is in the post-attrition driver subsample for analysis. All specifications include driver and month × year fixed effects. Controls in even-numbered columns are depot × vehicle, depot × route, and day of week fixed effects. Consistently patient=1 if time preferences were CoPa and =0 if time preferences were Hype, CoIm, or PNIL (see Table 2.2 for details.) Dice points > 99p indicates total dice points was above the 99th percentile of the theoretical distribution.

Table 1.9: Heterogeneity Analysis: Importance of the Salience and Habit Channels

	Age (years)		Risk aversion (score)		Max. donated	
	(1)	(2)	(3)	(4)	(5)	(6)
Post × Char	-0.00341**	-0.00205**	0.01000	0.00657	0.0210	0.0185
	(0.00164)	(0.000963)	(0.00852)	(0.00421)	(0.0238)	(0.0123)
Training × Post	0.0369	-0.0756	0.0601	0.0324	0.0141	0.00765
	(0.0920)	(0.0512)	(0.0426)	(0.0250)	(0.0296)	(0.0148)
Incentives × Post	-0.0732	-0.00831	0.0633*	0.0290	0.0283	0.0184
	(0.0820)	(0.0434)	(0.0381)	(0.0188)	(0.0250)	(0.0128)
Training × Incentives × Post	-0.0237	0.00148	-0.0333	0.0213	-0.00670	0.00369
	(0.125)	(0.0689)	(0.0588)	(0.0336)	(0.0412)	(0.0205)
Training × Post × Char	-0.000703	0.00218	-0.0170	-0.00769	-0.0100	0.00227
	(0.00233)	(0.00139)	(0.0132)	(0.00691)	(0.0386)	(0.0200)
Incentives × Post × Char	0.00238	0.000652	-0.0146	-0.00403	-0.0162	-0.00368
	(0.00207)	(0.00116)	(0.0112)	(0.00568)	(0.0330)	(0.0170)
Training × Incentives × Post × Char	0.00101	-0.0000151	0.0168	-0.00636	0.0408	-0.00372
	(0.00319)	(0.00186)	(0.0176)	(0.00939)	(0.0525)	(0.0272)
Observations (shifts)	384241	381375	384170	381303	383965	381100
Number of clusters (drivers)	1432	1432	1431	1431	1431	1431
Controls		Yes		Yes		Yes
.						
Characteristic Mean	38.61	38.61	3.055	3.055	0.604	0.604

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Parentheses contain robust standard errors clustered by driver. KMPL is winsorized at 1% and 99%. Post indicates shifts after March 1 2016. Char is the driver characteristic specified in the column title. The characteristic mean is in the post-attrition driver subsample for analysis. All specifications include driver and month × year fixed effects. Controls in even-numbered columns are depot × vehicle, depot × route, and day of week fixed effects. Max. donated indicates Rs.30, the maximum amount, was donated in a pro-socialness game. Risk aversion scores range from 1 (riskiest lottery selected) to 5 (safest lottery selected).

Third, I test the importance of the salience and habit channels based on the conceptual framework in Section 1.2.5, and find no heterogeneous effects based on age, risk aversion, or pro-socialness, as seen in Table 1.9. Based on the conceptual framework, I predicted that if habit is important, either intervention may have a larger effect on younger drivers who might find it easier to change habits, and if salience is importance, treatment effects of either intervention might be larger for risk averse and pro social drivers who may be intrinsically motivated to drive well and could be easily activated by salient cues. I use behavioral measures of risk aversion and pro-socialness, which I describe briefly in Section 1.2.5, and in detail in Chapter 2 of this dissertation.

1.4.5 Comparing the Conceptual Framework Predictions to the Results

In Section 1.2.5, I proposed a conceptual framework that the training and incentives interventions could change driver performance through four main channels: ability, effort, salience, and habits. In this section, I compare the theoretical predictions from the conceptual framework to the results for the overall intervention treatment effects, the time patterns of the treatment effects, and the heterogeneity of treatment effects, reported in Sections 1.4.2, 1.4.3, and 1.4.4 respectively. I find that the effort and salience channels are important, while the ability and habit channels are not.

First, the ability channel. Overall, the results suggest that variation in driving ability, or lack of ability, is not a main factor in driver under-performance. I find the training program had no overall statistically significant effect on fuel efficiency in the entire twelve month period. Turning to the time patterns, though I find a short run effect of training, it was not sustained, which implies the training program's effects may be through salience rather than changing ability. Finally, looking at the heterogeneous effects of training, I do not find that training had a larger effect on drivers with less training or less ability at baseline. Indeed, I instead find training was more effective as a repeat training for drivers who had been trained at KSRTC before. Taking these results together, and combining them with the conceptual framework predictions, I find that the impact of the interventions is not primarily through the ability channel.

Second, the effort channel. Overall, the results suggest that lack of effort or variation in effort is a key factor in driver performance. I find the financial incentives scheme had an overall statistically significant effect on fuel efficiency in the entire twelve month period. Turning to the time patterns, I find the incentives program had a positive effect while drivers were eligible for financial incentives which also persisted for some time post eligibility before fading. This time pattern suggests that while effort is important, salience matters as well, as discussed below. Finally, looking at the heterogeneous impact of incentives, I find that the financial incentives scheme had a much larger impact on drivers whose baseline KMPL was above-median, i.e. on high performing drivers who could achieve the target with only a little incremental effort, though I find no heterogeneity due to driver patience or dishonesty. Taking these results together with the conceptual framework, I find that the impact of the interventions works in part through the effort channel.

Third, the salience channel. Overall, the results suggest that the interventions worked in large part by increasing the salience of fuel efficiency. The overall results show that the incentives intervention, which had more embedded cues than the training intervention, had a larger effect - which could be due to either effort or salience. Looking at the time patterns, I find an immediate short run effect of training which rapidly faded away. This strongly supports the hypothesis that the training intervention worked because it increased the salience of fuel efficiency for a short period. For the incentives intervention, the treatment effect persisted post eligibility as long as drivers were in regular contact with the field staff. That is, even after the three month eligibility period ended, the incentives intervention continued to have a significant effect as long as drivers got regular salient cues about fuel efficiency. The incentives treatment effect faded once all implementation activities were complete - this pattern is very noticeable in the complementary event-time approach in Appendix A.2. Thus the time patterns of the incentives effect again give strong credence to the hypothesis that the incentives intervention worked at least in part by increasing salience[17]. Turning to the predictions on heterogeneous effects, I find

17. However, it is also possible that drivers were confused about how long they were eligible for incentives. They were told it was three months long during the baseline survey, and again informed via phone of their eligibility at the beginning of the three month period, but drivers may not have recalled the duration. Additionally, while collecting receipts, field staff gave reminders of how many more months drivers were eligible. Due to the slow disbursement of incentives, drivers may have received this reminder after eligibility ended, and thus mistakenly believed they remained eligible.

no evidence that the interventions were more effective for risk averse or pro-social drivers.

Fourth, the habit channel. Overall, the results suggest the habit channel is not a key channel through which the interventions affect driver behavior. The prediction from the conceptual framework was that if the habit channel is important, there should be a noticeable overall interaction effect, but I find no evidence of any complementarities between training and incentives. Looking at time patterns, I do not find any persistence of the short run effects of training, indicating training did not change habits. I do find some persistence of the incentives treatment effect as long as implementation activities are ongoing, but this starts to taper off post-implementation, particularly in the complementary event-time approach in Appendix A.2, which suggests the post-eligibility persistence is due to salience rather than changes in habits. Turning to heterogeneous effects, I find no evidence to support the prediction that younger drivers might find it easier to change driving habits.

1.4.6 Robustness Checks

The results are robust to alternative specifications and methodologies.

Along with the difference in difference specification in Equation 1.1, I also estimate an analogous model using post-intervention data only, while controlling for drivers' baseline KMPL, and find similar results. I estimate

$$\begin{aligned} Y_{imd} &= \beta_1 Training_d + \beta_2 Incentives_d + \beta_3 Training_d \times Incentives_d \\ &\quad + \beta_4 BaselineKMPL_d + \gamma_s + \delta_m + Controls_{imd} + \epsilon_{imd}, \forall Y_{imd} \in [m >= Mar2016] \end{aligned} \tag{1.5}$$

Where $BaselineKMPL_d$ is the average KMPL for all pre-intervention shifts before March 1 2016, γ_s controls for strata fixed effects, and the rest is as in Equation 1.1. Appendix Table A.6 presents the results. Equations 1.1 and 1.5 yield similar results in the specifications without controls (Column 1 of Table 1.4 and Appendix Table A.6). Results differ significantly in the models with controls, as Equation 1.1 estimates vehicle, route, and other fixed effects using the full panel dataset, while Equation 1.5 estimates vehicle and other fixed effects using post-intervention data only.

Similarly, as an analogue to Equation 1.2, in Equation 1.6 I separately estimate treatment effects for each month-year m in September 2015-February 2017 and find similar time patterns:

$$\begin{aligned} Y_{imd} = &\beta_{1,m} Training_d + \beta_{2,m} Incentives_d + \beta_{3,m} Training_d \times Incentives_d \\ &+ \beta_4 BaselineKMPL_d + \gamma_s + \epsilon_{imd} \end{aligned} \tag{1.6}$$

Where $BaselineKMPL_d$ is the average KMPL for shifts from January-August 2015, and the rest is as in Equation 1.5. Each $\beta_{1,m}$, $\beta_{2,m}$, and $\beta_{3,m}$ comes from separate regressions for each month-year m. Appendix Figure A.6 plots the monthly treatment effects from the difference-in-difference specification without controls (Column 1 of Appendix Table A.5), along with the separate estimates for each month from Equation 1.6. The results from both approaches are very consistent.

Results remain unchanged if I drop shifts with multiple drivers. As described in Section 1.3.3, I created a 26 month panel dataset with uniform start dates using KSRTC administrative data for all 385,310 unique shifts between January 1 2015 to February 20 2017. Of these shifts, 382,895 had only one driver participating in the randomized controlled trial (RCT), and 2,415 had two RCT drivers. The panel dataset is created at the RCT driver-shift level, so the 2,415 shifts with two RCT drivers appear twice in the panel dataset, for a total of 387,725 observations. Thus the dataset contains 385,310 unique shifts and 2,415 duplicates. Appendix Table A.7 demonstrates that the results are similar if all shifts with multiple RCT drivers are dropped. As an additional check, I also drop any shift with multiple drivers (including drivers who were not in the RCT). Results remain similar.

1.5 Cost-Effectiveness of the Interventions

Both interventions were cost-effective. A social cost benefit analysis might also show significant net welfare benefits due to safety co-benefits, reduced emissions, and other externalities. During the pilot, I estimate the cost-effectiveness, i.e. return per rupee spent, was Rs. 3.12 for training and Rs. 4.22 for incentives. Estimated fuel savings was about 0.19% and 0.35% respectively of baseline fuel consumption, implying total savings of \$13,230 and \$27,410 respectively. At scale, I estimate the cost-effectiveness of training would be Rs. 2.04 per rupee spent and the cost-effectiveness of incentives would be Rs. 1.46 per rupee spent.

1.5.1 Cost-effectiveness analysis

I present approximate estimates of cost-effectiveness for the two interventions, both for the pilot and at scale. Figure 1.10 has extensive detail about the calculations, caveats, and assumptions.

Cost-effectiveness analysis for the pilot:

I find that during the pilot, the cost-effectiveness, i.e. the return per rupee spent, was 3.12 for the training intervention and 4.22 for the incentives intervention. These numbers are on an intent-to-treat basis and represent the total costs and total benefits for the entire group randomly assigned to treatment, regardless of compliance.

To evaluate the cost-effectiveness of the pilot, I use the Column 4 treatment estimates from Table 1.4. I consider the treatment effect of training to be 0.009 KMPL over 12 months and ignore the statistical insignificance. Based on the full dataset, each driver drives an average of 5972 kilometers (KM) per month at 4.712 KMPL and hence consumes an average of 1267 liters of diesel a month. Post-treatment, drivers would have a KMPL of 4.7214 (4.712+0.009) and would hence consume 1264.879 liters of diesel (5972 KM/4.7214 KMPL), implying 2.433 liters of diesel saved per driver per month. For the incentives intervention, the treatment effect is 0.01675 KMPL and post-treatment KMPL would be 4.7291 KMPL (4.712+0.01675). Thus diesel consumed would be 1262.82 liters, implying 4.492 liters of diesel saved per driver per month.

Multiplying this by 760 drivers for 12 months and using a diesel price of Rs. 57 per liter, the estimated cost savings for the pilot are Rs. 1,264,773 and Rs. 2,335,119 for the training and incentives interventions. Estimated costs are Rs. 404,794 and Rs. 553,504 respectively. Thus net fuel cost savings were Rs. 859,979 (about $13,230[18]) and Rs. 1,781,615 (about $27,410) respectively. Cost-effectiveness, i.e. the return per rupee spent, was 3.124 and 4.219 respectively. Estimated fuel savings was about 0.19% and 0.35%[19] respectively of baseline fuel consumption.

18. Using 1 USD=65 INR

19. Fuel savings of 2.433 and 4.492 liters respectively as a percentage of the baseline 1267 liters consumed

Table 1.10: Cost-Effectiveness Analysis

	Training	**Incentives**
Costs and Savings for the Pilot		
Direct costs:	Rs. 320,094[1]	Rs. 426,000[2]
Indirect costs:	Rs. 84,700[3]	Rs. 127,504[4]
Total Costs for the Pilot	Rs. 404,794	Rs. 553,504
Diesel consumed per driver per month (average)[5]	1267.312 liters (5972 KM/4.712 KMPL)	1267.312 liters (5972 KM/4.712 KMPL)
Treatment estimate for 12 month period[6]	0.00906 KMPL	0.01675 KMPL
Estimated KMPL after treatment	4.7214 KMPL	4.7291 KMPL
Diesel consumed after treatment per driver per month	1264.879 liters (5972 KM/4.7214 KMPL)	1262.82 liters (5972 KM/4.7291 KMPL)
Diesel saved per driver per month[7]	2.433 liters (0.19% of baseline)	4.492 liters (0.35% of baseline)
Diesel saved over 12 month period for 760 treated drivers	22,189 liters	40,967 liters
Total Savings from the Pilot (at price Rs. 57 per liter)	Rs. 1,264,773	Rs. 2,335,119
Net Savings from the Pilot (Savings-Costs)[8]	**Rs. 859,979 ($13,230)**	**Rs.1,781,615 ($27,409)**
Cost-Effectiveness of Pilot (Savings/Costs)	**3.124**	**4.219**
Costs and Savings At Scale		
Costs per driver[9]	Rs.558[10] (one-time cost)	Rs. 268[11] (monthly cost)
Treatment estimate during implementation [12]	0.0186 KMPL	0.0264 KMPL
Diesel consumed per month during implementation	1262.329 liters (5972 KM/ (4.712+0.0186) KMPL)	1260.25 liters (5972 KM/ (4.712+0.0264) KMPL)
Diesel saved per month	4.98 liters (1267.312-1262.329)	7.06 liters (1267.312-1260.25)
Duration of treatment effect	4 months	Ongoing
Total diesel saved	19.93 liters (one-time)	7.06 liters per month

Table continued on next page

Table 1.10: (Continued)

	Training	**Incentives**
Cost savings per driver @ Rs. 57/liter	Rs.1136 (one-time benefit)	Rs. 402 (monthly benefit)
Net Savings at Scale (Savings-Costs)	**Rs. 578 ($8.89, one-time)**	**Rs. 134 ($2.06, monthly)**
Cost-Effectiveness at Scale (Savings/Cost)	**2.036**	**1.455**

Table Footnotes:

[1] The instructor remuneration (fees+ travel+ lodging) for the 592 drivers trained in Haneef sessions was Rs. 261,000, i.e. a per driver cost of Rs. 441. The exact costs for the 134 drivers trained in PCRA sessions was unavailable. Hence I estimate the PCRA instructor remuneration to be 134*Rs.441=Rs.59,094 (number of drivers trained in PCRA sessions* per driver cost of Haneef sessions).

[2] This was the total directly paid out as incentives.

[3] Some drivers had to travel to attend the training sessions. Based on research team costs, I estimate Rs.350 travel + lodging costs for these drivers, and estimate that about 1/3rd of drivers had to travel, so the average travel costs are Rs. 117 per driver, for 726 trained drivers. I expected the KSRTC staff costs to be minimal as there was little coordination required and do not include it, or the opportunity costs of driver's time.

[4] The total costs for the field research team was Rs. 1,275,035. While this was primarily for research activities, about 10% was for implementing the incentives intervention and I use this to estimate what KSRTC staff costs would be for in-house implementation.

[5] 5972 KM is the average KM per driver per month for the 25 complete months in the panel. 4.712 KMPL is the average winsorized KMPL for all shifts in the panel dataset.

[6] Treatment estimates from Column 4 of "Table 1.4: Impact on Fuel Efficiency in the Twelve Month Post-Intervention Period".

[7] Percentage calculated as percentage of the baseline fuel consumption of 1267.312 liters.

[8] Using 1 USD=65 INR.

[9] The total pilot costs are for the entire treatment groups, i.e. intent-to-treat costs. To estimate per driver costs I divide the costs by number of drivers who were actually treated, i.e. compliers.

[10] Total pilot costs of Rs. 404,794 divided by the 726 trained drivers.

[11] Pooling all months of incentive eligibility, 41.55% of drivers were paid Rs. 500, for an expected monthly payout of Rs. 208. Dividing the estimated indirect costs of Rs. 127,504 by the 703 compliant drivers in the incentives groups gives a per driver cost of Rs. 181 over 3 months, so the per driver per month cost is Rs. 60.

[12] Treatment estimates from Column 4 of "Table 1.5: Impact on Fuel Efficiency in Four Experiment Phases".

Cost-effectiveness projections at scale:

Next, I project back-of-the-envelope estimates of the cost-effectiveness for KSRTC if the intervention is scaled up. I estimate that at scale, the cost-effectiveness of training would be Rs. 2.04 per rupee spent and the cost-effectiveness of incentives would be Rs. 1.46 per rupee spent.

The training program would have a one-time cost of Rs. 558 per driver, and a scaled-up ongoing incentives scheme would have a recurring monthly cost of Rs. 268 per month. As per Column 4 of Table 1.5, the training intervention had a marginally significant short term effect of 0.0186 KMPL in March-June 2016 and a null effect thereafter. In the midst of implementation, the incentives scheme had a treatment effect of 0.0264 KMPL in July-September 2016. Using these numbers, I estimate that the training program has a one-time benefit of Rs. 1,136 while a scaled up incentives scheme would have a recurring benefit of Rs. 402 per month. Training would have a one-time net savings of Rs. 578 (about $8.89) and a cost-effectiveness of 2.036. An incentives scheme would have a recurring monthly savings of Rs. 134 (about $2.06) and a cost-effectiveness of 1.455. These estimates involve numerous assumptions and simplifications. In particular, the RCT recruited low performing drivers. A scaled up incentives scheme would have more high-performers who achieved the target, hence costs would be higher. The heterogeneity analysis in Section 1.4.4 suggests that benefits would also be larger for high-performers. I note that for the incentives intervention, cost-effectiveness is significantly higher for the pilot as costs were only during eligibility while benefits persisted post-eligibility.

1.5.2 Cost-benefit analysis

The previous cost-effectiveness section demonstrates that investments in bus driver skills and financial incentives for drivers can cost-effectively improve fuel efficiency. A complete cost-benefit analysis of these interventions would include other societal benefits and costs of these interventions, such as safety co-benefits, the time-costs to passengers from a slower and safer journey, and positive externalities from reduced emissions.

Safety co-benefits: Theory and evidence from case studies indicate that Safe and Fuel Efficient Driving (SAFED) has significant safety co-benefits. "...on the whole, a safe driver is a green driver, with anticipation, smoothness, and sensible speed being the defining characteristics", and evidence from case studies also suggests that eco-driving training improves accident rates (Young, Birrell, and Stanton 2011). Driving style affects safety through speeding, aggressive driving habits, risky driving maneuvers, breaking traffic rules, et cetera. Safety co-benefits arise because there is significant overlap in the techniques that improve fuel efficiency and safety.

While I cannot measure the safety co-benefits of the training and incentives interventions, these are likely large. The Global Burden of Disease Project estimated that globally in 2010, road injury was the 8th leading cause of death, with a death rate of about 19 per 100,000, and 1.3 million total deaths. The death rate for India was estimated as 22 per 100,000, and the gross domestic product losses due to road crashes in India for 2014 were estimated to be as high as 4.6% (The World Bank and IHME 2014). Buses and trucks cause disproportionate road accident fatalities in India. In one analysis of fatal highway accidents, the impacting vehicles were trucks and buses for 65% and 15% of accidents respectively (Mohan et al. 2009), though victims were typically on motorized two-wheelers, bicyles, or on foot. Another study of Bangalore, Karnataka found that public transport buses were involved in 12% of fatal crashes (Kharola, Tiwari, and Mohan 2010), despite being only about 1% of vehicles.

Positive externalities from reduced emissions: Reduced fuel consumption directly reduces carbon emissions. Since this was achieved through Safe and Fuel Efficient Driving (SAFED) techniques, local emissions of particulate matter and other pollutants were also reduced, with accompanying health benefits. Thus there are substantial positive externalities from averted carbon emissions and averted local air pollutant emissions.

Positive externalities on congestion through increased public transport bus services: Lowered operating costs could indirectly improve the quantity of bus services provided by KSRTC. In general, when operation and capital costs are lower, public transport agencies are able to provide more services overall.

Negative impact for passengers and drivers from time-costs: Safe and fuel efficient driving minimizes speeding and hence may lengthen the overall duration of the journey. As driver and passenger time is valuable, this is a cost of SAFED.

Other: SAFED driving may also have other minor costs and benefits. Positive externalities could include reduced congestion due to less lane switching and smoother driving by bus drivers. Costs could include the potential rental value of the bus for the longer duration. SAFED may also change the maintenance costs for the vehicles and could either increase costs due to the bus being operated for greater durations, or lower costs due to better driving causing less damage.

1.6 Conclusion

A sizable literature demonstrates that the quality of government services in developing countries, including India, is frequently poor (Finan, Olken, and Pande 2017). Public service providers such as teachers and health care workers display high rates of absenteeism (Chaudhury et al. 2006; Banerjee, Glennerster, and Duflo 2008; Duflo, Hanna, and Ryan 2012), often have limited qualifications and skills (Das et al. 2012; Das, Hammer, and Leonard 2008), and receive remuneration regardless of performance (Muralidharan and Sundararaman 2011). In India, citizens often 'vote with their feet' and turn to private service providers instead (Banerjee, Glennerster, and Duflo 2008; Kingdon 2007).

In this paper I analyze public transport provision, an important responsibility of the state. The bus driver's performance affects multiple dimensions of public transport quality, including accident rates, vehicular emissions, and fuel costs, which are the single largest cost and thus important for the quantity of bus services. I find that in this setting, human capital investment for service providers through bus driver training has minimal impact. The training program increased fuel efficiency by 0.0186 kilometers per liter, marginally significant, for four months and had no noticeable effect thereafter. In contrast, even small performance-based financial incentives lead to a demonstrable improvement. The incentives scheme improved fuel efficiency by 0.0168 kilometers per liter for a twelve month period.

This study contributes to the body of experimental evidence indicating that outcome-based financial incentives improve government worker performance (Finan, Olken, and Pande 2017). It also demonstrates successful interventions for reducing the negative externalities of motorized road transport, which is a leading global cause of death through vehicular pollution and road injuries. The incentives offered in this study were low-powered due to public-sector related constraints. Private sector bus or truck fleets could potentially provide much larger incentives to drivers and see significantly reduced fuel consumption. In the future, I hope to collect data on safety outcomes to provide direct evidence on how interventions like training or incentives can affect safety.

Chapter 2

Personality Traits and Job Performance: An Analysis of Public Sector Bus Drivers in India

2.1 Introduction

A growing body of evidence indicates that personality traits of government employees are an important determinant of job performance (Finan, Olken, and Pande 2017; Callen et al. 2015). Research also shows that personality traits are malleable (Kautz et al. 2014). Callen et al. (2015) point out that the importance of personality, together with the evidence that personality traits are malleable, suggests that public service delivery can be improved by taking into account non-cognitive attributes of employees. They suggest three channels to improve service delivery, namely, incorporating personality traits into recruitment methods, incorporating personality traits into hiring and firing decisions, and interventions that directly target non-cognitive attributes. The literature demonstrates that recruitment methods affect the personality traits of the applicant pool (Dal Bó, Finan, and Rossi 2013; Ashraf, Bandiera, and Lee 2016; Deserranno 2017), and that the success of interventions to improve job performance vary based on employee personality traits (Callen et al. 2015; Ashraf, Bandiera, and Jack 2014).

Designing interventions that incorporate personality traits to improve screening, recruitment, or incentive structures is a fruitful area of research (Finan, Olken, and Pande 2017; Callen et al. 2015). To design personality-based interventions on screening, recruitment, hiring and firing, or incentive structure, it is necessary to have a better understanding of the importance of personality traits under status quo policies in various settings. What is the distribution of personality traits among government employees under current screening, recruitment, and retention policies? In what settings, under what personnel policies, do personality traits affect job performance? When do changes in incentive structure have heterogeneous treatment effects based on employee personality traits?

This paper contributes to the emerging literature by analyzing how personality traits relate to attendance, job performance, changes in incentive structure, and human resource decisions among public sector bus drivers in Karnataka, India. I study 1522 bus driver employees of the Karnataka State Road Transport Corporation (KSRTC), India's 5th largest public sector bus service provider. This builds on previous research, detailed in Chapter 1 of this dissertation, in which I implemented a field experiment with KSRTC to evaluate the impact of two interventions to improve safety and fuel efficiency: a training program and financial incentives.

I pair extensive administrative data from KSRTC with two surveys of bus drivers. A September - October 2015 baseline survey has experimental measures of dishonesty through cheating on a random dice task, risk aversion through an ordered lottery selection task, and pro-socialness through a charity dictator game. I also measure time preferences through hypothetical time discounting questions. For January 2015-February 2017, I collect detailed administrative data on driver attendance at the workplace, as measured by the number of shifts an employee is driving, along with shift-level fuel efficiency data for each driver. Finally, for 761 drivers, I conduct a second follow-up survey, using workplace interviews in the period April - September 2016 to understand why drivers are not in attendance. For these 761 drivers, I collate a dataset of human resource decisions on firing, suspension, transfers, authorized leave, and duty reassignments.

Using these data sets, I first investigate the distribution of time preferences, risk aversion, dishonesty, and pro-socialness under the current screening, hiring, and firing decisions at KSRTC, and how these experimental measures correlate with each other. I analyze whether personality traits predict job performance, using two metrics: driver attendance and the driver's average fuel efficiency. I evaluate whether the training program or financial incentives had heterogeneous effects based on these behavioral traits. Finally, I investigate whether behavioral measures predict human resource decisions on transfers, suspensions, and duty reassignments.

I find that behavioral measures rarely predict job performance, nor do they predict who responds to interventions to improve fuel efficiency. I evaluate whether behavioral traits predict attendance, and find no correlations using a broad measure of attendance, the number of shifts an employee is driving. Risk aversion, time preferences, and dishonesty do not predict fuel efficiency. Pro-socialness is negatively correlated with fuel efficiency. I analyze whether either a training intervention or a financial incentives intervention has heterogeneous treatment effects based on the bus driver's baseline measures of risk aversion, dishonesty, time preferences, and pro-socialness, and find no differential effects.

To understand what distinguishes this setting from the many other contexts in which personality traits are correlated with job performance, I delve into the descriptive statistics for this group. In this population of public sector bus drivers in India, I find very high levels of pro-socialness and honesty, in comparison to the literature. Fully 60% of the drivers donate the maximum amount, 98.75% donate at least some amount, and the average donation is 82% of the endowment. These pro-socialness levels are very high compared to the literature using the charity dictator game (see Meier 2007 for a review; Hanna and Wang 2017; Ashraf, Bandiera, and Jack 2014). Honesty levels are also much higher than most other studied populations (Hanna and Wang 2017; Kröll and Rustagi 2016; Rosenbaum, Billinger, and Stieglitz 2014; Bucciol and Piovesan 2011; Houser, Vetter, and Winter 2012; Fischbacher and Föllmi-Heusi 2013; Barfort et al. 2015). Using individual measures of dishonesty, 18% of the drivers are dishonest at the 5% statistical significance level. At the group level, 67.5% of the respondents are unconditionally honest and do not cheat at all, 18.5% are partially dishonest, and 14% are strongly dishonest. I also find that risk averse drivers are less likely to cheat on a random dice task or have hyperbolic time preferences, and patient drivers are more likely to be pro-social.

Finally, I find that personality traits predict a number of human resource decisions. I use a detailed measure of attendance that incorporates the human resource decisions for why the employee is not driving, such as that the driver is suspended, retired, on leave, transferred, and so on. Dishonest drivers are more likely to be transferred to other branches or reassigned to other duties, pro-social drivers are less likely to be on authorized leave, and risk-averse drivers are more likely to be in attendance, less likely to be disciplined or absent, and more likely to be on authorized leave.

Thus, in summary, personality traits have minimal effect on job performance. Personality traits do not predict driver attendance. Pro-socialness predicts lower fuel efficiency performance, but dishonesty, time preferences, and risk aversion are not correlated. I do not find that a change in incentive structure has differential effects based on personality traits, nor does a training program.

My results suggest that human resource policies may explain the findings. KSRTC has a highly pro-social and honest group of employees, compared to the literature from other contexts. While there is heterogeneity among drivers, the entire distribution of outcomes, such as the mean, median, and the fraction of respondents who are pro-social or honest, is very positive. This is likely to be a consequence of recruitment and human resource management policies at KSRTC. I find evidence that dishonest drivers are more likely to be transferred or reassigned, and risk-tolerant drivers are more likely to be disciplined or absent. Overall, my results suggest that KSRTC manages its employee pool such that among retained drivers, personality traits largely do not predict job performance. I hypothesize that the high wage premium KSRTC drivers receive leads to a large applicant pool and allows management to be selective on multiple dimensions during recruitment, selection, and retention.

This paper contributes to several strands of the literature on personality traits, job performance, and public service delivery. First, I show that under selective human resource policies, organizations can hire and retain employees with beneficial personality traits, and subsequent to selection, personality traits do not predict job performance. This is related to Ashraf, Bandiera, and Lee (2018)'s finding that while recruitment posters advertising career benefits attract more talented but less pro-social candidates on average, subsequent to selection, the candidates who are actually hired are equally pro-social and more talented. It is also related to Callen et al. (2015)'s suggestion that recruitment and hiring should incorporate personality traits. At KSRTC, human resource decisions are presumably based on candidate potential and job performance, and not explicitly on personality traits, yet this indirectly creates an unusually honest and pro-social group. Thus, to the extent that personality traits affect job performance, recruitment and hiring do indirectly incorporate personality traits. Second, I add to the body of work using experimental measures of dishonesty, time preferences, risk aversion, and pro-socialness in developing countries. I compare the distribution of these traits in this Indian bus driver population to students and workers in other settings (Hanna and Wang 2017; Banerjee et al. 2016; Ashraf, Bandiera, and Jack 2014; Ashraf, Karlan, and Yin 2006). Third, I contribute to the literature demonstrating the relationship between behavioral traits and real-world outcomes (Hanna and Wang 2017; Lagarde and Blaauw 2014), and show that personality traits predict a number of human resource decisions in this setting. Fourth, I add to the research on whether interventions to improve job performance can have heterogeneous effects based on the worker's behavioral traits (Ashraf, Bandiera, and Jack 2014; Callen et al. 2015; Banerjee et al. 2016), and find no effect in this population.

Additionally, I also contribute to the literature on experimental measures of honesty in field settings using a dice task (Fischbacher and Föllmi-Heusi 2013; Hanna and Wang 2017; Kröll and Rustagi 2016; Rosenbaum, Billinger, and Stieglitz 2014). I demonstrate a relationship between risk aversion and honesty on the dice task, which was not found among students (Barfort et al. 2015) and may be specific to workplace settings. In Appendix B, I compare the measurement errors from several possible dice task metrics. I note that the sample median metric has substantial Type I error if the population is honest, and that metrics based on the 99th and 95th percentiles of the theoretical distribution are able to identify strongly dishonest respondents but have low power to identify partially dishonest respondents.

This paper is structured as follows. Section 2.2 describes the background, the field experiment with KSRTC, and KSRTC personnel practices. Section 2.3 describes the surveys, administrative data, and experimental measures. Section 2.4 presents descriptive statistics of the experimental measures and the correlation across measures. Section 2.5 analyzes whether baseline experimental measures of behavioral traits predict driver attendance or fuel efficiency. Section 2.6 investigates whether the training program or financial incentives had heterogeneous effects based on personality traits. Section 2.7 tests whether personality traits predict human resource decisions, and Section 2.8 concludes.

2.2 Background

I study public sector bus driver employees of the Karnataka State Road Transport Corporation (KSRTC). KSRTC is a public transport agency owned by the Government of Karnataka, one of India's southern states. It is the 5th largest public sector bus service provider in India and serves 2.6 million passengers a day[1]. KSRTC operates passenger buses for intercity travel and employs over 20,000 bus drivers across 79 branches.

Hiring at KSRTC is currently based on a standardized testing process for applicants who meet certain qualifications, followed by a two year probation period. However, hiring practices used to be more informal and patronage based, and many older drivers were hired through that system. Management cannot easily dismiss workers, but they can suspend them relatively easily. KSRTC management can also transfer drivers to other branches. Since some locations are favored, transfers can be used as both a carrot and a stick. KSRTC drivers are also occasionally reassigned to other roles, most commonly as a bus conductor.

KSRTC drivers are very well compensated, with total remuneration ranging from Rs. 30,000-50,000 per month. New entrants get Rs. 20,000-25,000 a month. In comparison, the salary range for privately employed car drivers in Bangalore is Rs. 10,000-20,000. KSRTC drivers are unionized, with a fairly powerful union with frequent strikes[2]. Promotion is primarily based on tenure, and wage increments are negotiated by the union and are about 8-10% a year with an additional annual increment after 15 years of service. Some depots have small performance-based bonuses such as annual prizes for safety.

This paper builds on a previous field experiment detailed at length in Chapter 1 of this dissertation, in which I evaluate two interventions: a Safe and Fuel Efficient Driving (SAFED) training program, and a financial incentives scheme for achieving fuel efficiency targets. 1522

1. "History of KSRTC" webpage, Karnataka State Road Transport Corporation. Available at http://www.ksrtc.in/pages/history.html. Accessed on 15th August 2017.

2. About 5-6 strikes over the 1.5 years of project implementation.

drivers from 34 branches were randomly assigned to a control group, training, incentives, or both interventions. I collected extensive administrative data on shift-level fuel efficiency for a 26 month period and used a difference-in-difference empirical strategy. The training program increased fuel efficiency in the short term for four months and had no effect thereafter. The incentives scheme increased fuel efficiency for a twelve month period. I find no evidence of any complementarities between training and incentives.

2.3 Data

2.3.1 Survey and Administrative Data

I combine three sources of data: a 2015 baseline survey of bus drivers, a second 2016 survey on employee attendance and human resource decisions, and administrative data on shift-level attendance and fuel efficiency for 2015-2017.

In September and October 2015, I recruited 1522 drivers from 34 KSRTC branches to participate in the research. I conducted a baseline survey of these drivers which included experimental measures of pro-socialness, dishonesty, risk aversion, and time preferences. These measures are described in detail in Section 2.3.2. The survey also collected demographic details, education history, questions on previous training experiences at KSRTC, questions on the status quo incentive schemes at KSRTC, a set of questions on baseline knowledge of the theory of safe and fuel efficient driving, a job satisfaction measure, and an accident history.

I also gathered extensive KSRTC administrative data on driver attendance and fuel efficiency. For each of the participating drivers, I collected data on every shift they drove between January 2015 and February 2017. For each shift, I collected data on kilometers traveled (KM), diesel consumption (liters), vehicle used, route traveled, number of drivers on the shift, date, and other shift-level variables. I use this to estimate shift level fuel efficiency, i.e. kilometers per liter (KMPL) (calculated as kilometers traveled in KM/diesel consumed in liters).

Finally, in 2016, for 761 drivers, I conducted a follow-up survey at the workplace. These drivers were tracked for three months each at some point in the period April - September 2016. The field survey team made contact with each of these 761 drivers, or, if they were not available at the workplace, asked the management team why the driver was missing. I use the responses provided by the KSRTC management to collate a dataset of human resource decisions for these drivers.

2.3.2 Experimental Measures of Behavioral Traits

I measure respondents' pro-socialness through a charity dictator game, risk aversion through an ordered lottery selection task, dishonesty through propensity to cheat on a random dice task, and time preferences through hypothetical time discounting questions. I describe these experimental measures in detail below.

Risk aversion: To elicit risk aversion, I used an ordered lottery selection task, following Eckel and Grossman (2002, 2008). Respondents selected one out of five ordered lotteries with linearly increasing expected values and greater standard deviations. Once the respondent picked a lottery, it was immediately played and respondents received their winnings. Each lottery involved a coin toss. The first lottery choice, Lottery 1, yielded Rs. 40 if the coin landed "Heads" or "Tails". The final lottery choice, Lottery 5, yielded Rs. 120 if the coin landed "Heads" and Rs. 0 if the coin landed "Tails". This ordered lottery selection task was selected after extensively piloting risk aversion measures. Respondents found most commonly used risk aversion measures difficult to comprehend. The ordered lottery selection task was carefully designed to be appropriate for this setting. The lottery was a coin toss as this is well understood due to its use in India's popular sport, cricket. Before respondents selected a lottery, they were questioned on what the outcomes would be if hypothetical people played the game. If respondents seemed unclear on how the game worked, the instructions were repeated. Appendix Figure B.1 shows the lottery task instructions.

Pro-socialness: To measure pro-socialness, I conducted a standard dictator game where respondents could divide Rs. 30 between themselves and a charity of their choice out of 7 options. The charity choices and task instructions followed Hanna and Wang (2017). Respondents were told they could keep the entire Rs. 30, donate the entire Rs. 30, or split it. Every rupee they donated was doubled by the research team. For instance, if a respondent donated Rs. 30, Rs. 60 was given to the charity. Appendix Figure B.2 shows the dictator game instructions.

Dishonesty: To measure dishonesty, I implemented a dice task, closely following Hanna and Wang (2017). Participants were asked to privately roll a six-sided die 42 times and record each outcome. Respondents received Rs. 0.50 for each point, i.e. Rs. 0.50 for a 1 and Rs. 3.00 for a 6, for a maximum possible payment of Rs. 126 and a minimum possible payment of Rs. 21. Appendix Figure B.3 shows the dice task instructions. Respondents were assured of privacy as the survey team was asked to step away during this task. I compare the individual's total dice points to the theoretical distribution of total points for 42 dice rolls to get an estimate of cheating. As privacy was ensured, it is not possible to measure dishonesty with certainty, but the total dice points measure is highly correlated with cheating.

Time Preferences: I measured time preferences through a series of hypothetical time discounting questions, closely following Ashraf, Karlan, and Yin (2006). Early in the survey, I asked respondents "Suppose you had the choice, would you prefer to receive Rs. 800 guaranteed today, or Rs. 1000 guaranteed in 1 month?". If they preferred Rs.800 today, I repeated the question offering Rs. 1200 in 1 month. If they still preferred Rs. 800 immediately, I asked how much money they would want to wait for a month. After about 20 minutes, I repeated the questions with reference to 6 and 7 months. Based on the responses, I characterize respondents as having consistently patient, consistently impatient, patient now-impatient later, and hyperbolic time preferences. Table 2.2 provides more detail on the questions, responses, and time preference categories.

2.4 Outcomes of Experimental Measures of Behavioral Traits and Correlation Across Measures

I present summary statistics on the experimental measures of behavioral traits, along with the correlations across measures. I find very high levels of pro-socialness and honesty, a uniform distribution of risk aversion, and substantial levels of patience. Risk averse drivers are less likely to cheat on the dice task. Unlike Hanna and Wang (2017) and Barfort et al. (2015), I find no negative correlation between pro-socialness and dishonesty. A correlation between risk aversion and honesty on the dice task may be specific to workplace settings, given that Barfort et al. (2015)

find no correlation among students. I also find that consistently impatient respondents are less pro-social, hyperbolic respondents are less risk-averse, and consistently patient respondents are more pro-social.

2.4.1 Descriptive Statistics on Experimental Measures of Behavioral Traits

Table 2.1 presents descriptive statistics on demographic data, risk aversion, pro-socialness, dishonesty, and time preferences.

Dishonesty: I find high levels of honesty in this population. I first report the individual-level dishonesty measures used by Hanna and Wang (2017). Across 42 dice rolls, the theoretical mean is 147 points. Among the 1511 respondents for the dice task, the mean and median is 153 points, 9% of the respondents have scores above the theoretical 99th percentile, and 18% have scores above the theoretical 95th percentile. Figure 2.1 shows the distribution of dice points in the sample and compares it to the theoretical distribution.

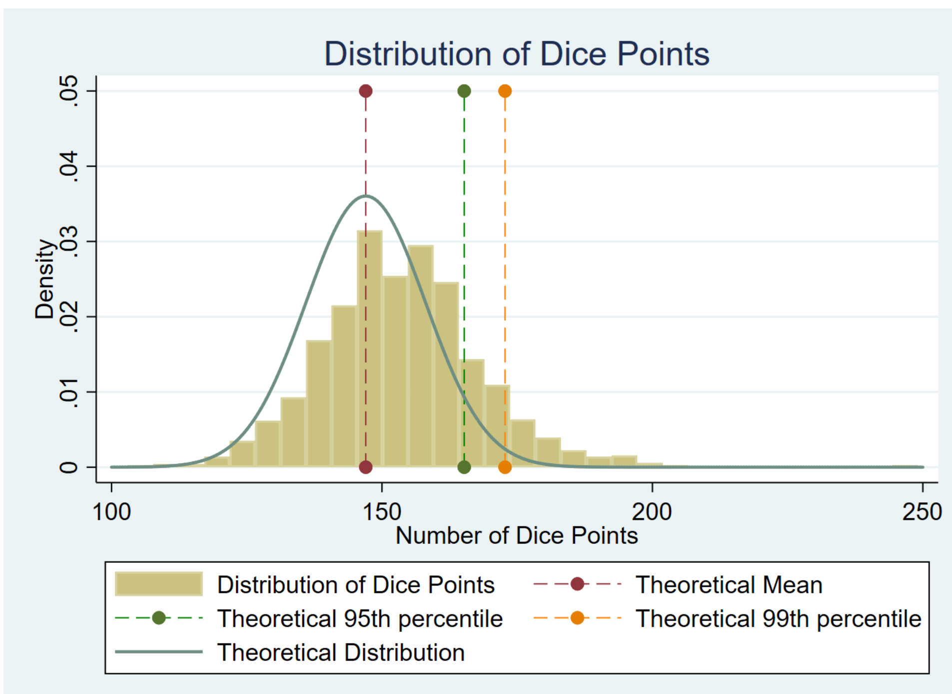

Figure 2.1: Distribution of Dice Points in Dishonesty Task

Table 2.1: Descriptive Statistics on Behavioral Traits

	Mean	Standard Deviation	Count
Demographic & Career:			
Age (years)	38.88	8.46	1522
Education (years)	10.27	1.92	1522
KSRTC tenure (years)	10.01	8.14	1522
Job satisfaction (score)	4.54	0.84	1520
Attended driving school	0.04	0.20	1522
Attended KSRTC training	0.86	0.35	1522
Has had accident	0.58	0.49	1520
Risk Aversion:			
Risk aversion (score)	3.06	1.47	1520
Pro-Socialness:			
Amount Donated (Rs.)	24.60	7.53	1521
Donated $>=$ Median	0.60	0.49	1521
Dishonesty:			
Dice Task points (points)	153.31	14.52	1511
Dice points $>$ 95p	0.18	0.38	1511
Dice points $>$ 99p	0.09	0.28	1511
Time Preferences:			
Consistently Patient	0.58	0.49	1509
Consistently Impatient	0.12	0.33	1509
Hyperbolic	0.22	0.41	1509
Patient Now Impatient Later	0.08	0.26	1509

Time Preferences are CoPa(Consistently Patient), CoIm(Consistently Impatient),Hype(Hyperbolic), and PNIL(Patient Now Impatient Later)(see Table 2.2 for details. Dice points $>$ 99p and $>$ 95p are indicators that the respondent's total dice points was above the 99th and 95th percentiles respectively of the theoretical distribution. Job satisfaction scores range from 1 (very dissatisfied) to 5 (very satisfied). Risk aversion scores range from 1 (riskiest lottery selected) to 5 (safest lottery selected). 'Amount Donated (Rs.)' indicates the amount donated (between Rs. 0-30) in a pro-socialness game. 'Donated $>=$ median' indicates the driver donated the median amount, Rs.30 (the maximum).

I next report group-level measures of dishonesty. I estimate that 67.5% of the respondents are unconditionally honest and do not cheat at all, 18.5% are partially dishonest, and 14% are strongly dishonest, i.e. they report dice points above the 95th percentile. I present estimated frequency distributions for honest and dishonest respondents in Appendix Table B.3. Appendix B has a detailed discussion of how I estimate frequency distributions, along with a discussion of the measurement error from several possible metrics.

This population, with 67.5% unconditionally honest respondents, is more honest than most other groups noted in the literature. In a review of experimental evidence on honesty, Rosenbaum, Billinger, and Stieglitz (2014) summarize outcomes from different types of honesty games, and the vast majority of study populations report honesty levels either lower than or similar to my sample. Looking specifically at dice tasks and other games of chance, reported honesty levels have been lower in almost all other samples. One exception is that Hanna and Wang (2017) find similar levels of dishonesty among government nurses in Karnataka, India. Using the same task, they find that the mean dice points among nurses was 152 points, and 9.4% reported points above the theoretical 99th percentile. The nurses were paid in candy, while the bus drivers were paid in cash, yet cheating levels are strikingly similar. It is possible that nurses would have cheated to a different extent with a cash reward. Using the same dice task, Kröll and Rustagi (2016) find in a study of 72 informal milkmen in Delhi that 50% are unconditionally honest, 37.5% are weakly dishonest, and 12.5% are strongly dishonest. The honesty among the KSRTC bus drivers is much higher than what students demonstrate on the dice task. Hanna and Wang (2017) study college students in Karnataka and find very high levels of cheating. The mean of dice points was 168 and 34% of students reported scores above the theoretical 99th percentile. Fischbacher and Föllmi-Heusi (2013) also find high cheating levels among Swiss students on a similar dice task. 20% of the students lied to the full extent possible, 41% were partial liars, and 39% were honest. In another variation of this task, Barfort et al. (2015) study students in Denmark and find that 13% of students cheat all the time, 10% barely cheat, and 77% are in-between. Reported honesty levels are also lower in other games of chance. In one coin toss game with children, Bucciol and Piovesan (2011) find around 85% report the high outcome, relative to a theoretical 50%, implying that 30% were fully honest[3]. Among students in Munich, 74.5% reported the better outcome in a coin toss game, implying that about 50% were fully honest (Houser, Vetter, and Winter 2012).

Pro-socialness: I find very high levels of pro-socialness in this population. The median amount donated is Rs. 30, the maximum possible, which was selected by fully 60% of the respondents. The mean donation is Rs. 24.60. Thus, on average, drivers kept 18% of their endowment and donated 82%. 98.75% of the drivers donate at least some money. Figure 2.2 shows the entire distribution of amounts donated. The distribution is concentrated around round numbers - Rs. 30 is the most common, followed by Rs. 20, Rs. 10, and Rs. 15. In the development economics literature, donation rates have usually been lower in charity dictator games. Hanna and Wang (2017) find that students in Karnataka kept 59% of their endowment on average and donated 41%. Ashraf, Bandiera, and Jack (2014) find that community health workers kept 55% on average and donated 45% of their endowment. Turning to the broader literature, the donation levels found among the Karnataka bus drivers are large compared to most estimates (see Meier (2007) and Hanna and Wang (2017) for summaries of the literature). There is significant heterogeneity in the amount donated in dictator games, based on the setting, the exact framing of the experiment, and so on (see Meier (2007) for a review of the economic theories and field evidence on pro-social behavior). Dictator games in which the donated amount goes to a charity, rather than an anonymous individual, have higher donation rates (Eckel and Grossman 1996).

3. Assuming that the distribution among honest people matches the theoretical distribution, and assuming that those who report the low outcome are all honest, I expect 15%*2 to be unconditionally honest. See Appendix B.2 for more details.

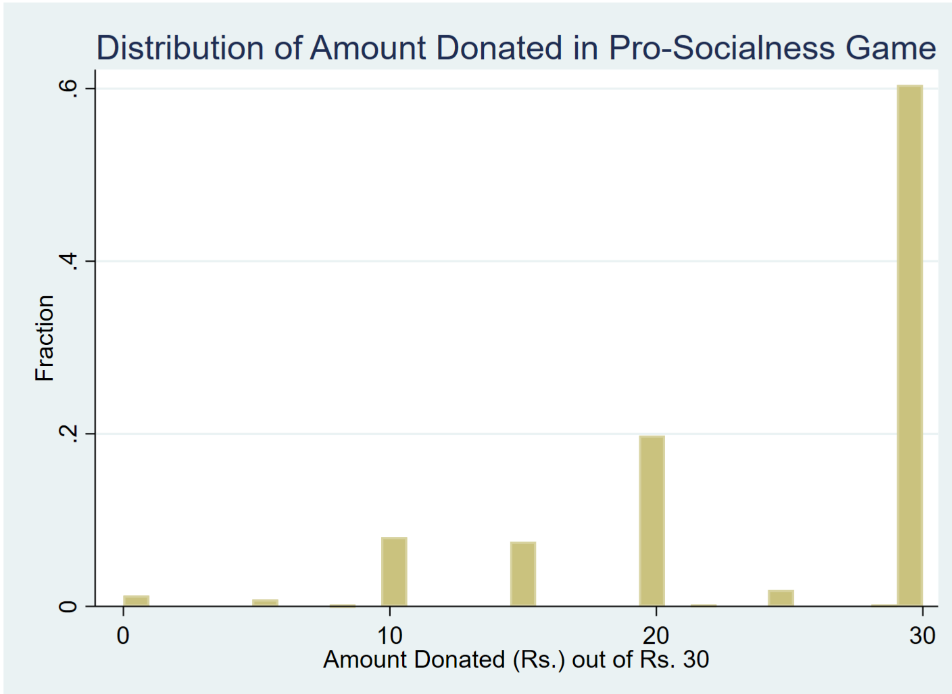

Figure 2.2: Amount Donated in Pro-Socialness Game

Risk aversion: The distribution of risk aversion scores in this population is close to uniform. As seen in Figure 2.3, 20.8% of drivers chose the riskiest lottery and had a risk aversion score of 1. 18% had a score of 2, 20.7% had a score of 3, 15.8% had a score of 4, and 24.7% chose the safest lottery and had a score of 5. I use the Eckel-Grossman ordered lottery measure of risk aversion (Eckel and Grossman 2002, 2008). Comparing my results to the literature, I note that the bus driver sample is more risk averse than the original samples by Eckel and Grossman, and has similar levels of risk aversion to a sample of Indian students. In a lab experiment among students in the USA, Eckel and Grossman (2002) found that 5% chose the safest lottery, 17.5% the next safest, 32.5% the middle lottery, 20.5% the second riskiest, and 24.5% chose the riskiest lottery[4]. In a later study among a similar population, they found that 4.3% chose the safest lottery, 16.4% the next safest, 32.8% the middle lottery, 22.7% the second riskiest, and 23.8% chose the riskiest lottery (Eckel and Grossman 2008). Among Indian undergraduate students doing the Eckel-Grossman task, Dasgupta et al. (2016) find that 20% chose the safest lottery, 22%

4. I use the reported frequencies of lottery choice in Table 2. 10 participants chose lottery 1, the safest lottery, 35 chose lottery 2, 65 chose lottery 3, 41 chose lottery 4, and 49 chose lottery 5, out of 200 participants.

the next safest, 21% the middle lottery, 18% the fourth lottery, and 19% chose lotteries 5 and 6 (which are similar to lottery 5 in my study).

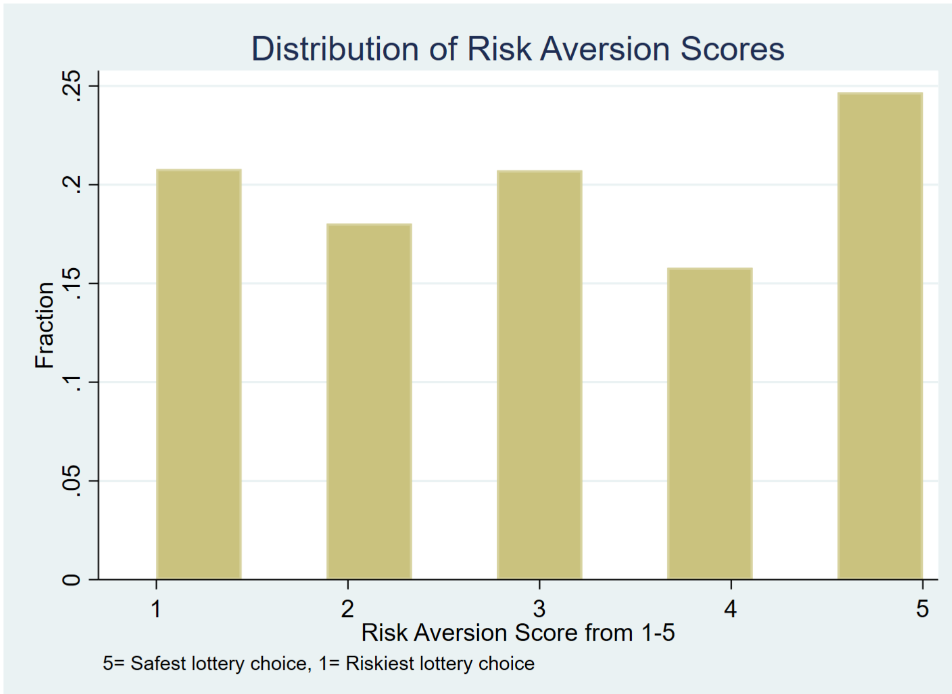

Figure 2.3: Distribution of Risk Aversion Scores

Time preferences: Table 2.1 summarizes the distribution of time preferences, while Table 2.2 shows details of the responses to specific hypothetical questions. Overall, 58% of the drivers are consistently patient and always choose to wait one month to receive a larger sum, regardless of whether the hypothetical scenario is in the present or future. Around 12% are consistently impatient. 22% demonstrate hyperbolic preferences, and are impatient in a hypothetical scenario in the present and are patient in a hypothetical scenario in the future. Finally, 8% have 'Patient Now Impatient Later' preferences, where they are impatient in a hypothetical future scenario but patient in the present. The time preferences questions I use are based on Ashraf, Karlan, and Yin (2006), where they study clients of a Philippine bank. In that population, patience levels are much lower. Using the classification described above, 34.4% are consistently patient, 25.3% are consistently impatient, 25.7% have hyperbolic time preferences, and 14.6% have 'Patient Now Impatient Later' preferences.

Table 2.2: Details of Responses to Time Preference Questions

		Q2. Indifferent between Rs. 800 in 6 months and Rs. X in 7 months				
		X < 1000	1000 < X < 1200	1200 < X	Missing	Total
Q1. Indifferent between Rs. 800 now and Rs. X in 1 month	X < 1000	875 (57.49) [CoPa]	40 (2.63) [PNIL]	67 (4.40) [PNIL]	2 (0.13) [Miss]	984 (64.65)
	1000 < X < 1200	82 (5.39) [Hype]	27 (1.77) [CoIm]	7 (0.46) [PNIL]	0 (0.00) [Miss]	116 (7.62)
	1200 < X	225 (14.78) [Hype]	25 (1.64) [Hype]	161 (10.58) [CoIm]	0 (0.00) [Miss]	411 (27.00)
	Missing	1 (0.07) [Miss]	0 (0.00) [Miss]	0 (0.00) [Miss]	10 (0.66) [Miss]	11 (0.72)
	Total	1183 (77.73)	92 (6.04)	235 (15.44)	12 (0.79)	1522 (100.00)

In each cell I report the number of drivers, with percentages in parentheses and time preference category in square brackets. CoPa is Consistently Patient, Hype is Hyperbolic, CoIm is Consistently Impatient, PNIL is Patient Now, Impatient Later, and Miss indicates Missing data.

The X in the rows are determined by responses to Questions 1a and 1b. 1a: 'Suppose you had the choice, would you prefer to receive Rs. 800 guaranteed today, or Rs.1000 guaranteed in 1 month?' If the respondent preferred Rs. 1000 in 1 month, I assume X < 1000. If the respondent preferred Rs.800 now, I asked Question 1b: 'Suppose you had the choice, would you prefer to receive Rs. 800 guaranteed today, or Rs.1200 guaranteed in 1 month?' If the respondent preferred Rs. 1200 in 1 month, I assume 1000 < X < 1200. If the respondent preferred Rs. 800 now, I assume 1200 < X.

About 20 minutes later in the survey, I repeated the questions with respect to 6 and 7 months (Question 2a: 'Suppose you had the choice, would you prefer to receive Rs. 800 guaranteed in 6 months, or Rs.1000 guaranteed in 7 months?') The X in the columns are determined by responses to Questions 2a and 2b.

2.4.2 Correlation Across Experimental Measures

In this section, I study which factors are correlated with dishonesty, risk aversion, time preferences, and pro-socialness in this setting. In particular, unlike the dictator game or ordered lottery task, the dice task to measure dishonesty is fairly new (Fischbacher and Föllmi-Heusi

2013; Hanna and Wang 2017). Since it is increasingly used in experimental economics (Banerjee et al. 2016; Barfort et al. 2015; Kröll and Rustagi 2016), it is helpful to understand the factors that correlate with dice task outcomes in different field settings. Table 2.3 shows the factors that predict dishonesty and pro-socialness, and Table 2.4 shows the factors that predict risk aversion and consistently patient time preferences. I use OLS regressions, where odd numbered columns analyze the correlation of the behavioral trait with basic demographic variables and with other behavioral measures, and even numbered columns also include career characteristics.

Dishonesty: I find that risk averse drivers are less likely to cheat on the dice task. Unlike Hanna and Wang (2017) and Barfort et al. (2015), I do not find that less pro-social drivers are more likely to cheat on the dice task. Cheating on the dice task does not correlate with age, education, time preferences, or career characteristics such as job tenure or job satisfaction, with the exception that dishonest respondents are more likely to have attended KSRTC training in the past. The outcome variable in Table 2.3 is an indicator for drivers who reported dice points above the 95th percentile of the theoretical distribution (see Appendix B.2 for a discussion of alternate metrics).

My results suggest that workplace respondents, especially risk averse respondents, may fear hidden consequences to cheating. I conduct the dice task in the workplace and not a university, and find that risk aversion is negatively correlated with cheating. In contrast, Barfort et al. (2015) implement a variation of this dice task among students in Denmark, along with an experimental measure of risk aversion, and find no correlation. Hanna and Wang (2017), who report much higher levels of cheating among students than nurses, note that one possible explanation of the differences is that the nurses were in a workplace setting. Given the difference between my result and Barfort et al. (2015), this explanation seems likely.

Table 2.3: Factors that Predict Dishonesty and Pro-Socialness

	Dishonesty: Dice points > 95p		Pro-Socialness: Donated >= Median	
	(1)	(2)	(3)	(4)
Behavioral Traits:				
Risk aversion (score)	-0.0215***	-0.0218***	-0.0123	-0.0108
	(0.00693)	(0.00695)	(0.00848)	(0.00847)
Amount donated >= median	-0.0211	-0.0222		
	(0.0203)	(0.0203)		
Dice points > 95p			-0.0348	-0.0364
			(0.0333)	(0.0332)
TP=Consistently Impatient	-0.00712	-0.0106	-0.171***	-0.174***
	(0.0310)	(0.0308)	(0.0399)	(0.0399)
TP=Hyperbolic	-0.00193	-0.00383	-0.0254	-0.0213
	(0.0248)	(0.0251)	(0.0316)	(0.0316)
TP=Patient Now Impatient Later	-0.0429	-0.0414	-0.0378	-0.0279
	(0.0341)	(0.0342)	(0.0490)	(0.0492)
Demographic Characteristics:				
Age (years)	-0.00139	-0.00166	0.00141	-0.00751**
	(0.00137)	(0.00274)	(0.00185)	(0.00348)
Education (years)	-0.000921	-0.00161	0.0158*	0.0227***
	(0.00647)	(0.00676)	(0.00828)	(0.00850)
Career Characteristics:				
KSRTC tenure (years)		0.000147		0.0110***
		(0.00299)		(0.00373)
Job satisfaction (score)		0.00216		0.0293*
		(0.0123)		(0.0158)
Attended driving school		0.0731		-0.00162
		(0.0551)		(0.0631)
Attended KSRTC training		0.0637**		0.0177
		(0.0247)		(0.0364)
Has had accident		-0.0136		-0.00719
		(0.0205)		(0.0262)
Constant	0.321***	0.281**	0.459***	0.476***
	(0.110)	(0.135)	(0.143)	(0.177)
N	1500	1500	1500	1500

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Parentheses contain robust standard errors. Dice points > 95p indicates total dice points was above the 95th percentile of the theoretical distribution. Donated >= median indicates total amount donated in the pro-soialness game was at or above the median.

Table 2.4: Factors that Predict Risk Aversion and Time Preferences

	Risk Aversion: Risk Aversion (score)		Time Preferences: Consistently Patient	
	(1)	(2)	(3)	(4)
Behavioral Traits:				
Risk aversion (score)			0.00791	0.0101
			(0.00866)	(0.00865)
Amount donated >= median	-0.112	-0.0985	0.0726***	0.0693***
	(0.0768)	(0.0770)	(0.0262)	(0.0262)
Dice points > 95p	-0.320***	-0.324***	0.0183	0.0211
	(0.103)	(0.103)	(0.0334)	(0.0333)
TP=Consistently Impatient	-0.0394	-0.0537		
	(0.119)	(0.119)		
TP=Hyperbolic	-0.172*	-0.189**		
	(0.0946)	(0.0947)		
TP=Patient Now Impatient Later	0.179	0.145		
	(0.138)	(0.139)		
Demographic Characteristics:				
Age (years)	0.00778	0.0123	0.00161	0.00318
	(0.00553)	(0.0106)	(0.00186)	(0.00352)
Education (years)	-0.00188	-0.0101	-0.00381	-0.000839
	(0.0246)	(0.0255)	(0.00820)	(0.00850)
Career Characteristics:				
KSRTC tenure (years)		-0.00499		-0.00222
		(0.0113)		(0.00378)
Job satisfaction (score)		-0.125***		0.0437***
		(0.0436)		(0.0154)
Attended driving school		0.178		-0.175***
		(0.197)		(0.0628)
Attended KSRTC training		0.0749		0.00711
		(0.104)		(0.0365)
Has had accident		-0.00882		0.0186
		(0.0789)		(0.0265)
Constant	2.923***	3.385***	0.485***	0.202
	(0.417)	(0.503)	(0.143)	(0.174)
N	1500	1500	1500	1500

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Parentheses contain robust standard errors. Consistently patient=1 if time preferences were CoPa and =0 for Hype, CoIm, or PNIL (see Table 2.2 for details.). Risk aversion scores range from 1 to 5 (safest lottery selected). Dice points > 95p indicates total dice points was above the 95th percentile of the theoretical distribution.

Pro-socialness: I find that consistently impatient respondents donate less in the dictator game, relative to the reference group of consistently patient drivers, and find no correlation between pro-socialness and dishonesty, or pro-socialness and risk aversion. I also find that pro-socialness correlates positively with education, tenure at KSRTC, and job satisfaction, and negatively with age. Table 2.3 shows the factors that predict pro-socialness, measured as an indicator for drivers who donated the median amount, which was the maximum amount, Rs. 30.

Risk Aversion: Risk aversion is negatively correlated with reporting dice points above the 95th percentile of the theoretical distribution, negatively correlated with having hyperbolic preferences relative to the consistently patient drivers, and negatively correlated with job satisfaction. Table 2.4 shows the factors that predict risk aversion, as measured by the lottery selected in the Eckel-Grossman ordered lottery selection task.

Time Preferences: I find that consistently patient respondents are around 7 percentage points more likely to donate the median (maximum in this case) amount in the pro-socialness dictator game. I also find that consistently patient drivers have higher job satisfaction scores, and are less likely to have attended formal driving school. Table 2.4 shows the factors that predict consistently patient time preferences. The outcome variable is an indicator for being consistently patient, with the reference group being respondents with a mixture of consistently impatient, hyperbolic, and 'Patient Now Impatient Later' (PNIL) time preferences.

To further investigate variation based on time preferences, I conduct a series of pair-wise mean comparison tests in Appendix Table B.1. For two of the three experimental measures - risk aversion and pro-socialness - there are statistically significant differences based on time preferences. For risk aversion, respondents with hyperbolic time preferences report the lowest risk aversion scores, with a mean of 2.90 (maximum score of 5). Consistently patient and consistently impatient drivers have similar average risk aversion scores, 3.08 and 3.06 respectively. PNIL drivers have the highest risk aversion scores, an average of 3.26. The difference between the consistently patient and hyperbolic groups is significant at the 10% level, and the difference between the hyperbolic and PNIL groups is statistically significant at the 5% level. In the case of pro-socialness, consistently patient, hyperbolic, and PNIL drivers report similar scores, with 63%, 60.5%, and 59.6% of drivers respectively donating the median amount (the maximum of Rs. 30).

In contrast, consistently impatient drivers donate substantially less, with only 46.3% of drivers donating the maximum (the median) amount. All three pair-wise comparisons of consistently impatient drivers with the other groups are statistically significant. For the dice task, consistently patient, consistently impatient, hyperbolic, and PNIL drivers report fairly similar outcomes, with 17.8%, 17.6%, 18.2%, and 13.3% respectively reporting scores above the 95th percentile of the theoretical distribution. There are no significant differences. Turning to demographic data, I find no differences of age or education. There are some differences on career characteristics including tenure at KSRTC, job satisfaction, driving school history, and accident history, which are detailed in Appendix Table B.1.

2.5 Do Behavioral Measures Predict Job Performance?

I investigate whether behavioral measures predict two job performance metrics: attendance and fuel efficiency. I find no correlations using a broad measure of attendance, the number of shifts an employee is driving. Risk aversion, time preferences, and dishonesty do not predict fuel efficiency, however, there is a negative correlation between pro-socialness and fuel efficiency.

Personality traits have been shown to predict job performance in a number of contexts. For instance, Lagarde and Blaauw (2014) show that pro-socialness measured by the dictator game predicts the take-up of rural postings by nurses, Hanna and Wang (2017) show that cheating on the dice task predicts absenteeism among government nurses, and Callen et al. (2015) show that doctors with higher personality scores attend work more and falsify reports less.

Table 2.5: Factors that Predict Driver Attendance and Fuel Efficiency

	Attendance in Jan2015-Feb2016		Attendance in Mar2016-Feb2017		Mean KMPL in Jan2015-Feb2016	
	(1)	(2)	(3)	(4)	(5)	(6)
Behavioral Traits:						
Risk aversion (score)	0.617	0.658	0.236	0.277	0.00564	0.00362
	(1.421)	(1.409)	(1.298)	(1.291)	(0.00967)	(0.00962)
Amount donated >= median	0.754	1.196	-1.111	-0.890	-0.0769***	-0.0695**
	(4.218)	(4.173)	(3.878)	(3.867)	(0.0289)	(0.0289)
Dice points > 95p	-1.975	-2.537	0.368	-0.0771	0.0445	0.0509
	(5.424)	(5.357)	(4.974)	(4.947)	(0.0382)	(0.0381)
Consistently patient	-0.975	-2.480	-0.280	-1.423	0.0246	0.0300
	(4.165)	(4.132)	(3.823)	(3.818)	(0.0294)	(0.0292)
Demographic Characteristics:						
Age (years)	-0.483	0.813	0.152	0.970*	-0.0112***	-0.00446
	(0.307)	(0.548)	(0.287)	(0.533)	(0.00213)	(0.00377)
Education (years)	-4.807***	-4.907***	-2.735**	-2.685**	-0.0212**	-0.0286***
	(1.394)	(1.453)	(1.271)	(1.324)	(0.00915)	(0.00947)
Career Characteristics:						
KSRTC tenure (years)		-1.664**		-1.057*		-0.00734*
		(0.649)		(0.606)		(0.00417)
Job satisfaction (score)		1.846		1.520		-0.0487***
		(2.225)		(2.029)		(0.0168)
Attended driving school		-34.04***		-25.92***		-0.0223
		(9.968)		(8.287)		(0.0775)
Attended KSRTC training		24.97***		19.18***		-0.116***
		(5.874)		(5.104)		(0.0411)
Has had accident		3.303		2.658		-0.0557*
		(4.234)		(3.899)		(0.0299)
Constant	209.9***	147.5***	133.8***	88.84***	5.329***	5.565***
	(23.04)	(26.54)	(21.64)	(25.83)	(0.160)	(0.189)
Number of drivers	1500	1500	1500	1500	1411	1411
.						
Dependent Variable Mean	143.2	143.2	111.6	111.6	4.669	4.669

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Parentheses contain robust standard errors. Attendance is measured as the number of shifts for the period, Mean KMPL is the simple average of kilometers per liter in all shifts over the period. Risk aversion scores range from 1 (riskiest lottery selected) to 5 (safest lottery selected).'Donated >= median' indicates the driver donated the median amount, Rs.30 (the maximum possible), in the pro-socialness dictator game. Dice points > 95p indicates that the respondent's total dice points was above the 95th percentile respectively of the theoretical distribution. Consistently patient indicates the reported time preferences were consistently patient (see Table 2.2 for details.) Job satisfaction scores range from 1 (very dissatisfied) to 5 (very satisfied).

I use a broad measure of attendance, the number of shifts an employee works as a driver in that branch. With this metric, I find no correlations between behavioral characteristics and attendance. Table 2.5 analyzes which factors predict driver attendance. I consider two outcome variables: attendance between January 2015 and February 2016, before the field experiment started, and attendance between March 2016 and February 2017 in the post-intervention period. Attendance is defined as the driver being present at work for a shift, in that branch, and performing work as a driver. This is a broad measure of attendance, and drivers could show lack of attendance for a number of reasons, which are a mixture of absenteeism, authorized leave, and other short term factors, as well as long term or permanent factors such as resignations and retirements.

I also analyze fuel efficiency for the period January 2015 to February 2016, before the field experiment started, as shown in Table 2.5. The outcome variable is the simple average of shift-level fuel efficiency for the driver's shifts, measured in kilometers per liter (KMPL). Risk aversion, time preferences, and dishonesty do not predict performance on the fuel efficiency metric. However, pro-social drivers have a lower average fuel efficiency in the pre-intervention period. It is not entirely clear why this relationship exists. It is possible that pro-social drivers are assigned to, or request, a different type of route, such as driving in cities rather than highways. City driving is less fuel efficient. It is also possible that pro-social drivers drive differently, and are more willing to sacrifice fuel efficiency for passenger comfort - for instance, they might keep air conditioning on for longer, which is better for passengers but lowers fuel efficiency.

2.6 Do Behavioral Measures Predict Responses to Job Performance Interventions?

I analyze whether either the training intervention or the financial incentives intervention have heterogeneous treatment effects on fuel efficiency based on the bus driver's baseline measures of risk aversion, dishonesty, time preferences, and pro-socialness. The results are presented in Table 2.6. I find no evidence of any heterogeneous treatment effects. In Chapter 1 of this dissertation I present the main effects, which are summarized in Section 2.2.

Table 2.6: Treatment Effect Heterogeneity Analysis by Measures of Behavioral Traits

	Dice > 95p		Donated >= Median		Risk Aversion (score)		Consistently Patient	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post × Char	-0.0155	0.0106	0.0210	0.0185	0.01000	0.00657	-0.00690	0.00568
	(0.0305)	(0.0160)	(0.0238)	(0.0123)	(0.00852)	(0.00421)	(0.0241)	(0.0123)
Training × Post	0.00697	0.0135	0.0141	0.00765	0.0601	0.0324	0.00846	0.0347**
	(0.0205)	(0.0111)	(0.0296)	(0.0148)	(0.0426)	(0.0250)	(0.0299)	(0.0162)
Incentives × Post	0.00902	0.0196**	0.0283	0.0184	0.0633*	0.0290	0.00186	0.0127
	(0.0178)	(0.00908)	(0.0250)	(0.0128)	(0.0381)	(0.0188)	(0.0241)	(0.0130)
Training × Incentives × Post	0.0189	-0.00864	-0.00670	0.00369	-0.0333	0.0213	0.00923	-0.0328
	(0.0277)	(0.0150)	(0.0412)	(0.0205)	(0.0588)	(0.0336)	(0.0388)	(0.0213)
Training × Post × Char	0.0115	-0.0208	-0.0100	0.00227	-0.0170	-0.00769	-0.000817	-0.0430**
	(0.0533)	(0.0280)	(0.0386)	(0.0200)	(0.0132)	(0.00691)	(0.0386)	(0.0208)
Incentives × Post × Char	0.0624	-0.0121	-0.0162	-0.00368	-0.0146	-0.00403	0.0324	0.00413
	(0.0448)	(0.0228)	(0.0330)	(0.0170)	(0.0112)	(0.00568)	(0.0326)	(0.0169)
Training × Incentives × Post × Char	-0.00117	0.0505	0.0408	-0.00372	0.0168	-0.00636	0.0127	0.0595**
	(0.0706)	(0.0361)	(0.0525)	(0.0272)	(0.0176)	(0.00939)	(0.0515)	(0.0277)
Observations (shifts)	381437	378578	383965	381100	384170	381303	381128	378270
Number of clusters (drivers)	1423	1423	1431	1431	1431	1431	1422	1422
Controls		Yes		Yes		Yes		Yes
.								
Characteristic Mean	0.175	0.175	0.604	0.604	3.055	3.055	0.579	0.579

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Parentheses contain robust standard errors clustered by driver. See Table 2.1 notes for variable descriptions. See Chapter 1 of this dissertation for more details on the heterogeneity analysis. KMPL is winsorized at 1% and 99%. Post indicates shifts after March 1 2016. Char is the driver characteristic specified in the column title. The characteristic mean is in the post-attrition driver subsample for analysis. All specifications include driver and month × year fixed effects. Controls in even-numbered columns are depot × vehicle, depot × route, and day of week fixed effects.

A substantial and growing literature demonstrates that interventions to improve job performance can have heterogeneous effects based on the worker's personality traits. For instance, Ashraf, Bandiera, and Jack (2014) find that both financial and non-financial incentives for promoting public health were more effective among pro-socially motivated Zambian agents. In another study in Pakistan, Callen et al. (2015) find that health inspectors who had higher scores on two personality measures - the Big Five and Perry Public Sector Motivation scores- had a larger treatment response to a randomized intervention that increased monitoring. In Indonesia, Banerjee et al. (2016) conduct a randomized field experiment that reforms the implementation of a subsidized rice program. They use the same dice task I used in Karnataka, and find that individuals who show higher levels of dishonesty on the dice task are more likely to block the reforms and protect the rent-seeking present in the status quo.

Treatment effects did not vary by pro-socialness in this study. In many settings, financial incentives have caused perverse effects among the pro-socially motivated (Bénabou and Tirole 2006), though as Finan, Olken, and Pande (2017) point out, it is not clear if this is an important issue in the context of governments. In fact, Ashraf, Bandiera, and Jack (2014) find that the effect of both financial and non-financial incentives are greater for pro-socially motivated agents. They recruited hairdressers and barbers in Zambia to provide information on HIV prevention and to sell condoms. They implement a randomized controlled trial evaluating small financial incentives, large financial incentives, and non-financial incentives to sell condoms. Pro-socially motivated participants, who donated more to an HIV charity in a baseline dictator game, respond more strongly to both financial and non-financial rewards.

I do not find any evidence that there were heterogeneous responses to the interventions based on cheating on the dice task. I evaluate if the training program was less effective among dishonest workers due to shirking, and if financial incentives were more effective among drivers who cheated on the dice task, who may have a large marginal value for money. As Hanna and Wang (2017) show, cheating on the dice task is correlated with absenteeism among government nurses in Karnataka, where they note that an absentee worker is fraudulently collecting a salary for a day not worked. There could be similar shirking in the training program, if drivers treated it as a form of paid leisure and ignored the work purposes. The training program involved two or three days off regular work. The training schedule had fewer hours than a typical shift. Often, the training program was held out of station, with lodging and food provision for the participating drivers. If drivers treated it as paid leisure, it would be a type of shirking, and if it did occur, it was plausibly more common among drivers who cheated on the dice task. I also study whether dishonest individuals may have disproportionately strong responses to financial incentives. Cheating on the dice task implies the respondent prioritized the money available from cheating over the intrinsic benefits of staying honest. The dice task thus reveals people who are relatively dishonest. However, cheating on the dice task may also represent a very high value on money.

I hypothesize that the training intervention might have a larger impact on risk-averse drivers, but do not find any evidence that this is the case. I expect risk averse drivers to place a greater value on personal safety. Risk averse drivers who are taught safety techniques in the SAFED training program may thus be more intrinsically motivated and demonstrate a larger treatment response when given training.

I find some evidence of heterogeneous treatment responses based on time preferences, but it is inconsistent and specification-sensitive. I hypothesize that consistently patient drivers face fewer intrinsic costs if they have to drive at moderate speeds. Safe and fuel efficient driving inherently involves driving at moderate speeds and not speeding. Impatient drivers, in contrast, may face intangible costs if they are constrained from driving fast. Table 2.6 compares the treatment effect for consistently patient drivers, relative to drivers in the consistently impatient, patient now-impatient later, and hyperbolic categories. The specification without controls has no statistically significant differences. The specification with controls does find that training was less effective in the consistently patient subgroup, and that the interaction effect was more effective. This result is not robust, given that the estimates differ dramatically with and without controls. In Table 1.7 I implement an additional subgroup analysis by time preference, for the specification without controls. None of the treatment effects are statistically significant in subgroups and the results overall do not suggest significant heterogeneity by time preference categories.

In Appendix Table B.2, I verify that the randomized interventions do not predict attendance in the pre-intervention period, had no overall treatment effect on attendance in the post-intervention period, and had no heterogeneous treatment effects on attendance in the post-intervention period.

2.7 Do Behavioral Measures Predict Human Resource Decisions?

I find that personality traits predict a number of human resource decisions. Dishonest drivers are more likely to be transferred to other branches or reassigned to other duties, pro-social drivers are less likely to be on authorized leave, and risk-averse drivers are more likely to be in attendance, less likely to be disciplined or absent, and more likely to be on authorized leave.

I collate a dataset of human resource decisions for 761 drivers using the 2016 follow-up survey of 761 drivers, detailed in Section 2.3. I build this dataset from a detailed metric of attendance which incorporates departures from the driver workforce. Drivers were tracked for three months each, and are categorized as present for all three months or missing for at least one of the three months. Drivers who are missing are classified as dismissed or suspended, i.e. disciplined in some way; absent, indicating unauthorized absenteeism; on temporary authorized leave due to short term disability, holiday, or leave; permanently departed from KSRTC due to death or retirement; transferred to other branches of KSRTC; reassigned to other duties in the same branch, such as being a bus conductor; and attrition, indicating we could not find the driver or find out why he was missing from the workplace. I combine these into 5 categories: Present, Disciplined/Absent (dismissed, suspended, or absent), Authorized Leave (disability, holiday, leave, death, retirement), Transferred/Reassigned, and Attrition.

Table 2.7 evaluates whether there are differences in types of driver attendance/non-attendance based on behavioral traits. For each behavioral trait, I present the means for the five categories described above. I also present the differences from pair-wise mean comparison tests (ttests). Panel A analyzes heterogeneity based on dishonesty and pro-socialness. Panel B analyzes heterogeneity based on risk aversion and time preferences.

Table 2.7: Heterogeneity Analysis of Causes of Driver Absence

	Means		Difference	Means		Difference
	(1) No	(2) Yes	(3) Diff	(4) No	(5) Yes	(6) Diff
Panel A: Dishonesty and Pro-socialness:						
	Dice points>95p			Donated >= median		
Present	0.826	0.788	0.0377	0.801	0.830	-0.0287
	(0.379)	(0.410)	(0.0364)	(0.400)	(0.376)	(0.0286)
Disciplined/Absent	0.0293	0.0219	0.00737	0.0232	0.0305	-0.00732
	(0.169)	(0.147)	(0.0156)	(0.151)	(0.172)	(0.0122)
Authorized Leave	0.0163	0	0.0163	0.0232	0.00654	0.0166**
	(0.127)	(0)	(0.0108)	(0.151)	(0.0807)	(0.00843)
Transferred/Reassigned	0.0959	0.153	-0.0573**	0.109	0.105	0.00470
	(0.295)	(0.362)	(0.0291)	(0.312)	(0.306)	(0.0229)
Attrition/Unknown	0.0325	0.0365	-0.00398	0.0430	0.0283	0.0147
	(0.178)	(0.188)	(0.0170)	(0.203)	(0.166)	(0.0135)
Number of drivers	615	137	752	302	459	761
Panel B: Risk Aversion and Time Preferences:						
	Risk-averse			Consistently Patient		
Present	0.798	0.849	-0.0514*	0.816	0.819	-0.00277
	(0.402)	(0.358)	(0.0285)	(0.388)	(0.385)	(0.0285)
Disciplined/Absent	0.0374	0.0131	0.0242**	0.0253	0.0297	-0.00443
	(0.190)	(0.114)	(0.0121)	(0.157)	(0.170)	(0.0122)
Authorized Leave	0.00659	0.0230	-0.0164*	0.0127	0.0137	-0.00107
	(0.0810)	(0.150)	(0.00842)	(0.112)	(0.117)	(0.00846)
Transferred/Reassigned	0.116	0.0918	0.0247	0.120	0.0961	0.0241
	(0.321)	(0.289)	(0.0228)	(0.326)	(0.295)	(0.0228)
Attrition/Unknown	0.0418	0.0230	0.0188	0.0253	0.0412	-0.0159
	(0.200)	(0.150)	(0.0135)	(0.157)	(0.199)	(0.0135)
Number of Drivers	455	305	760	316	437	753

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Risk-averse indicates the two highest risk aversion scores. See Table 2.1 notes for remaining variable descriptions. For each variable, the first and second columns report means with standard deviations in parenthesis, and the third column reports differences from pair-wise mean comparison tests (ttests) with standard errors in parenthesis. 'Present' - continued employment about 9 months after the survey. 'Disciplined/Absent' - driver is either dismissed, suspended, or absent from work. 'Authorized Leave' - death, accident-related disability, retirement, or authorized leave/holiday. 'Transferred/Reassigned' - driver was transferred to another branch or reassigned to alternate duties. 'Attrition/Unknown' indicates updated information is missing.

Personality traits are related to a number of statistically significant differences in why drivers are absent from the workplace. Drivers who are dishonest on the dice task are 5.73 percentage points more likely to be transferred to another branch or reassigned to alternate duties in the same branch. Among honest drivers, 9.59% are transferred or reassigned, while 15.3% of the dishonest drivers are, and the difference is statistically significant. Since many of the reassignments are to conductor duty, and drivers are thus removed from driving duties, it is possibly a demotion in many cases. However, it is not clear whether conductor duty is preferred to driver duty. In the case of transfers to other branches, whether this is a type of reward or punishment will be location-specific. Pro-social drivers are less likely to be on authorized leave. Among less pro-social drivers, 2.32% are on authorized leave during the follow-up survey, while 0.65% of the pro-social drivers are, leading to a statistically significant difference of 1.66 percentage points. Risk-averse drivers are 5.14 percentage points more likely to be present in the workplace. Among less risk-averse drivers, 79.8% are present, while 84.9% of the risk-averse drivers are present, and this difference is significant at the 10% level. Risk-averse drivers are 2.42 percentage points less likely to be either disciplined or absent, which is statistically significant. Among less risk-averse drivers, 3.74% are disciplined or absent, compared to 1.3% of the risk-averse drivers. Risk-averse drivers are also more 1.64 percentage points more likely to be on authorized leave, significant at the 10% level. Among less risk-averse drivers, 0.66% are on authorized leave, compared to 2.3% of the risk-averse drivers. I find no relationship between time preferences and drivers being absent from the workplace.

2.8 Conclusion

Overall, personality traits do not appear to affect job performance among KSRTC bus drivers, with one exception. I do not find that either a training program intervention or a financial incentives scheme has heterogeneous treatment effects based on personality traits, nor do I find that personality traits predict a broad measure of driver attendance. Risk aversion, time preferences, and dishonesty do not predict fuel efficiency. However, pro-socialness is negatively correlated with fuel efficiency.

My results suggest this is a consequence of KSRTC's effective use of recruitment, hiring, transfers, and other human resource policies. I find that dishonest drivers are more likely to be transferred or reassigned, and risk-tolerant drivers are more likely to be dismissed, suspended, or absent from work, indicating that KSRTC actively uses human resources policies to retain or reject drivers. This bus driver population is more pro-social and more honest than most other populations studied in the literature, which is likely to be a result of selection and retention policies. Among retained drivers, who remain in the driver workforce at the time of the 2015 survey using experimental measures of behavior, personality traits do not predict job performance. These are the drivers who have made it through the hiring process, the two year probation period, and have not been reassigned to non-driver duties, suspended, or otherwise removed.

Despite being a government bureaucracy with constraints on promotion, salaries, and dismissals, KSRTC is able to reward or punish employees using alternate tools of suspensions and transfers. This is congruent with findings from other bureaucracies in India. For instance, Banerjee et al. (2012) show how police reforms in Rajasthan, India, successfully used transfers to sought-after postings as an incentive, given that financial incentives were infeasible. Dhaliwal and Hanna (2017) show that location postings and liberal implicit permission for absenteeism are used to recruit and retain doctors in Karnataka's public health system, which struggles to fill vacancies due to relatively low salaries.

I hypothesize that the high wage premium KSRTC drivers receive leads to a large applicant pool and allows management to be selective on multiple dimensions during recruitment, selection, and retention. When there are multiple applicants for a post, the organization need not make tradeoffs between multiple dimensions of candidate quality, as shown by Ashraf, Bandiera, and Lee (2018). They conduct a field experiment among community health workers in Zambia, and show that recruitment posters that highlight career advances, rather than community benefits, on average attract applicants with lower pro-socialness and higher talent. Though the treatment changes the applicant pool, and there is a tradeoff between talent and pro-socialness on average, candidates are not hired randomly. Rather, the best candidates are hired in both the treatment and control groups. As a result, the treatment recruits are more talented and equally pro-social (Ashraf, Bandiera, and Lee 2018).

KSRTC drivers have a total remuneration ranging from Rs. 30,000-50,000 per month, and new entrants get Rs. 20,000-25,000 a month. In comparison, the salary range for privately employed car drivers in Bangalore is Rs. 10,000-20,000. It is common for the public sector to pay a premium over the private sector, especially in the developing world (Finan, Olken, and Pande 2017). However, the wage premium at KSRTC is particularly stark, especially given that the job only requires an education up to 10th grade. Other jobs, such as nurses, doctors, or teachers, may also have a premium, but these jobs have far more stringent educational requirements. With such a large wage premium, KSRTC should find it easy to fill job vacancies and can be highly selective on multiple aspects of job performance.

Chapter 3

Trapping Dust: How Much Would Pollution Control Cost Indian Industrial Plants?[1]

3.1 Introduction

Environmental pollution, air or water, has extraordinary public health costs. Ambient fine particulate matter - $PM_{2.5}$ - air pollution is a particular health hazard and is one of the leading causes of death globally (HEI 2017). India's particulate matter pollution is one of the worst among the world's populous countries, with a population-weighted annual average $PM_{2.5}$ concentration of 74.3 microgram/m3 in 2015 (HEI 2017; GBD MAPS 2018). With fine particulate pollution, an increase in concentration of 10 microgram/m3 reduces life expectancy by about a year on average (Chen et al. 2013, as calculated in Greenstone et al. 2015). This implies that improving India's population-weighted $PM_{2.5}$ concentration to the WHO standard of 10 microgram/m3 (WHO 2006) would increase average life expectancy by 6.4 years.

Thus a rich literature indicates that increasing the stringency of environmental pollution regulations would have substantial health benefits in highly polluted middle income countries like India. However, there are few estimates of how much regulation would increase pollution

1. Co-authored with Santosh Harish

abatement (reduction) costs for polluting sectors, such as industry, transport, and power generation, and there are few cost-benefit analyses of alternate pollution control policies. Middle income countries often have limited capacity to bear pollution abatement costs, and it is thus important to design cost-effective environmental regulation. Cost-effective regulation makes abatement cheaper for polluters. It can also enable greater overall abatement, as regulators can set more stringent pollution reduction goals when abatement costs are lower. Theory and experience from developed countries suggest that market-based instruments, such as cap and trade or emissions taxes, reduce abatement costs compared to command and control regulations like performance standards and technology mandates.

In India, the majority of existing particulate matter regulations are command and control ones, so a shift to market based instruments could significantly reduce abatement costs for the polluters. On the other hand, there are political, legal, and administrative capacity barriers to shifting to more cost-effective policies. Reforms are not always feasible, or are difficult to implement. Moreover, policies that minimize abatement costs may not minimize monitoring and enforcement costs for the regulator (Malik 1992), which is a particular concern in developing countries. The status quo monitoring and enforcement environment in India is weak, and industrial plants routinely violate environmental regulations. In such an environment, the regulator's efforts may be better directed to reforms around monitoring and enforcement. To evaluate if the regulator should devote resources to regulatory reforms, it is thus valuable to have quantitative estimates of how reforms would change pollution abatement, abatement costs borne by the polluter, the distribution of costs, and monitoring and enforcement costs borne by the regulator.

In this paper, we focus on Indian industrial plants, which contribute 7.7% percent of India's $PM_{2.5}$ pollution (GBD MAPS 2018). We use a rich plant-level dataset on particulate matter emissions, air pollution control devices, and operating practices from 311 industrial plants in the Surat region of the state of Gujarat. We conduct a descriptive analysis of the status quo pollution control at these industrial plants. We then empirically estimate particulate matter abatement costs for these plants using a bottom-up engineering model and detailed plant-level data. We pair the engineering model with an economic framework in which plants minimize abatement costs subject to regulatory constraints. Using the engineering model and economic framework, we estimate plant-level and aggregate outcomes on costs and abatement under several pollution control policies. We analyze the costs of compliance to polluters, and the pollution emissions, under status quo and alternate policies. We also evaluate the relative cost-effectiveness and distributional impact of alternate policies, and how compliance costs, abatement, and net benefits change as regulation becomes more stringent. Although it would be tremendously valuable to also estimate the monitoring and enforcement costs borne by the regulator under alternate policies, we are not able to do so in this paper.

Under existing regulation, Indian industrial plants have a limit on the concentration of pollutants that can be emitted from the plant (in mg/Nm^3), a policy known as an emissions concentration standard, as well as a technology mandate. In industrial plants, emissions are generated as a byproduct of combustion, partially captured through an air pollution control device system, and then released to the ambient air through a stack (chimney). India sets a limit on the pollution concentration that can be emitted through the stack. In Surat it is currently set at 150 mg/Nm3 of total suspended particulate matter, which is a combination of fine particulates such as $PM_{2.5}$ and larger particulates. Surat industrial plants also face an technology standard that requires plants to install air pollution control devices before beginning operations.

First, using descriptive analysis, we find that nearly all plants in our survey have at least two air pollution control devices per combustion source. Nonetheless, particulate matter emission concentrations remain high and usually above the regulatory standard. The presence of air pollution control devices together with high particulate matter emissions implies the technology mandate is well enforced but the emissions concentration standard is not. We note that despite weak enforcement, many plants do comply with the emissions concentration standard.

We next use our detailed dataset to build a bottom-up engineering model that combines survey data and expert engineering estimates on air pollution control technology. We estimate how much emissions each plant generates, how much emissions the plant currently captures given its air pollution control devices and operating practices, and how much emissions are released into the ambient air through the stack. We then make air pollution control technology and operating practices into optimization variables. For each plant, we estimate what total abatement costs and total pollution abatement would be under possible alternate combinations of air pollution control technology and operating practices. We use these estimates to build an abatement cost function and an emission function, each with two optimization variables: the air pollution control technology selected, and how often to operate the air pollution control system.

We pair the engineering model with a simple economic framework in which each plant selects abatement technologies and operating practices in order to minimize expected abatement costs, subject to regulatory constraints. This is a static analysis which assumes each plant is compliant, risk-neutral, and prioritizes cost-minimization. Based on this framework, we use the abatement cost and emissions functions to set up constrained optimization cost minimization problems for each plant stack under four potential pollution control policies. We evaluate the existing emissions concentration standard, an alternative quantity based performance standard called an emissions load standard, a cap and trade program, and an emissions tax. We also consider several potential regulation stringencies, from 200 mg/Nm3 (a loosening of existing standards) to 50 mg/Nm3. Assuming each firm cost-minimizes subject to constraints, we estimate outcomes under the different constraints and incentives posed by different pollution control policies.

Using the framework, we first estimate the changes in abatement costs and pollution reduction if the status quo emissions concentration standard of 150 mg/Nm3 was rigorously enforced and all plants were in compliance. Fully enforcing the current regulation would lead to a 66% reduction in average annual emissions, at a average abatement cost of Rs. 36,150 per year ($556), though there is significant variation in costs across plants. Abatement costs are very modest, as our industries report an average annual turnover of Rs. 502,000,000 ($7.72 million) and an average annual electricity bill of Rs. 28,000,000 ($430,000). Since the technology mandate is well enforced, plants already have the technology required to reduce pollution. Thus the costs of complying with existing regulations are modest, with relatively little capital costs and installation of new equipment. Instead, plants upgrade their existing technology and improve operations and maintenance.

Next, we compare the status quo to alternate policies and alternate regulation stringencies. For any fixed quantity of pollution reduction, both market based instruments - a cap and trade program and an emissions tax - reduce abatement costs substantially, while the alternate command and control regulation we evaluate, the emissions load standard, increases abatement costs relative to an emissions concentration standard. Aggregate emissions are significantly less than baseline emissions under every scenario we evaluate. Strictly enforced emissions concentration standards lead to significant over-control. For all scenarios we evaluate, plants are more likely to retrofit existing equipment than purchase new equipment. This again highlights that all industries already have pollution control equipment installed, which are being poorly maintained and operated at baseline.

We also look at relative cost-effectiveness, that is, the costs of different policies for a fixed quantity of emissions reduction, and also analyze the factors that lead to cost differences between policies. Strictly enforcing the 150 mg/Nm3 emissions concentration standard would have an average annual abatement cost of Rs. 1.31 per kilogram of particulates abated, for a total 66% reduction in baseline emissions. For the same 66% reduction, switching to either market based instrument would lead to a 39% reduction in abatement costs, while switching to an emissions load standard would increase costs by 21%. We present cost comparisons for several different levels of emissions reduction in Section 3.6.2. Under an emissions concentration standard, average annual abatement costs per kilogram rise from Rs. 1.31 for a 66% reduction in baseline emissions, to Rs. 1.80 to achieve a 84% reduction in emissions. For market based instruments, average costs rise from Rs. 0.80 for a 66% reduction to Rs. 1.31 for a 84% reduction. We demonstrate what factors drive these cost differences by providing a case study of a representative industrial plant.

We then look at the distributive impact of regulatory reform. Each of the three alternate policies, an emissions load standard, cap and trade, or an emissions tax, is based on emissions quantity rather than emissions concentrations. As a result, reform consistently disadvantages plants that are compliant at baseline, since these plants are already optimized for the existing emissions concentration standard, and often incur some costs from a change to a quantity based policy. Other distributive impacts vary based on the details of policy reform.

We turn to the cost-benefit analysis for different regulation stringencies. In every scenario we analyze, the health benefits of abatement substantially exceed abatement costs, and the emissions tax is cheaper than an emissions concentration standard. We estimate the benefits of improved life expectancy from particulate matter abatement in Surat, and compare these to our estimated costs. Our central estimates, holding total baseline emissions reduction fixed at 66%, are that the net benefits under a concentration standard are Rs. 580-583 billion, and the net benefits under an emissions tax are Rs. 584-586 billion. We also present net benefit estimates for different levels of emissions reduction. As expected, abatement costs increase significantly as the regulation stringency increases, however, the total monetized benefits increase even more. Thus, for the range we consider, net benefits increase with increased stringency. While we cannot identify the optimum stringency, we find it should be equivalent to at least an 84% reduction in baseline emissions. We present benefit to cost ratios as well. Our estimates for the emissions concentration standard are within the ratios for the original Clean Air Act, which also relied heavily on command and control regulation.

Given the relatively low abatement costs, it is apparent that the key barriers to particulate matter abatement in Indian industrial plants are not the abatement cost to the polluter, but rather the regulator's constraints regarding bureaucratic capacity, monitoring, and enforcement. We discuss ongoing reforms to monitoring and potential reforms to enforcement. We conclude that since any reforms to enforcement will involve new legislation and a political process, a shift to a market based instrument should be done simultaneously as part of the reform process.

This paper contributes new, detailed empirical estimates of particulate matter abatement costs in small and medium industrial plants using granular plant level data. To our knowledge, this is the first estimate of particulate matter abatement costs for the small and medium enterprises that form the bulk of India's industrial sector. While a rich literature demonstrates large health benefits to reducing particulate matter pollution in India (HEI 2017; GBD MAPS 2018; Bank 2013; S. K. Guttikunda and Puja Jawahar 2014), and numerous papers and reports evaluate the success of India's environmental policies and environmental reforms (Duflo et al. 2013, 2014; Greenstone and Hanna 2014; MoEFCC 2014), there are relatively few estimates of abatement costs. Much of the existing literature focuses on abatement costs at thermal power plants, or on macroeconomic sector-level costs (Cropper et al. 2017; Gupta 2002).

We highlight several new contributions to the existing literature. First, we demonstrate that although most industrial plants are not compliant with the existing emissions concentration standards, they do indeed have the necessary pollution control technologies installed, due to successful implementation of an informal technology mandate. However, these pollution control devices are improperly maintained and operated, and thus perform very poorly. Second, we show that if the current regulations were strictly enforced, abatement costs would be modest for most plants, in part because they only need to improve existing pollution control equipment. Third, we show that there is substantial heterogeneity in plant-level emissions and abatement costs, even though the industrial plants in our sample are very homogeneous. Due to this heterogeneity of abatement costs, switching from the existing command and control regulation to market based instruments reduces costs substantially, by 39% in our main estimate. Due to technology characteristics, the existing emissions concentration standard outperforms the alternate command and control regulation of an emissions load standard. Fourth, we analyze distributional effects, and show that while plants that are compliant at baseline face the least absolute increase in costs, they are relatively disadvantaged by a shift to a quantity based policy since they are already optimized for a concentration policy. Thus, the distributional consequences of a policy change fall primarily on plants that are compliant at baseline. Finally, we present net benefit estimates for a variety of policy scenarios.

This dissertation chapter is structured as follows. Section 3.2 provides background on industrial particulate matter pollution in India, how industrial plants generate emissions, what the status quo pollution control policy is, possible alternatives, the history of environmental regulation in India, and the current status of monitoring and enforcement. Section 3.3 provides details on the data, including industrial plant survey data and expert inputs on pollution control technology. Section 3.4 describes descriptive results about the industrial plants. Section 3.5 describes the features of the engineering model, the main functions and cost minimization problems used in the model, and how the data is used. Section 3.6 presents results on the costs of enforcing status quo regulations, outcomes under alternate policies, the factors driving cost differences among alternate policies, and the distributive impact of alternate policies. Section 3.7 provides a cost-benefit analysis for pollution abatement in our setting. Section 3.8 contains a discussion and conclusion.

3.2 Background

3.2.1 Industrial Particulate Matter Pollution in India, Emitted From Stacks

Industrial emissions contribute 7.7% percent of India's fine particulate matter pollution (GBD MAPS 2018). Emissions are generated in industrial plants, partially captured by air pollution control systems, and then released to the atmosphere through a stack.

Much of India's fine particulate matter pollution comes from the industrial sector. At a national level, the Global Burden of Disease Maps Project estimates that industrial sector coal burning causes 7.7% of ambient $PM_{2.5}$ pollution in India, and is one of the major sources of PM pollution alongside residential biomass burning, coal use in power plants, anthropogenic dust, and other sources (GBD MAPS 2018). The industrial sector's contribution to particulate pollution varies by region. Industrial emissions contribute to both primary particulate matter, which is directly emitted, and to secondary particulates, which are produced by chemical reaction in the atmosphere from precursor pollutants such as oxides of nitrogen and sulphur.

Industrial emissions are a byproduct of combustion. Plants burn fuel, such as coal, in combustion sources, such as boilers, to generate energy. The heat generated is used to convert water to steam, which is then transported across the plant to power various processes and equipment. Meanwhile, combustion generates pollutant byproducts including carbon dioxide, oxides of sulphur, carbon monoxide, and particulate matter (soot), which leave the boiler as a cocktail called 'flue gas'. The flue gas is passed through a series of air pollution control devices (APCDs) that together form an APCD system. Air pollution control equipment can reduce particulate matter emissions, though not necessarily gaseous pollutants such as carbon dioxide. Common APCDs include cyclones, scrubbers, bag filters and electrostatic precipitators, which use different methods to remove the particulates from the flue gas. The resulting cleaner flue gas passes through the stack outlet to the ambient air. Industrial plants can have two or more parallel chains attached to each stack, where a parallel chain consists of a combustion source followed by air pollution control devices (APCDs). The final emissions from the stack is a function of several factors, including the amount of fuel burned, the efficiency of the combustion equipment, and the efficiency of the APCD system.

3.2.2 Status Quo and Alternate Pollution Control Policies

Pollution control policies primarily consist of two categories: command-and-control policies, such as performance standards and technology mandates, and market based instruments, which include cap and trade and emission taxes. Market based instruments are expected to minimize abatement costs. In Surat, the status quo pollution control policy is a emissions concentration-based performance standard. The emissions concentration standard mandates that polluted air exiting the industrial plant's stack must have a total suspended particulate matter concentration less than 150 mg/Nm3. Surat industrial plants also face a informal technology standard that mandates air pollution control devices (APCDs) be installed at the plants.

Status Quo and Command and Control Regulations

For industrial emissions, the State Pollution Control Boards in India (SPCBs) set a permissible limit on the concentration of pollutants that can be emitted from the stacks (i.e. the chimneys) of each industrial plant. In Surat, for plants for the size we study, the emissions concentration standard for total suspended particulate matter (SPM) is 150 mg/Nm3. Surat industrial plants also face a informal technology standard that mandates Air Pollution Control Devices (APCDs) be installed at the plants, without which the Gujarat State Pollution Control Board does not give the industrial plant 'consent' to start operations.

An alternate performance standard could set limits on the pollutant load from a plant over a period of time - we refer to such a quantity-based policy as an emissions load standard. The existing pollution norms are in terms of concentrations: the mass of pollutants in a unit volume of air leaving a stack. As a result, industries are regulated based on their instantaneous emissions and not their pollutant load over the course of a stipulated duration. This is worth noting because industrial emissions can vary substantially over the course of the day and year depending on industrial operations and whether and the extent to which the installed air pollution control equipment are run.

Market Based Instruments

The choice of policy instrument, such as command-and-control regulations (CACs) or market based instruments (MBIs) affects several outcomes, including the level of pollution and industry profits and productivity (Helfand 1991). Theory suggests that shifting from CACs to MBIs can significantly reduce abatement costs. However, as performance and technology standards often lead to "over control" beyond the standards, CACs may have greater benefits than MBIs, so the net benefits must be compared (Oates, Portney, and McGartland 1989).

The primary motivation for market-based instruments is that they minimize the abatement costs of attaining any given level of emissions. Appropriately designed market-based instruments are expected to minimize overall abatement costs by equalizing the marginal cost of abatement from each plant (Stavins 2003). Market-based instruments are likely to have greater benefits compared to command-and-control regulations as the heterogeneity of abatement costs between polluters grows larger and as the degree of mixing of the pollutant increases (Stavins 2003) [2]. The relative advantages of quantity-based instruments (like emissions trading) versus price-based instruments (like taxes) depend on the polluter's cost and benefit functions and the degree of uncertainty about each (Weitzman 1974; Stavins 2003), and on the relative responsiveness to change of each instrument (Stavins 2003).

Empirical evidence supports the theoretical expectation that market based instruments reduce costs. In evaluating the US SOx markets of 1995 (expanded in 2000), Carlson et al. (2000) estimate savings of 45-55% compared to a uniform standard regulating emissions rates. Burtraw et al. (1998) and Muller and Mendelsohn (2009) estimate that the improvements in public health and reduced acidification from these markets outweigh the costs by an order of magnitude. Fowlie, Holland, and Mansur (2012) find that emissions from firms under the RECLAIM NOx market were on average 24 percent lower than those regulated under nearby firms subject to command and control.

An additional advantage of economic instruments such as trading is that polluters have dynamic incentives to continue abating their emissions and innovate cleaner equipment and

2. Externalities are spatially differentiated and in theory every emitting source might impose different social costs from pollution (Muller and Mendelsohn 2009).

processes (Jaffe and Stavins 1995). Absolute emission norms, like emissions concentration standards in India, do not provide any incentive for industrial plants to reduce emissions beyond the minimum they are required to, even if the marginal costs of additional abatement are negligibly small in comparison with the externalities they impose.

However, policies that minimize abatement costs may not minimize enforcement costs for the regulator (Malik 1992). Monitoring and enforcement costs could be lower for market based instruments (MBIs) if participants in a trading scheme are incentivized to report their emissions (Stavins 2003). On the other hand, monitoring and enforcement costs may be higher for MBIs, particularly in developing countries (Blackman and Harrington 2000). In developing countries in general, institutional readiness is a potential barrier for trading to be as efficient and cost-effective as it could be in theory. Coria and Sterner (2008) review the lessons from the trading program in Santiago launched in 1997 (the first application of emissions trading outside the OECD countries) and conclude that Chile's experience demonstrates that a middle income country is capable of implementing a trading scheme, but caution that "the challenges of designing successful environmental programs in less developed countries should not be underestimated".

3.2.3 History of Successes and Failures in Indian Environmental Regulation

Most pollution control policies in India are command-and-control (CAC) regulations, which have often had success at controlling pollution generated, compared to a counterfactual in which these policies were not enacted. There have been nascent efforts at market based instruments as well. However, the absolute quantity of pollution generated from power plants, industry, transport, and other sectors has grown so fast that pollution control policies have been inadequate, leading to ever worsening air quality in India.

Environmental regulation in India remains reliant on command and control policies: emissions standards, technology mandates, and bans on operations and processes, which have often been successful. These regulations date back to Air and Water Acts from the 1970s, which themselves were based on the original Clean Air Act in the US. Past successes include the mandated installation of catalytic converters, an end-of-pipe technology to reduce emissions from vehicular exhaust. Greenstone and Hanna (2014) estimate that the catalytic converter policy resulted in declines of 19 percent, 69 percent, and 17 percent of the 1987–1990 nationwide mean concentrations, for PM, SO2 and NO2 respectively. Narain and Krupnick (2007) find significant decreases in both fine particulate emissions and sulphates after a court mandated that all public transportation in New Delhi use compressed natural gas (CNG) and stop using diesel fuel. Another successful technology mandate was through Supreme Court Action Plans, which required pollution control equipment in industrial plants (Harrison et al. 2015).

India has rarely used markets as a means of environmental regulation, with the Renewable Energy Certificates and the Bureau of Energy Efficiency's Perform, Achieve and Trade scheme being notable exceptions. These schemes were introduced by India's Ministry of Power, and there exist no similar examples in the sphere of environmental regulation. India does have a tax on coal that was raised to INR 400 per tonne in the 2016-17 budget.

3.3 Data

Our paper uses two main data sources. First, we use expert inputs on air pollution control technologies, their costs, and their pollution control efficiency. Second, we use a unique plant-level survey on emission sources, emission levels, and air pollution control equipment, conducted among 311 industrial plants in Surat, Gujarat.

3.3.1 Engineering Inputs on Air Pollution Control Technologies

First, we use expert inputs on air pollution control technologies from environmental consultants at a national industrial federation - the Federation of Indian Chambers of Commerce and Industry (FICCI). FICCI was consulted to prepare detailed energy and air pollution control audits for eight sample industrial plants in Surat. FICCI also helped develop a detailed methodology to conduct an automated, scaled up audit analysis for all surveyed plants. This audit analysis method is used in our engineering model. We also use information provided by FICCI on Air Pollution Control Device (APCD) maximum collection efficiencies, typical lifetimes, and capital, operations, and maintenance costs for different emission source capacities. Appendix Section C.2 contains details on all the inputs provided by FICCI that are used in the model.

FICCI primarily provided data on three types of APCDs commonly used in Surat: cyclones, bag filters, and scrubbers. Cyclones are the least expensive and least effective of the APCDs. Cyclones force the flue gas to flow in an upward spiral (that is, in a cyclonic motion), and a fraction of the suspended particles settle down due to the centrifugal force. While cyclones have an efficiency of 80% overall, they are highly efficient in removing relatively larger particles, and are placed as the first APCD in an APCD chain. In most industries in Surat, a cyclone is followed by a second, more efficient APCD. The most common second APCD is a bag filter, which is much more expensive than a cyclone. As the name suggests, the flue gas is made to pass through a series of porous bags, which filter out the particles. Bag filters are very efficient and if designed and maintained properly, can remove 99% of the particles from the flue gas. An APCD system consisting of a cyclone plus bag filter has a maximum theoretical efficiency of 99.8%; put another way, if this system were to be maintained well the industrial plant would be comfortably within regulatory norms. The other APCD commonly used in Surat is a scrubber, where a water spray is used to scrub the particles out of the flue gas. The maximum efficiency of a scrubber is 94%, somewhat less than a bag filter. An APCD system consisting of a cyclone plus scrubber has a maximum theoretical efficiency of 98.8%. Scrubbers are cheaper than bag filters but more expensive than cyclones.

3.3.2 Survey of Industrial Plants in Surat, Gujarat

Second, we use data from a survey commissioned by the Central Pollution Control Board (CPCB) on a sample of 311 industrial plants in the Surat industrial cluster[3]. The CPCB survey collected data on each parallel chain attached to the stack, including the size and operating conditions of the emissions source and air pollution control equipment. The survey included data on various operating conditions at the plant, such as annual working hours. Most importantly, the CPCB survey also included manual sampling of the stacks in order to measure pollution levels and flow rates under normal operating conditions.

This survey was done in Surat, a large industrial city in the state of Gujarat in western India,

3. The survey was implemented for over 350 plants, but we use the information for the 311 plants that had sufficient data for our model.

with a population of 4.5 million. Surat is an important port city, and serves as a major hub for small and medium enterprises that process and export diamonds, and process textiles. Surat now has a large number of information technology (IT) firms. Most of the industrial units in the area are small and medium sized enterprises that process textiles: typically, processing and dyeing cotton cloth. Some of the larger industries manufacture the dyes themselves. These enterprises are largely homogeneous in nature, and are situated in five dense industrial clusters scattered within and around the city.

3.4 Descriptive Analysis of Indian Industrial Plants in Surat

We present descriptive results on our sample of industrial plants in Surat. We show that most stacks have multiple combustion sources, and nearly all stacks have at least two air pollution control devices (APCDs) per combustion source. Nonetheless, particulate matter emission concentrations remain high and usually above the regulatory standard, though there are also cases where plants exceed the standards. The presence of air pollution control devices together with high particulate matter emissions implies the technology mandate is well enforced but the concentration standard is not.

First, most industrial plant stacks in Surat have multiple combustion sources and 'parallel chains'. Table 3.1 presents the common configurations of combustion sources and APCDS, and shows that 77.8% of the surveyed industries have multiple combustion sources. Combustion sources are combined with APCDs into 'parallel chains', that is, an emissions source followed by a system of APCDs.

Second, every industrial plant in our sample already has air pollution control devices installed. As seen in Table 3.1, 96.5% stacks have two or more APCDs per combustion source. Every industrial plant in our sample has at least one APCD installed at baseline.

Table 3.1: Configurations of Combustion Source and Air Pollution Control Devices

Combustion and APCD Configuration -	Percentage of Plants -	Observed Efficiency -	Maximum Theoretical Efficiency
Single combustion source + cyclone + bag filter	7.70%	93.30%	99.80%
Single combustion source + cyclone + scrubber	7.40%	86.90%	98.80%
Single combustion source + other APCD configurations	7.10%	-	-
Two combustion sources + two cyclones + two bag filters	17.00%	89.30%	99.80%
Two combustion sources + two cyclones + two scrubbers	16.40%	86.00%	98.80%
Two combustion sources + two cyclones + two bag filters + two scrubbers	7.40%	91.90%	99.99%
Two combustion sources + other APCD configurations	37.00%	-	-

Third, despite the presence of air pollution control devices, rates of non compliance with pollution regulations are high across plants. Most plants do not meet the regulatory concentration standard of 150 mg/Nm3. Figure 3.1 shows the distribution of particulate matter concentration across plants. Only 26% of plants meet the standard. The median plant has a PM concentration of 232 mg/Nm3, and the mean PM concentration across all plants is 346 mg/Nm3. Thus, monitoring and enforcement is weak - a result also found by other studies, such as Duflo et al. (2013).

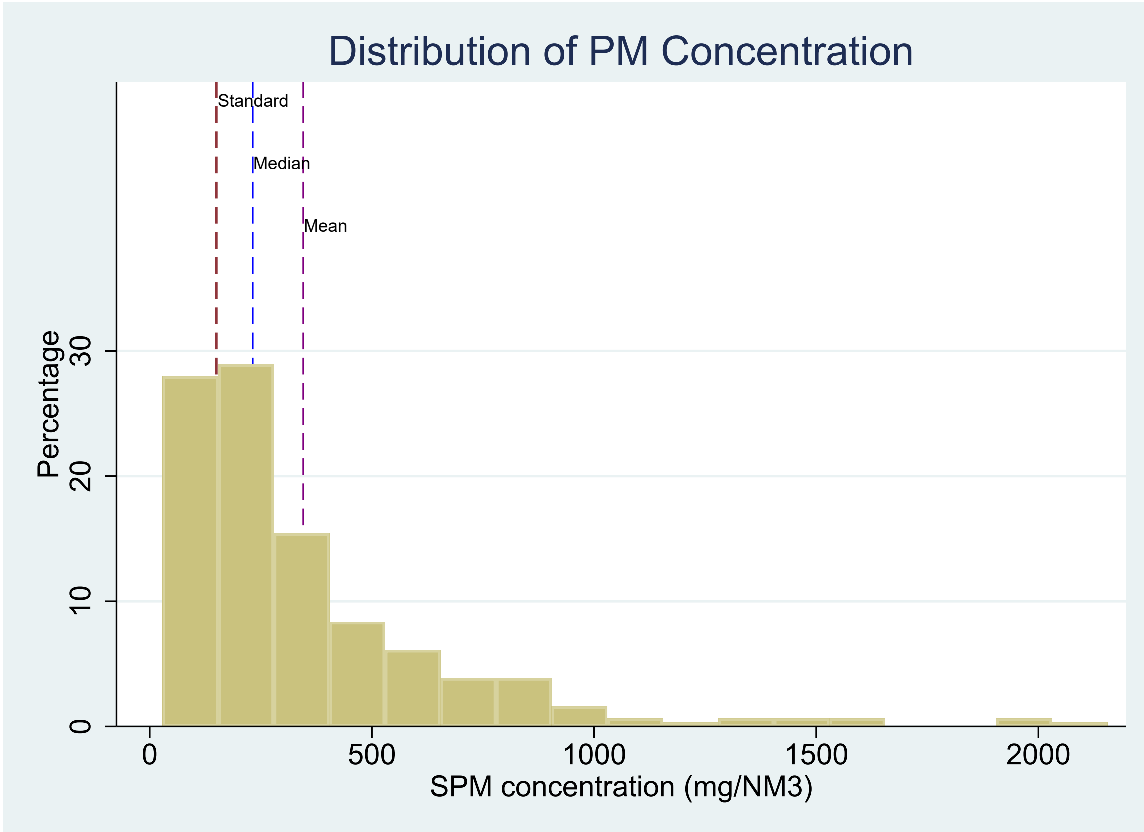

Figure 3.1: Distribution of PM Concentration Among Surat Industrial Plants

Fourth, since emissions exceed standards despite the presence of APCDs, this implies that the technology mandate is effectively enforced but the emissions concentration standard is not enforced, and also implies that the efficiency of the existing APCDs must be significantly less than the theoretical maximum efficiencies. If these installed APCDs were operating optimally, the plants would be able to meet standards. Table 3.1 presents the estimated actual efficiencies for the common APCD configurations along with the maximum theoretical efficiencies. As expected, the average actual APCD efficiency at baseline is significantly less than the theoretical maximum.

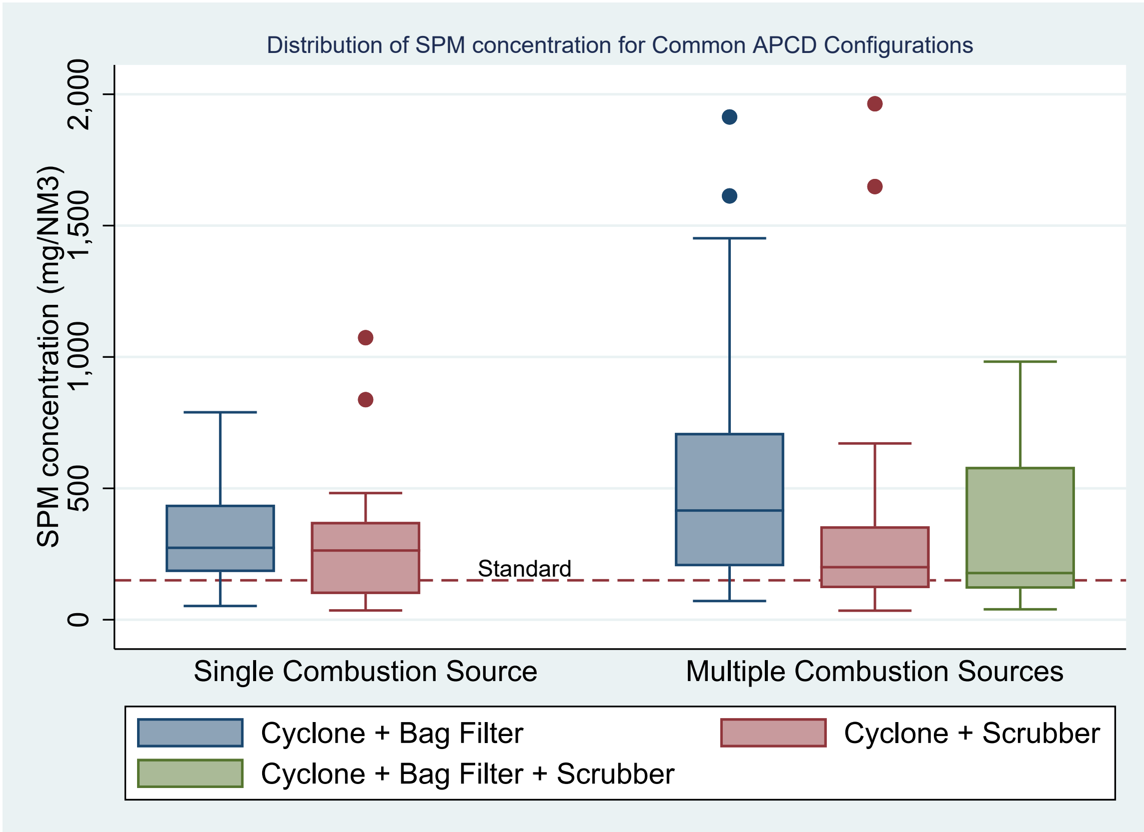

Figure 3.2: PM Concentrations For Common Configurations of Air Pollution Control Devices

The poor performance of the existing APCDs can also be seen by looking at the variation in PM emissions concentrations among plants with the same APCD configurations. Figure 3.2 plots the PM concentration distribution for the five most common configurations of combustion sources and APCDs. Each of these configurations include at least 2 APCDs per combution source. Yet, even within a particular APCD configuration, there is a wide dispersion in the emissions from the stack, and the bulk of the stacks are well above the concentration standard of 150 mg/Nm3.

Fifth and finally, there are some industrial plants who comply or overcomply despite the weak enforcement. Indeed, among the complying plants, most overcomply and exceed the concentration standard. This is an artifact of the technology options. For a particular set of air pollution control devices (APCDs), which are designed to have a specific concentration, it is not easy to adjust the concentration to achieve the standard exactly. A concentration standard need not inherently lead to overcompliance, but it does do so in this setting due to technology characteristics.

3.5 Engineering Model and Economic Framework on Particulate Matter Abatement in Indian Industrial Plants

3.5.1 Features of the Engineering Model based on Descriptive Results

We design a plant level engineering model on particulate matter abatement, with features and assumptions that incorporate the insights from the descriptive results, the baseline survey, and the expert engineering estimates on pollution control technology. In our model, the plant can optimize over two variables: the air pollution control technology selected, and how often to operate the pollution control system. We estimate particulate matter abatement and abatement costs for every combination of these two variables.

Insights from the Descriptive Results and Engineering Estimates

Our plant-level marginal abatement cost function incorporates the following insights from the descriptive results, qualitative research, and expert inputs.

Using our descriptive analysis, detailed in Section 3.4 , we note four relevant insights. First, in a single stack, plants can attach multiple chains of combustion sources followed by Air Pollution Control Devices (APCDs). These 'parallel chains' are common. Second, all plants have existing APCDs due to informal technology mandates, though these are degraded due to poor maintenance. Thus, to reduce emissions, a plant can invest in new APCDs or can simply retrofit and upgrade existing ones. Third, if an APCD is not maintained appropriately, it quickly deteriorates in efficiency. It is apparent from the baseline survey that many plants have subobtimal, low-efficiency APCDs that operate at much less than their theoretical efficiency. Fourth, many plants overcomply with concentration standards. Air pollution control devices (APCDs) are designed to have specific particulate matter (PM) outlet concentrations. Once a set of APCDs is selected, it is not easy for a plant to adjust the outlet concentration. Scrubbers, bag filters, cyclones, and other technologies vary in designed outlet concentration.

We note a fifth relevant insight based on our qualitative research and interviews with experts. Plants can 'bypass' the entire APCD system by allowing polluted air to be released directly into the atmosphere, without passing it through pollution control systems. Experts indicate this practice is common. Thus, though the survey data includes information on PM concentrations when APCD systems are operating, it is possible that APCD systems are bypassed at other times.

Features and Limitations of the Model

Model Feature 1 - Selecting a Technology:

For each stack k the plant can choose technology t out of T_k choices ($t = 1, 2, 3, \ldots, T_k, t = 0$ at baseline).

The plant can invest in new air pollution control devices (APCDs) or can retrofit and upgrade existing APCDs. Each distinct and technologically feasible combination of APCD equipment and quality is treated as a unique technology choice t. Plants can select abatement technologies that are a combination of new APCDs and retrofits. Thus, if the plant choses to install a new bag filter, that is one potential technology. If the plant chooses to upgrade the existing cyclone and add a scrubber, that combination of abatement actions is a separate technology. Each technology t, i.e. each discrete combination of abatement actions, yields a specific PM outlet concentration.

The T_k choices for stack k depend on the existing stack configuration, provided by the baseline data. The model incorporates one or two parallel chains, and does not incorporate three or more parallel chains. Each parallel chain can have up to two cyclones, one bag filter, one scrubber, and one electrostatic precipitator. Given baseline data on the existing parallel chain configuration and their efficiencies, the model calculates every combination of installing new APCDs or retrofitting existing APCDs, to form the set of potential technologies T_k. This is described in more detail in Appendix Section C.2.

Model Feature 2 - Choosing whether to Operate the APCD system:

For each stack k the plant can choose Fr_k, the fraction of time the APCD system operates ($Fr_k \epsilon [0, 1]$). Since the APCD system can be bypassed, plants can choose whether or not to operate the APCD system. The APCD system is bypassed for $(1 - Fr_k)$ fraction of the time.

Model Feature 3 - An Operating APCD system must be Fully Maintained:

In our model, if a plant chooses to operate an APCD system, the plant must pay the full costs of maintenance each year, in order to ensure the APCD system retains its intended efficiency. From the baseline survey and qualitative research, it is clear that an unmaintained APCD system rapidly deteriorates.

Limitations of the Model:

The engineering model solely considers the costs and abatement of using pollution control equipment to capture emissions generated by a combustion source, and thus reduce emissions released from the stack outlet. It does not consider ways to reduce the emissions generated from the combustion source, such as by reducing production, increasing the efficiency of the combustion source, or changing the fuel used by the combustion source. Additionally, we do not include potential installations or retrofits of electrostatic precipitators (ESPs), which are rarely used in our Surat setting.

3.5.2 Model Functions to Estimate Marginal Abatement Costs and Emissions

Based on the above key model features, we design stack-level marginal abatement cost functions and stack-level particulate matter emissions functions.

Marginal Abatement Cost Function:

For each stack k, the firm selects a technology t and chooses Fr_k, the fraction of time the APCD system operates, to minimize expected annual costs $C_k(t, Fr_k)$.

$$C_k(t, Fr_k) = CC_k(t) + MC_k(t, Fr_k) + OC_k(t, Fr_k)$$

Where $CC_k(t)$ is the expected annualized capital costs of technology option t for stack k, $MC_k(t, Fr_k)$ is the expected annual maintenance costs of technology t at Fr_k for stack k, and $OC_k(t, Fr_k)$ is the expected annual operations costs of technology t at Fr_k for stack k.

Capital costs $CC_k(t)$ depend only on the technology selected.

Maintenance costs $MC_k(t, Fr_k)$ depend on the technology selected and also depend on whether the APCD system is operated. As described above, in the model, if the firm operates an APCD system at all, it must maintain it in full. We also assume that if the APCD system is not used, the firm does not choose to maintain it. Thus,

$$MC_k(t, Fr_k) = 1(Fr_k = 0) * 0 + 1(Fr_k > 0) * APCDAnnualMaintenanceCost_k(t)$$

where $APCDAnnualMaintenanceCost_k(t)$ is the annual maintenance costs for the APCD system of stack k given technology t.

Operations costs $OC_k(t, Fr_k)$ depend on the technology selected and whether the APCD system is operated. Operations costs increase linearly with the hours the APCD system is operated.

$$OC_k(t, Fr_k) = Workinghours_k * Fr_k * APCDHourlyOperatingCosts_k(t)$$

Stack Emissions Functions:

The stack emissions can be measured in two main ways - as the stack outlet particulate matter concentration at a given moment in time, or as the total annual emissions per year.

Stack Annual Emissions:

$AnnualEmissions_k(t, Fr_k)$ is the annual emissions released from the smoke stack each year, in $kg/year$, for stack k, and is a function of the APCD technology t selected and the fraction of time Fr_k that the APCD system is operated, as follows:

$$\begin{aligned} &AnnualEmissions_k(t, Fr_k) \\ &\qquad = Workinghours_k * [Fr_k * HourlyEmissionsControlled_k(t) \\ &\qquad + (1 - Fr_k) * HourlyEmissionsGenerated_k] \end{aligned}$$

Where $HourlyEmissionsGenerated_k$ is the hourly emissions generated from the smoke stack, in

$kg/hour$, which is the sum of hourly emissions generated from up to two parallel chains,

$$\begin{aligned}&HourlyEmissionsGenerated_k\\&= HourlyEmissionsGenerated_kChain_1 + HourlyEmissionsGenerated_kChain_2\end{aligned}$$

And $HourlyEmissionsControlled_k(t)$ is the hourly emissions released from the smoke stack, in $kg/hour$, when the APCD system is operating. This is a function of the hourly emissions generated, and the efficiency of the APCD system given technology t. Thus,

$$\begin{aligned}&HourlyEmissionsControlled_k(t)\\&= HourlyEmissionsGenerated_kChain_1 * (1 - APCDefficiency_kChain1(t))\\&+ HourlyEmissionsGenerated_kChain_2 * (1 - APCDefficiency_kChain2(t))\end{aligned}$$

Where $APCDefficiency_kChain_i(t)$ is the net efficiency of the APCD system for chain i in stack k with technology t. Finally, the annual emissions is proportionate to the total working hours for stack k.

Stack Outlet Concentration:

$EmissionsConcentration_k(t)$ is the outlet particulate matter emissions concentration for a given technology t for stack k when the APCD system is operating, in units of mg/Nm^3. This is fixed given technology type t, and is a function of stack and technology characteristics as follows:

$$EmissionsConcentration_k(t) = \frac{HourlyEmissionsControlled_k(t) * 10^6}{HourlyFlowRate_k}$$

where $HourlyFlowRate_k$ is the hourly flow rate of air from the stack in units on $Nm^3/hour$, as measured in the baseline survey.

3.5.3 Using the Data to Generate Empirical Inputs for the Model

We use the two data sources, the inputs from FICCI and the industrial plant survey, to generate empirical inputs for the engineering model. The model estimates the costs of pollution reduction, where abatement could be in the form of installation of new pollution control equipment, increasing efficiency of existing equipment through retrofits, and varying the number of operating hours of the pollution control equipment. The model uses empirical inputs on several industrial parameters. Many parameters are directly taken from the CPCB survey, such as the parallel chain configuration, emission levels, flue gas rates and annual working hours. Other parameters are estimated using the survey data and the engineering estimates from expert environmental consultants at FICCI. We briefly describe how these parameters are estimated below, and then provide a detailed description in Appendix Section C.2.

We empirically determine the efficiencies of the current pollution control equipment using the stack emissions concentration data from the CPCB survey, and the theoretical maximum efficiencies of each air pollution control equipment type. For every industrial plant, we measure the parallel chain configuration, which is the number of parallel chains and the specific devices in the chain. Industrial plants with the same parallel chain configurations are pooled together, and the plant with the lowest emissions is assumed to have the maximum theoretical APCD system efficiency. We calculate the inlet emissions for this plant based on the APCD system efficiency and the stack outlet emissions (Using the formula $InletEmissions \times (1 - APCDSystemEfficiency) = OutletEmissions$). We assume this is the inlet emissions for every plant with this parallel chain configuration, and use this to estimate the APCD system efficiencies for every other plant with this parallel chain configuration. The efficiency of each individual air pollution control device in the system is then computed by assuming that all the components of the chain have equal efficiencies. Appendix Section C.2 describes the estimation of efficiencies in more detail.

We combine the FICCI cost inputs with the empirically determined efficiencies to estimate the abatement costs of installing, retrofitting, or operating air pollution control equipment. For new equipment, the capital costs of installing new pollution control devices are estimated based on the emissions source capacity and the expert inputs on market prices of pollution control technology. The annual operations and maintenance costs of the equipment are taken as fixed fractions (3% and 6%, respectively) of the upfront capital costs. Capital costs of retrofitting equipment are computed in a similar manner as for the new equipment. However, instead of the full capital costs, retrofit cost is estimated as a function of the relative increase in efficiency from actual to maximum possible. Increase in operations and maintenance cost are estimated as 3% and 6% respectively for the retrofit costs. Capital costs for all equipment upgrades are annualized based on the lifetime of the equipment. We present details in Appendix Section C.2.

3.5.4 Plant Level Abatement Cost Minimization Problems under Different Pollution Control Regulations

We next pair the engineering model with an economic framework in which each plant minimizes costs subject to regulatory constraints. Using the stack level marginal abatement cost and emission functions from the engineering model, we set up constrained optimization cost minimization problems for each stack k under four potential pollution control policies: an emissions concentration standard, an emissions load standard, a cap and trade program, and an emissions tax. We also consider several potential regulation stringencies, from 200 mg/Nm3 (a loosening of existing standards) up to 50 mg/Nm3. Appendix Section C.3 describes the solutions to these constrained optimization problems in detail, which are solved empirically using the survey data inputs.

Each potential regulation stringency is expressed as an emissions concentration standard. To convert the emissions concentration standards to quantity caps for the emissions load standard, cap and trade, and emission tax policies, we use the following formula:

$$EmissionsCap = \frac{ConcentrationStandardStringency}{MeanBaselineLoadWeightedConcentration} \times BaselineEmissionsLoad$$

Emissions Concentration Standard:

Under an emissions concentration standard regime, each stack k is required to have an emissions concentration $EmissionConcentration_k(t)$ at the stack outlet that is less than the emissions concentration standard $ConcentrationStandard$, at every moment of time. Thus, the APCD system must always be operational, as it is otherwise technologically impossible to meet the standard at every moment. Each firm's stack-level constrained optimization problem is:

$$Minimize_{t,Fr_k} C_k(t, Fr_k)$$

$$subject\ to$$

1. $EmissionConcentration_k(t) <= ConcentrationStandard$
2. $Fr_k = 1$

Emissions Load Standard:

Under an emissions load standard regime, each stack k is allocated a quantity of emissions permits
$EmissionsPermits_k$ for the year, in $kg/year$. We allocate permits as per each stack's combustion source capacity. Emissions are not allowed to exceed the allocated amount. Each firm's stack-level constrained optimization problem is:

$$Minimize_{t,Fr_k} C_k(t, Fr_k)$$

$$subject\ to$$

$$1.\ AnnualEmissions_k(t, Fr_k) <= EmissionsPermits_k$$

Cap and Trade:

Under cap and trade, each stack k is allocated emissions permits $EmissionsPermits_k$ for the year, in $kg/year$. Permit allocation is based on each stack's combustion source capacity. Stacks can buy and sell permits at the permit price Pr, and must ultimately hold permits equal to emissions. Each firm's stack-level constrained optimization problem is:

$$Minimize_{t,Fr_k} C_k(t, Fr_k) + Pr * [AnnualEmissions_k(t, Fr_k) - EmissionsPermits_k]$$

For each stack k, $[AnnualEmissions_k(t, Fr_k) - EmissionsPermits_k]$ is the number of permits purchased, that is, it is the quantity of emissions produced less the initial permit allocation.

Emissions Tax:

Under an emissions tax, every stack k pays price Tax for every unit of emissions released. We model a policy where there are no free allowances and every unit emitted is taxed. We set the tax at the level at which the aggregate emissions cap is expected to be achieved, so that the emissions tax has the same total emissions as the load standard and cap and trade policies. Each firm's stack-level constrained optimization problem is:

$$Minimize_{t,Fr_k} C_k(t, Fr_k) + Tax * AnnualEmissions_k(t, Fr_k)$$

3.6 Empirical Estimates of Abatement Costs

3.6.1 Costs and Emissions Reductions of Complying with Existing Policies

If the current regulation was strictly enforced and all plants were in compliance, we estimate a 66% reduction in average annual emissions, at a modest average cost of Rs. 36,150 per year, though there is significant heterogeneity in costs.

Table 3.2 compares the outcomes at baseline to the outcomes if the current regulation, an emissions concentration standard of 150mg/Nm3, was strictly enforced. At baseline, particulate matter concentrations are well above the standard, with a mean concentration of 346 mg/Nm3 and a median of 232 mg/Nm3. Concentrations drop drastically if the standard is enforced, to a mean of 109 mg/Nm3 109 mg/Nm3 and a median of 121 mg/Nm3. The mean annual quantity of emissions per year drops from a baseline of 41,756 kg/year to 14,047 kg/year, a 66% reduction in average annual emissions. This dramatic reduction of emissions would come at a mean cost of Rs. 36,150 per year, and a median cost of Rs. 23,231. These costs are extremely modest. For comparison, our sample industrial plants have an average annual salary expenditure of Rs. 32 million, an average annual turnover of Rs. 502 million, an average annual electricity bill of Rs. 28 million, and annual average expenditures on raw materials of Rs. 199 million.

Table 3.2: Baseline Outcomes Compared to Outcomes under Compliance

Panel A:

Variable (unit)	Mean at Baseline	Mean with Compliance
PM Concentration (mg/Nm3)	346	109
Annual Emissions(kg/year)	41756	14047
Increase in Costs (Rs./year)	0	36150

Panel B:

Variable (unit)	Median at Baseline	Median with Compliance
PM Concentration (mg/Nm3)	232	121
Annual Emissions(kg/year)	24934	11931
Increase in Costs (Rs./year)	0	23231

Panel C:

APCDs	Number Retrofitted	Number Installed
Cyclones	190	106
Bag filters	23	0
Scrubbers	67	2

There is a wide variation in how much costs would increase for specific plants. Figure 3.3 shows how costs vary. Though the median is Rs. 23,231, the upper half of the distribution has a long tail and includes plants with costs over Rs. 70,000 per year. 38% of all plants have cost increases of less than Rs. 10,000 per year. Plants that were compliant at baseline have no increases in costs.

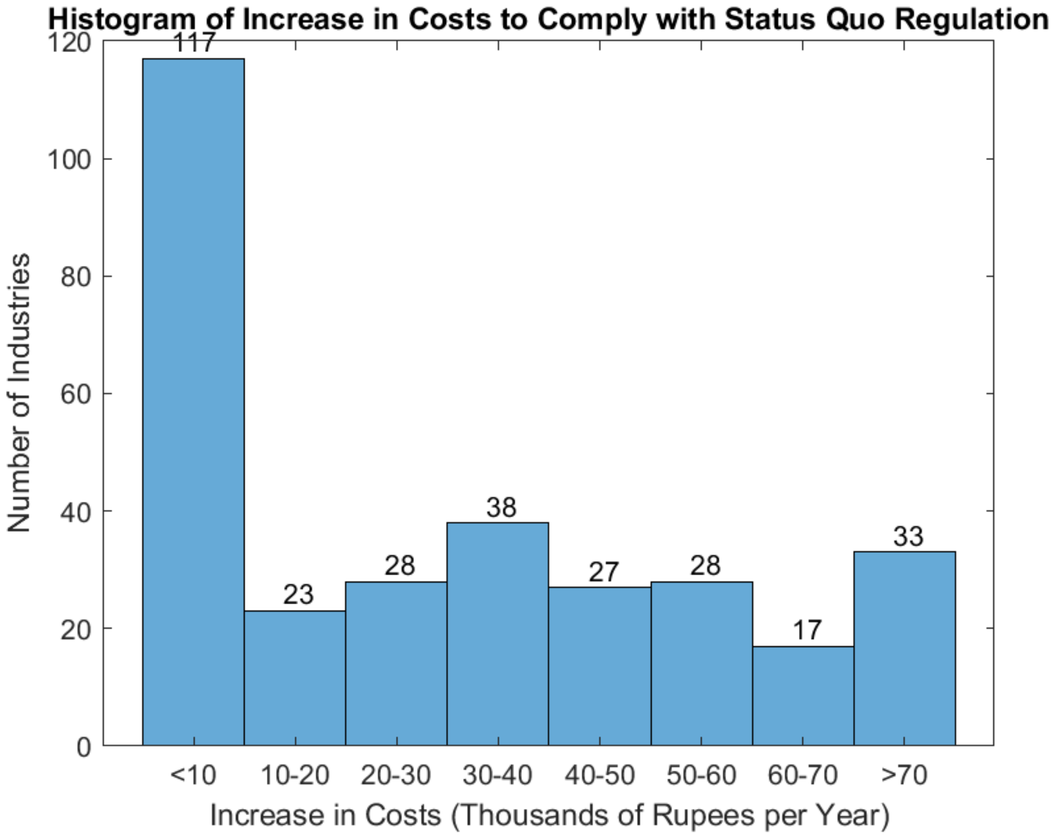

Figure 3.3: Distribution of Cost Increases from Universal Compliance with the Current Standard

Costs of abatement for the plants are usually modest because most abatement actions involve upgrades to and improved maintenance of existing equipment. Plants already have the technology needed to abate. Table 3.2 shows how pollution control technology is improved in order to achieve universal compliance with the standard. Because all the plants have existing air pollution control devices (APCDs), there are far more retrofits of existing technology - cyclones, bag filters, and scrubbers - than installations of new equipment.

3.6.2 Costs and Emissions Reductions of Alternate Policies and Regulation Stringencies

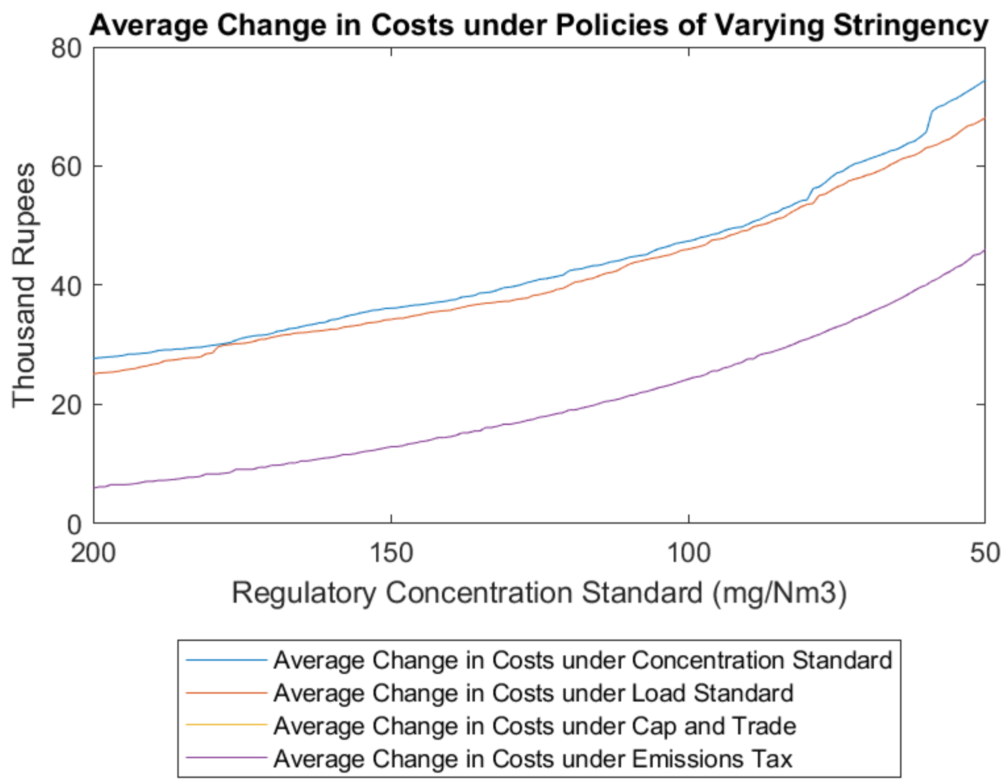

Figure 3.4: Average Change in Costs under Policies of Varying Stringency

We estimate costs and abatement for an emissions concentration standard, emissions load standard, cap and trade, and emissions taxes, for emissions concentration standard stringencies ranging from 200 mg/Nm3 to 50 mg/Nm3. Each stringency is converted to a quantity cap as described in Section 3.5.4. Aggregate emissions are significantly less than baseline emissions under every scenario we evaluate. Strictly enforced concentration standards lead to significant over-control. For any specific quantity of emissions reduction, market based instruments are significantly cheaper than concentration standards and load standards are costlier. For the status quo stringency of 150 mg/Nm3, switching to market based instruments would lead to a 39% reduction in abatement costs, while switching to a load standard would increase costs by 21%.

The increase in costs from alternate policies and alternate stringencies are presented in Table 3.3 and graphed in Figure 3.4. As Table 3.3 shows, the two market based instruments, cap and trade and emissions taxes, have the same costs and same emissions for a given emissions quantity cap. [4] Thus, in Figure 3.4, the emissions tax line overlays the cap and trade line.

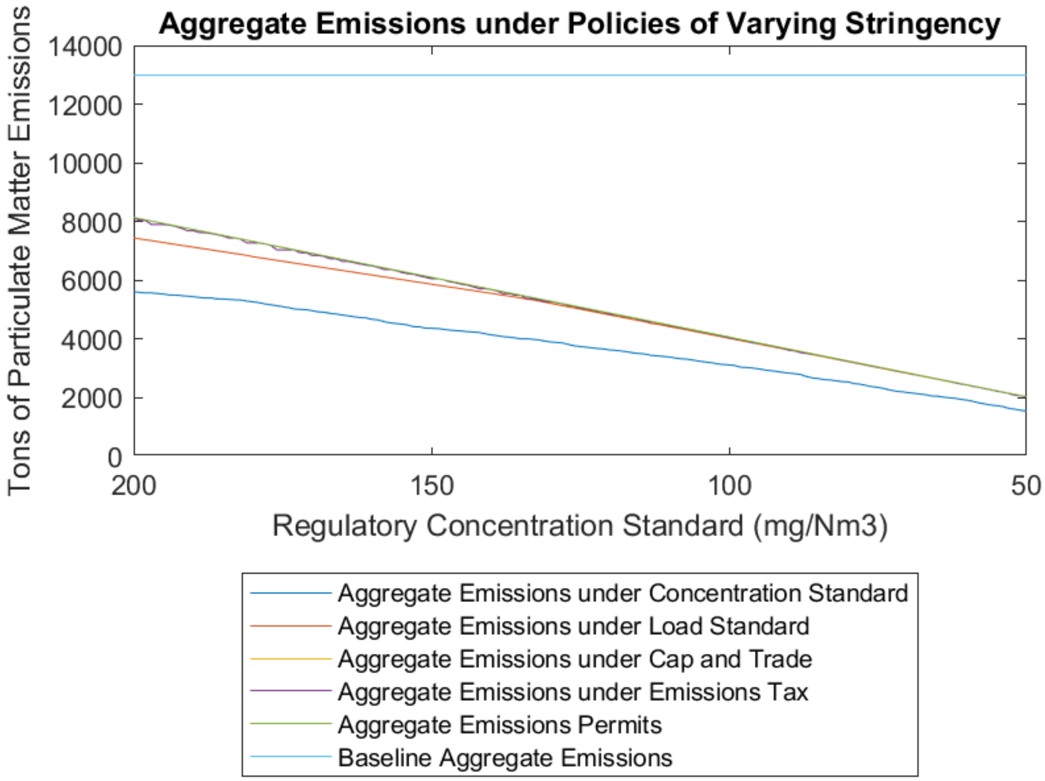

Figure 3.5: Aggregate Emissions under Policies of Varying Stringency

4. Given a specific quantity, quantity instruments and price instruments will have the same outcomes (Weitzman 1974) in a model like this that does not incorporate uncertainty in cost or benefit estimates.

Table 3.3: Change in Costs under Policies of Varying Stringencies

Concentration Standard (mg/Nm3)	Aggregate Change in Costs (Thousand Rupees/Year)				Average Change in Costs (Thousand Rupees/Year)			
	Concentration Standard	Load Standard	Cap and Trade	Emissions Tax	Concentration Standard	Load Standard	Cap and Trade	Emissions Tax
160	10642	10143	3459	3459	34	33	11	11
150	11243	10655	4017	4017	36	34	13	13
140	11664	11134	4538	4538	38	36	15	15
130	12345	11609	5187	5187	40	37	17	17
120	13200	12442	5938	5938	42	40	19	19
110	13890	13530	6675	6675	45	44	21	21
100	14742	14323	7537	7537	47	46	24	24
90	15614	15307	8604	8604	50	49	28	28
80	16900	16667	9629	9629	54	54	31	31
70	18984	18189	10917	10917	61	58	35	35
60	20419	19618	12456	12456	66	63	40	40
50	23163	21177	14309	14309	74	68	46	46

Table 3.4: Aggregate Emissions under Policies of Varying Stringency

Concentration Standard Stringency (mg/Nm3)	Aggregate Emissions under Different Policies (tons/year)				Aggregate Permits tons/year	Baseline Emissions tons/year
	Concentration Standard	Load Standard	Cap and Trade	Emissions Tax		
160	4677	6183	6494	6494	6511	12986
150	4369	5871	6057	6057	6104	12986
140	4145	5559	5689	5689	5697	12986
130	3903	5221	5276	5276	5290	12986
120	3630	4825	4841	4841	4883	12986
110	3392	4430	4457	4457	4476	12986
100	3122	4033	4052	4052	4069	12986
90	2842	3634	3620	3620	3662	12986
80	2542	3235	3251	3251	3256	12986
70	2181	2835	2847	2847	2849	12986
60	1922	2434	2438	2438	2442	12986
50	1549	2030	2026	2026	2035	12986

Under any policy, aggregate emissions are significantly less than the emissions at baseline. The aggregate emissions under various scenarios are presented in Table 3.4 and graphed in Figure 3.5. The emissions tax line overlays the cap and trade line in Figure 3.5. The load standard, cap and trade, and emissions tax policies lead to aggregate emissions that are very close to the aggregate emissions cap. In contrast, a concentration standard policy leads to substantial over-control, with aggregate emissions significantly below the equivalent emissions cap for every possible concentration standard stringency.

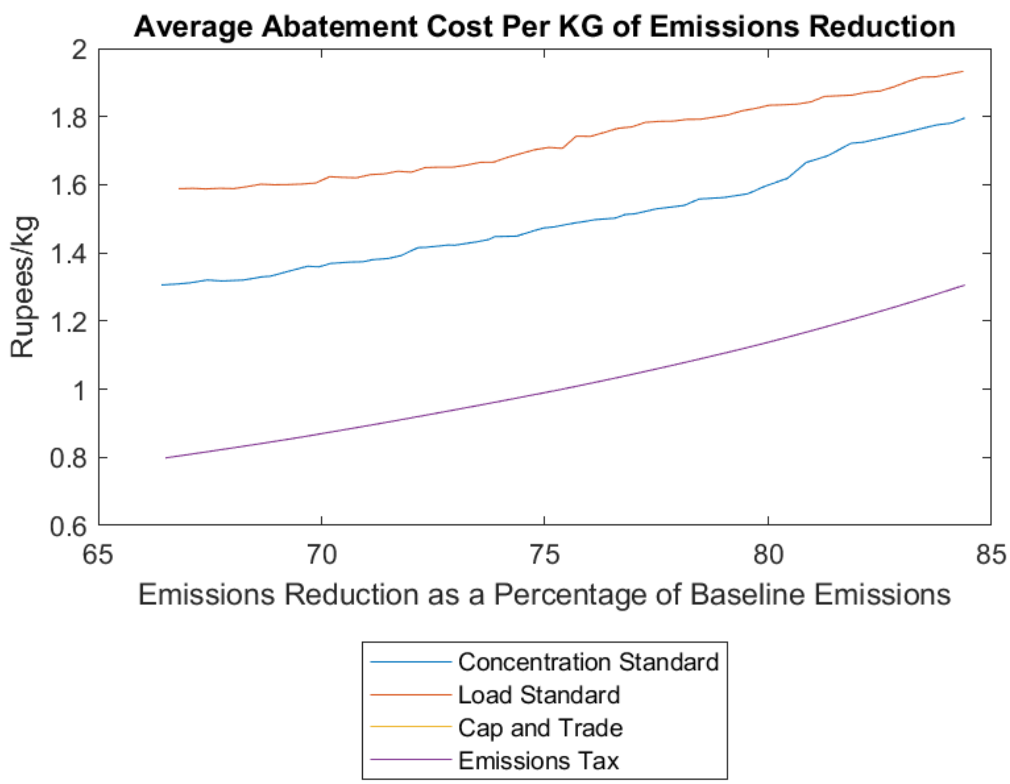

Figure 3.6: Average Abatement Cost per KG under Policies of Varying Stringency

Concentration standards, if enforced strictly, lead to significant over-control due to the limited discrete technologies available. As the concentration standard applies at every moment, stacks must always run the APCD system to be in compliance. The stack concentration is wholly a function of technology choice. The pollution control technology market in India has a limited number of discrete technologies, so firms cannot perfectly meet the standard of 150 mg/Nm3 standard. For example, the choice may be between a technology with 200 mg/Nm3 and 100 mg/Nm3.

In Figure 3.6 and Table 3.5, to compare the abatement costs of different policies, we estimate the cost per kilogram of emissions reductions, for a given percentage reduction of emissions relative to baseline. This makes the different policies comparable. Looking at total cost alone is misleading, since a given concentration standard leads to lower emissions than the equivalent emissions cap.

Table 3.5: Average Abatement Cost Per Unit for Different Levels of Abatement

Percentage Reduction in Baseline Emissions	Average Abatement Cost Per KG (Rupees/KG)			
	Concentration Standard	Load Standard	Cap and Trade	Emissions Tax
66.42%	1.31	1.59	0.80	0.80
68.08%	1.32	1.59	0.83	0.83
69.69%	1.36	1.61	0.86	0.86
71.14%	1.38	1.63	0.90	0.90
72.84%	1.42	1.65	0.94	0.94
74.37%	1.45	1.69	0.97	0.97
75.84%	1.49	1.74	1.01	1.01
77.49%	1.53	1.79	1.06	1.06
79.36%	1.57	1.82	1.12	1.12
80.85%	1.67	1.84	1.17	1.17
82.84%	1.75	1.89	1.24	1.24
84.40%	1.80	1.93	1.31	1.31

As predicted by theory, for any given level of emission reductions, market based instruments are significantly cheaper than either type of command and control regulation. If the current concentration standard of 150 mg/Nm3 was fully enforced, there would be a 66% reduction in baseline emissions, for an average abatement cost of Rs. 1.31/kg. Achieving a 66% reduction in baseline emissions with cap and trade or emissions taxes would be possible for an average abatement cost of Rs. 0.80 per kg. Thus plants would see 39% reduction in costs from switching to a market based instrument. Alternatively, for the same average abatement cost of Rs. 1.31/kg, there would be a 84% reduction in baseline emissions under a market based instrument, as shown in Table 3.5 which presents the differences in costs at several stringency levels.

Within command and control regulations, concentration standards significantly outperform load standards, as shown in Figure 3.6. Achieving a 66% reduction in emissions with a load standard would cost Rs. 1.59 per kg, a cost increase of 21%, relative to the existing concentration standard policy.

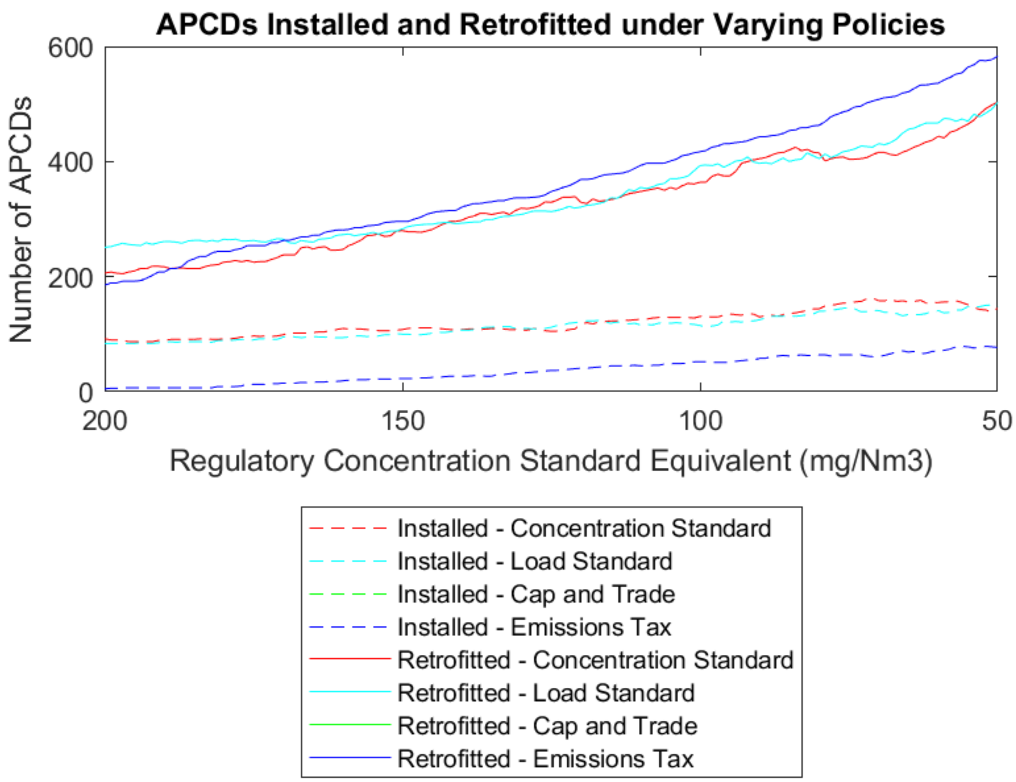

Figure 3.7: Number of APCDs Installed and Retrofitted under Policies of Varying Stringency

For all policies and stringencies that we evaluate, the number of retrofits remain significantly larger than the number of new installations of APCDs, as shown in Figure 3.7. This again highlights that all industries already have pollution control equipment installed, which are being poorly maintained and operated at baseline.

3.6.3 What Factors Drive the Cost Differences Between Policies?

Cost Differences Within Command and Control Policies:

Within command and control policies, the emissions load standard is more expensive than an emissions concentration standard because it does not lead to over control. To be in compliance under an emissions load standard, every plant must have high performing APCD systems, but there is no incentive to operate the APCD system beyond the minimum necessary to be in compliance with the emissions cap. Thus, emissions load standards have high fixed costs, but the APCD systems are underutilized, so total costs are significantly higher per unit of emissions reduction.

Table 3.6 provides a detailed breakdown comparing outcomes under the baseline, an enforced emissions concentration standard of 150 mg/Nm3, and enforced alternate policies at the same stringency level. Under the emissions concentration standard, cap and trade, and emissions tax policies, the mean and median 'Fraction of the time the APCD system is operated' equals one, that is, the APCD system is always operated. Under an emissions concentration standard policy, this is necessary to be in compliance. Under cap and trade or emissions tax policies, once the high fixed costs of an APCD system are incurred, it is optimal to operate it at all times and sell emissions permits or reduce the total tax burden. However, under an emissions load standard, the plant cannot utilize excess permits in anyway, so in order to save on operational costs, the APCD system is not operated all the time. The mean fraction of the time the APCD system is operated is 0.97, and the median fraction of time is 0.99. Though the plant has a small savings from operational costs, the fixed capital costs remain large, making emissions load standards a costly policy.

Table 3.6: Outcomes under Full Enforcement of Status Quo Policies

Variable (unit)	Baseline	Enforced Concentration Standard	Enforced Load Standard	Enforced Cap and Trade	Enforced Emissions Tax
Mean PM Concentration (mg/Nm3)	346	109	115	191	191
Median PM Concentration (mg/Nm3)	232	121	94	165	165
Mean Annual Emissions (tons/year)	41.76	14.05	18.88	19.48	19.48
Median Annual Emissions (tons/year)	24.93	11.93	14.56	17.56	17.56
Mean Increase in Costs (Thousand Rupees/Year)		36.15	34.26	12.92	12.92
Median Increase in Costs (Thousand Rupees/Year)		23.23	25.62	3.91	3.91
Mean Increase in Capital Costs (Thousand Rupees/Year)		24.17	22.95	8.62	8.62
Median Increase in Capital Costs (Thousand Rupees/Year)		15.51	16.56	2.6	2.6
Mean Increase in Maintenance Costs (Thousand Rupees/Year)		3.99	3.79	1.43	1.43
Median Increase in Maintenance Costs (Thousand Rupees/Year)		2.62	2.7	0.44	0.44
Mean Increase in Operations Costs (Thousand Rupees/Year)		7.98	7.52	2.86	2.86
Median Increase in Operations Costs (Thousand Rupees/Year)		5.23	5.39	0.87	0.87
Mean Fraction of Time APCD System is Operated		1	0.97	1	1
Median Fraction of Time APCD System is Operated		1	0.99	1	1

Cost Differences Between Command and Control Policies and Market Based Instruments: Case Study of a Sample Plant

Under market based instruments like cap and trade policies and emissions taxes, costs are significantly less than an emissions concentration standard because abatement takes place in the plants that have the cheapest abatement costs. This is the key theoretical reason that market based instruments minimize aggregate abatement costs. Plants with relatively high abatement costs can buy permits or pay higher taxes, rather than invest in APCDs. Plants with relatively low abatement costs can do excess emissions reduction, and then sell permits or pay lower taxes.

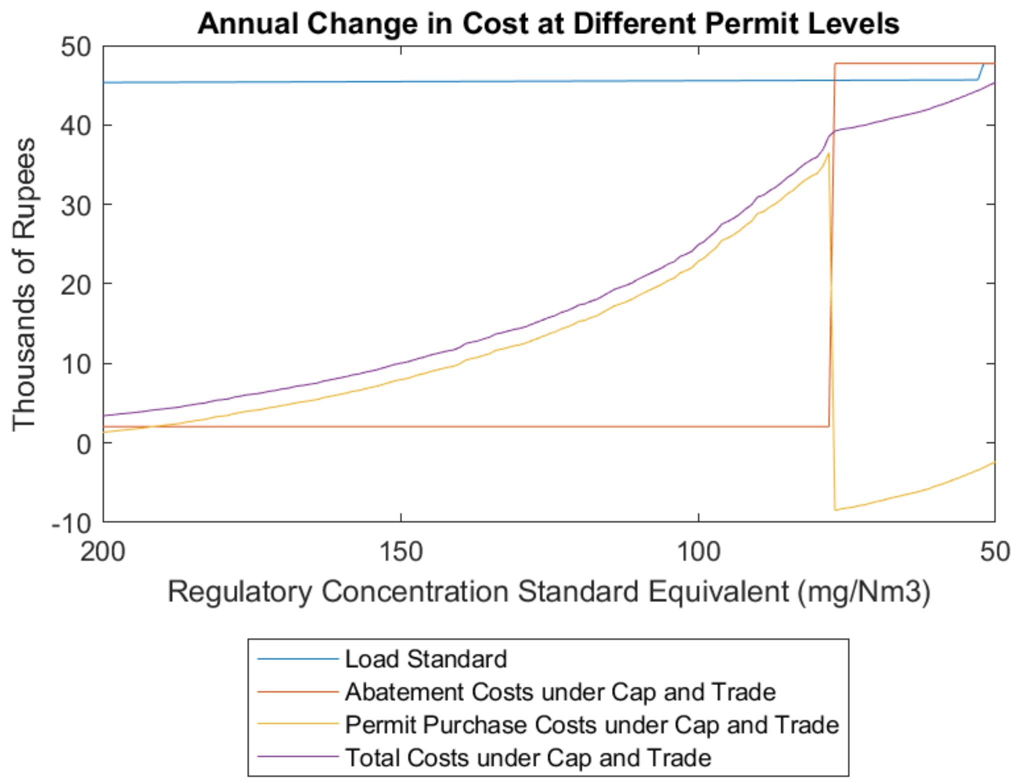

Figure 3.8: Sample Plant - Comparing Annual Changes in Cost under a Load Standard and Cap and Trade

To demonstrate how market based instruments reduce costs, we show a case study of a sample plant in our study. At baseline, the sample plant has a cyclone and bag filter attached to a single combustion source, a baseline PM concentration of 238 mg/Nm3, and baseline annual emissions of 21,700kg. Figure 3.8 shows the annual change in costs for this plant under the emissions load standard and cap and trade policies. Under the emissions load standard policy, for regulation stringencies of 200 mg/Nm3 to 53 mg/Nm3, the sample plant meets the emissions cap by installing a new second cyclone. Annual total increase in costs ranges from

Rs.45,340 for the 200 mg/Nm3 standard, to Rs. 45,664 for the 53 mg/Nm3 case, since the plant's operations costs increase as the APCD system is operated for more of the time. For the emissions concentration standards from 52-50 mg/Nm3, the plant installs the new second cyclone and also retrofits the original cyclone.

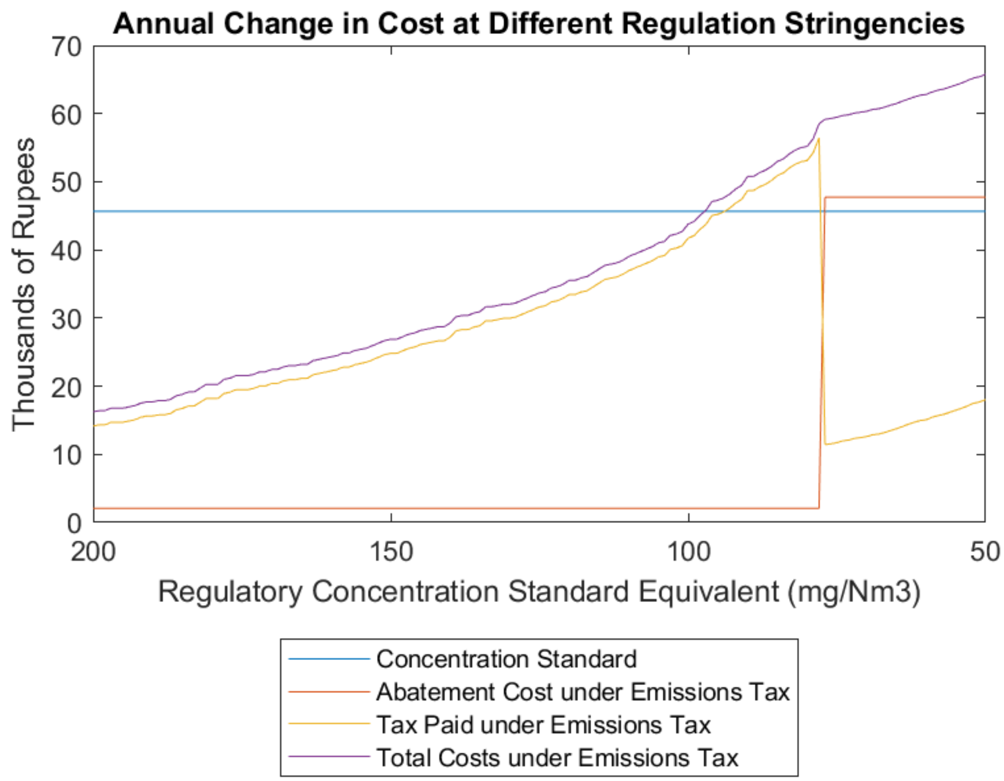

Figure 3.9: Sample Plant - Comparing Annual Changes in Cost under a Concentration Standard and Emissions Taxes

In contrast, under a cap and trade policy, for regulation stringencies of 200 mg/Nm3 to 78 mg/Nm3, the plant retrofits the original cyclone and operates it continuously, for a constant increase in abatement costs of Rs. 2067, as seen in the orange line. To meet the emissions cap, the plant also purchases permits and pays an increasing permit cost in the yellow line. At 200 mg/Nm3, the plant pays Rs. 1,343 to buy permits. This increases to Rs. 36,511 at 78 mg/Nm3. For regulation stringencies of 77 mg/Nm3 to 50 mg/Nm3, the plant installs a new second cyclone as well as retrofitting the original cyclone. The APCD system is operated continuously, at a constant increase in abatement costs of Rs. 47,734. The plant moves from being a buyer of permits, to selling its surplus permits, thus incurring negative costs in the yellow line depicting permit costs. Thus, due to trading of emissions permits, total costs for the plant in the purple line are much lower than the total costs under an emissions load standard depicted depicted in the blue line.

Figure 3.9 shows the annual change in costs for this plant under the emissions concentration standard and emissions tax policies. Under the emissions concentration standard policy, for all regulation stringencies from 200 mg/Nm3 to 50 mg/Nm3, the plant has to install a new, second cyclone to be in compliance, and operate it continuously, for a constant total increase in abatement costs of Rs. 45,666. Under the emissions tax, the abatement decisions and costs are identical to the cap and trade case. Instead of buying permits, the plant pays an increasing emissions tax from 200 mg/Nm3 to 78 mg/Nm3. Because this particular tax policy taxes every unit of emissions, the plant pays significantly more in emissions taxes than on permits under cap and trade. At 200 mg/Nm3, the plan pays Rs. 14,154 in taxes, which increases to Rs. 56,431 at 78 mg/Nm3. At 77 mg/Nm3, the total tax paid drops to Rs. 11,433 due to the dramatic reduction in emissions once the second cyclone is installed. Because the cap and trade policy provides a free permit allocation, while the emissions tax program taxes every unit of emissions, the emissions tax program involves a transfer from the plant to the government. Alternate emissions tax policies could be designed that provide a certain quantity of tax-free emissions to each plant. Because of this transfer, the total costs to the plant under emissions taxes is higher than in the emissions concentration standard policy for regulation stringencies of 97 mg/Nm3 to 50 mg/Nm3.

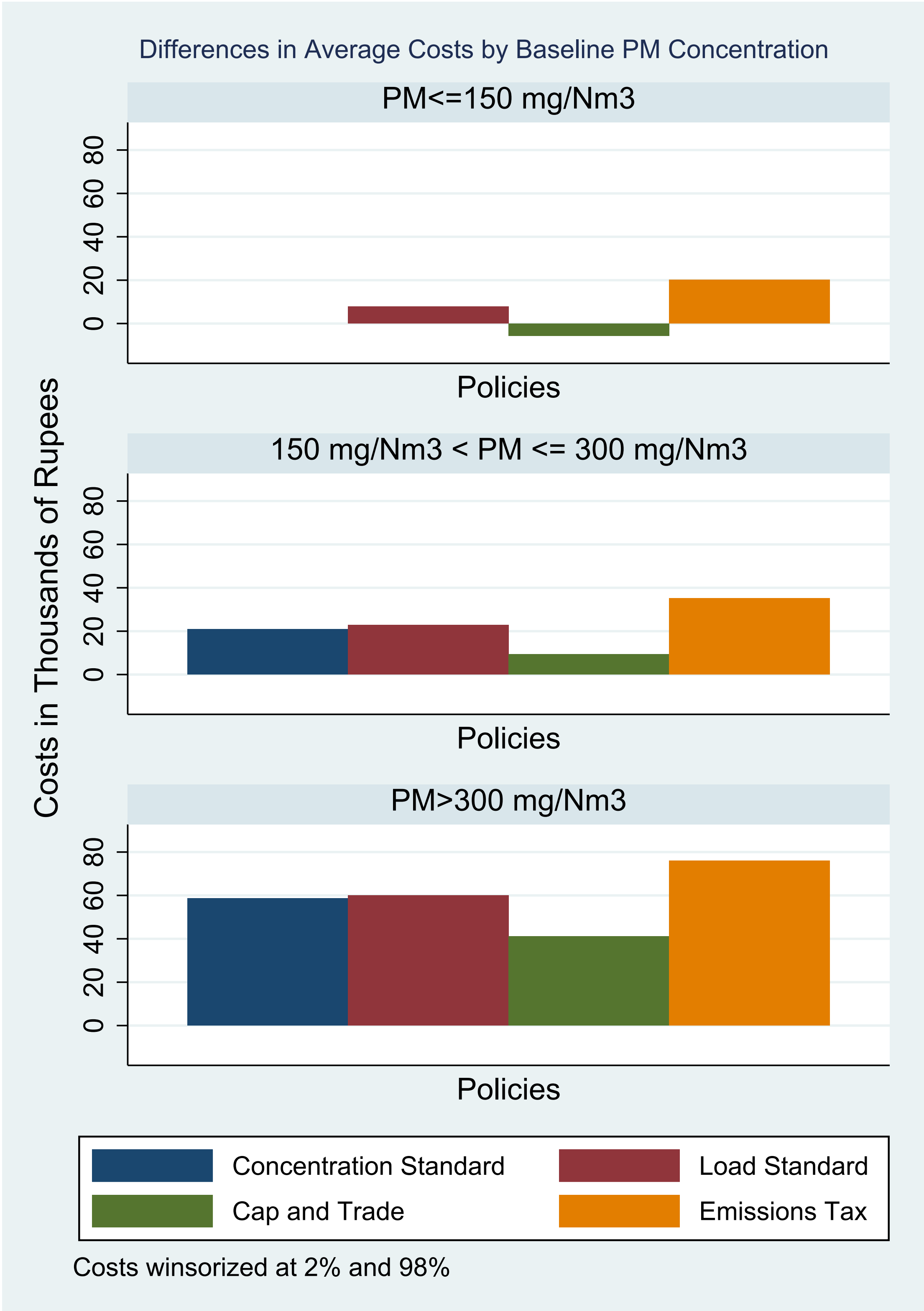

Figure 3.10: Variation in Increases in Costs by Baseline PM Concentration

3.6.4 Distribution of Changes in Costs under Alternative Policies

Figure 3.10 shows how the increase in total costs varies by baseline PM concentration. The three panels depict the average increase in total costs for the current 150mg/Nm3 regulation stringency under the four policies. These graphs represent costs to the firm. Total costs to the firm only include abatement costs for the emissions concentration and emissions load standards. For cap and trade, total costs include abatement costs plus costs spent on buying permits. At the aggregate level, these cancel out, and permit purchase costs represent a transfer from buyers of permits to sellers. For the emission tax program, total costs include abatement costs plus the total tax paid on emissions, the latter of which is a transfer to the government.

Plants that are compliant at baseline face the least absolute increase in costs, but are relatively disadvantaged by a shift to a quantity based policy. Under an emissions concentration standard, compliant plants face no changes in costs - they continue to do what they were doing at baseline. Each of the three alternate policies, an emissions load standard, cap and trade, or an emissions tax, is based on emissions quantity rather than emissions concentrations. There are distributional consequences of switching to quantity based policies, which allocate emissions permits to plants based on combustion source capacity. Some plants will be disadvantaged by this, leading to an increase in average costs. In short, reform consistently disadvantages plants that are compliant at baseline, since these plants are already optimized for the existing emissions concentration standard, and often incur some costs from a change to a quantity based policy. As Figure 3.10 shows, under an emissions load standard, this leads to a small increase in average costs. Under cap and trade, the disadvantage of the shift to a quantity policy is offset by permit sales. Overall, compliant firms actually profit, as they are able to sell excess permits to plants who are non-compliant at baseline. Under an emissions tax, the plant makes a sizable transfer to the government by paying for every unit of PM emitted, so total costs to the plant go up significantly.

For plants that are non-compliant at baseline, the increase in costs is nearly identical for the concentration standard and load standard policies. Depending on whether a concentration standard or load standard is enforced, plants optimize their APCD systems for either concentration or quantity. Since these plants are not already optimized for compliance with a concentration standard at baseline, they are minimally affected by the switch to a load standard. Under a cap and trade program, these industries are able to lower their overall costs. However, unlike the plants who are compliant at baseline, they cannot make absolute profits through cap and trade. Under emissions taxes, the plants pay higher total costs due to the tax burden. The pattern of cost increases is similar for plants in the 150-300 and above-300 mg/Nm3 categories, however, absolute costs are much higher for plants with baseline PM concentration over 300 mg/Nm3, who have to make large investments in APCD systems to become compliant.

3.7 Cost Benefit Analysis

To compare our estimated abatement costs to the potential benefits, we estimate the health benefits of improved life expectancy from particulate matter abatement in Surat district. For a 66% reduction in baseline emissions, our central estimates suggest that the net benefits under a concentration standard are Rs. 580-583 billion, and the net benefits under an emissions tax are Rs. 584-586 billion. We also present net benefit estimates for different levels of emissions reduction. These estimates involve numerous assumptions and uncertainties. They only include the costs of abatement, and not monitoring and enforcement costs to the regulator, or costs to the plants of purchasing continuous monitoring systems, or other associated costs of reform. We also ignore benefits unrelated to mortality, such as days of work lost, morbidity and quality of life, etc. Nonetheless, we can draw the overall conclusion that the health benefits of abatement substantially exceed abatement costs.

3.7.1 Benefits of Abatement

First, we analyze the benefits of abatement. We monetize the life expectancy benefits if 66% of industrial pollution was abated. We use multiple estimates of the life expectancy benefits from PM abatement, measured in years. We monetize this using Value of Life Years, which is the economic value of a one year increase in life expectancy. Our India-specific Value of Life Years estimates are based on the Lancet's 'Global Health 2035' report (Jamison et al. 2013). Our central estimate of total monetized life expectancy benefits for Surat District is Rs. 590,645 million, using Chen et al. (2013)'s estimates of life expectancy benefits from PM abatement. The lower and upper bounds of total benefits would be Rs. 209,584 million and Rs. 1,066,972 million, based on estimates from Correia et al. (2013) and Laden et al. (2006) respectively.

Industrial pollution is responsible for about 6.2 microgram/m3 of Surat's fine particulate pollution. In 2014-2015, the Central Pollution Control Board (CPCB) estimated that $PM_{2.5}$ concentrations in Surat was between 29-33 microgram/m3, based on data from seven monitoring stations, indicating a mean and median $PM_{2.5}$ concentration of 31 microgram/m3[5]. This number is within the national standard of 40 microgram/m3, but is three times higher than the 10 microgram/m3 level recommended by the World Health Organization. S. Guttikunda and Pooja Jawahar (2011) and S. K. Guttikunda and Puja Jawahar (2012) estimate that the industrial contribution to ambient $PM_{2.5}$ concentrations in Surat is about 20%, so the industrial sector is responsible for about 6.2 microgram/m3 of $PM_{2.5}$.

Our central estimate of life expectancy gains from reducing particulate matter by 6.2 microgram/m3 is 0.62 years. There are not many estimates in the Indian context for life expectancy gains with PM reduction. Most estimates (e.g. Pope, Ezzati, and Dockery 2009; Correia et al. 2013) are from the USA where the ambient concentrations are significantly lower. A more comparable estimate can be derived from Chen et al. (2013) for China, where ambient concentrations are similar to India. As Chen et al. (2013)'s estimates are for total suspended particulate matter (particles of sizes up to 100 micrograms), we rely on Greenstone et al. (2015)'s conversion that a 10 microgram/m3 reduction of fine particulates ($PM_{2.5}$) improves life expectancy by 1.0 years. Greenstone et al. (2015) also consider several of the main estimates used in the literature, which range from life expectancy gains of 0.35 to 1.8 years per 10 microgram/m3 of $PM_{2.5}$, from Correia et al. (2013) and Laden et al. (2006) respectively. Our central estimate based on Chen et al. (2013) suggests that a 6.2 microgram/m3 reduction in $PM_{2.5}$ would increase life expectancy by 0.62 years. The estimates of 0.35 and 1.8 years per 10 microgram/m3 of $PM_{2.5}$ would correspond to life expectancy gains in Surat district of 0.22 years and 1.12 years respectively.

Our central monetized estimate of life expectancy gains is that 0.62 years of life expectancy gains is worth Rs. 147,190 per person. Life expectancy gains of 0.22 years and 1.12 years would

5. Gujarat Pollution Control Board. "Ambient Air Quality Monitoring Programme: National Air Quality Monitoring Programme". http://gpcb.gov.in/ambient-air-quality-monitoring-programme.htm. Accessed 28th February 2018.

be worth Rs. 52,229 and Rs. 265,892 respectively. We use Value of Life Years (VLY)- that is, the economic value of a one year increase in life expectancy. The Lancet's 'Global Health 2035' report estimates that in low-income and middle-income countries, one VLY is 2.3 times the per person income (Jamison et al. 2013). Per capita income in India was Rs. 1,03,219 in 2016-2017 (GoI 2018). This implies the current VLY in India is Rs. 237,404. This is comparable to commonly used Value of a Statistical Life (VSL) estimates in India[6].

Our central estimate of total monetized life expectancy benefits for Surat District is Rs. 894,917 million rupees for 0.62 years of life expectancy gains. Life expectancy gains of 0.22 years and 1.12 years would have total monetized life expectancy benefits of Rs. 317,551 million rupees and Rs. 1,616,624 million rupees respectively. To get these estimates, we multiply the per person life expectancy gain by total population. Surat district has a population of 6.08 million people as of the 2011 Census of India[7].

For a 66% reduction in baseline emissions, benefits would be 66% of the above numbers. That is, a 66% reduction in baseline emissions would reduce $PM_{2.5}$ concentration by 4.092 microgram/m3 (66%*6.2). This would be total monetized life expectancy benefits of Rs. 590,645 million using our central estimate based on Chen et al. (2013). The lower and upper bounds of total benefits would be Rs. 209,584 million and Rs. 1,066,972 million, based on Correia et al. (2013) and Laden et al. (2006) respectively. Our estimates assume that the reduction in suspended particulate matter pollution from the plants would lead to a proportionate reduction in the industrial component of Surat's PM 2.5 emissions.

3.7.2 Costs of Abatement

We estimate the district-wide costs of a 66% reduction of baseline emissions, which is our estimate of total abatement in the surveyed plants if status quo regulations were fully enforced. We consider a scenario where surveyed plants are 3.6% of Surat's industrial emissions, in which

6. For instance, Cropper et al. (2017) study costs and benefits of air pollution control technology at coal-based thermal power plants in India, and use two estimates of VSL: Rs. 1.3 million (Bhattacharya, Alberini, and Cropper 2007) and Rs. 15 million (Madheswaran 2007). The Rs. 1.3 million estimate is based on 2005 data, and the Rs. 15 million estimate is based on 2001 data, when India's per capita income and hence VLY was much lower than today.

7. "Surat District: Census 2011 data". Census2011. https://www.census2011.co.in/census/district/206-surat.html. Accessed 28th February 2018

case net present abatement costs are Rs. 10,452 million under a concentration standard and Rs. 6,383 million under emissions taxes. We also consider a scenario where surveyed plants are 4.8% of Surat's industrial emissions, in which case net present abatement costs are Rs. 7,839 million under a concentration standard and Rs. 4,788 million under emissions taxes.

If a 66% reduction in baseline emissions were achieved through the current concentration standard regulations, the aggregate cost for the 311 plants in our survey would be Rs. 11.288 million a year, while the same abatement achieved through an emissions tax would cost Rs. 6.894 million a year, a 39% reduction in total costs[8].

Next, we estimate how much our 331 surveyed plants contribute to Surat's industrial emissions, and consider two scenarios where they contribute 3.6% or 4.8% of Surat's emissions. The plants in the CPCB survey were selected because they were the highest emitters of particulate matter in the area, and contribute disproportionately to Surat's industrial pollution. The Gujarat Pollution Control Board categorizes polluting industries as Red (worst), Orange, Green, and White (non-polluting). 99% of the plants in our sample are in the 'Red' category. Gujarat Pollution Control Board data indicates (ENVIS, n.d.) there are 16770 'Red' category units, 6468 'Orange' category, and 4654 'Green category' in Gujarat as of 2011[9]. We estimate Surat has about 5534, 2134, and 1536 industrial plants in the red, orange, and green categories[10]. Thus, Surat's polluting industrial units are about 60% Red, 23% Orange, and 17% Green. Since Red industrial units are more polluting than Orange and Green ones, they generate more than 60%

8. As shown in Table 3.4, under a 150 mg/Nm3 enforced concentration standard, aggregate emissions would be 4369 tons, a reduction of 8617 tons (66%) relative to the baseline of 12986 tons. This cost comes at Rs.1.31 per kg, i.e. Rs. 1310 per ton abated, as shown in Table 3.5, implying a total cost of Rs. 11,288,270, I.e. Rs. 11.288 million per year. The same abatement under an emissions tax would cost Rs. 0.80 per kg, I.e. Rs. 800 per ton abated, implying a total cost of Rs. 6.8936 million per year.

9. The Government of India reports that there are over 41,300 small and medium industrial units in Surat district, of which nearly 24,000 are in textiles, as well as around 800 large industrial units (MSME, n.d.). However, many of the 42100 industrial units in Surat contribute little to the ambient air quality as they are in non-polluting industrial sectors. Hence, we are only concerned with the industrial units in the Gujarat Pollution Control Board's data on polluting plants, and we ignore other non-polluting industrial units in the 'White' category. In Surat overall, about 2% of industrial units are large (800 of 42300). In our data, we have 68% small units, 26% medium units, and 6% large units - so our surveyed plants are larger than a purely random sample would indicate. However, this size difference is small and we ignore it.

10. In 2010 there were 55,170 Small and Medium Enterprises in Gujarat, around 34% of which were in Surat, and 5384 large enterprises, of which around 26% were in Surat. Taking a weighted average of small, medium, and large enterprises, this implies 33% of Gujarat's industrial units are in Surat. We assume that 33% of Gujarat's Red, Orange, and Green industrial units are in Surat

of industrial pollution. Hence, we consider a scenario that Red industries form 80% of industrial pollution. However, it is possible that 'Red' industries are under more scrutiny and have better pollution control, so we also consider a scenario where Red industries form 60% of industrial pollution (assuming the improved pollution control cancels out the higher pollution generation). Our 331 surveyed plants are 6% of the Red industrial units. If Red plants contribute 80% of emissions, then our surveyed plants contribute 4.8% of industrial pollution (6%*80%). If Red plants contribute 60% of baseline emissions, then our surveyed plants contribute 3.6% (6%*60%) of industrial pollution.

If our surveyed plants are 3.6% of Surat's emissions, then aggregate abatement costs in Surat would be Rs. 313.56 million annually under a concentration standard and Rs. 191.5 million annually under an emissions tax, for a 66% reduction in emissions[11]. If our surveyed plants are 4.8% of Surat's emissions, then aggregate abatement costs in Surat would be Rs.235.17 million annually under a concentration standard and Rs. 143.63 million annually under an emissions tax, for a 66% reduction in emissions[12]. For these estimates, we assume that the abatement costs incurred at our surveyed plants are representative of costs at other industrial units.

To estimate a net present value for these annual costs, we use a 3% discount rate, which is the discount rate used to monetize the health benefits (as per Jamison et al. (2013))[13]. For the 3.6% scenario, net present abatement costs are Rs. 10,452 million under a concentration standard and Rs. 6,383 million under emissions taxes. For the 4.8% scenario, net present abatement costs are Rs. 7,839 million under a concentration standard and Rs. 4,788 million under emissions taxes.

11. Rs.11.288 million/0.036=Rs. 313.56 million, Rs. 6.894 million/0.036=Rs.191.5 million.

12. Rs.11.288 million/0.048=Rs. 235.17 million, Rs. 6.894 million/0.048=Rs.143.63 million.

13. We treat the annual costs as a perpetuity and divide it by the discount rate to get the net present value

Table 3.7: Alternate Net Benefits Estimates

	Alternative Net Benefit Estimates		
Increase in Life Expectancy (years)	1.00 (Chen et al. 2013)	0.35 (Correia et al. 2013)	1.80 (Laden et al. 2006)
per 10 microgram/m3 Decrease in PM2.5			
Total Monetized Life Expectancy Benefits in Surat District	Rs. 590,645 million	Rs. 209,584 million	Rs. 1,066,972 million
Concentration Standard:			
Costs under 3.6% Scenario	Rs. 10,452 million	Rs. 10,452 million	Rs. 10,452 million
Costs under 4.8% Scenario	Rs. 7,839 million	Rs. 7,839 million	Rs. 7,839 million
Net Benefits under 3.6% Scenario	Rs. 580,193 million	Rs. 199,132 million	Rs. 1,056,520 million
Net Benefits under 4.8% Scenario	Rs. 582,806 million	Rs. 201,745 million	Rs. 1,059,133 million
Emissions Tax:			
Costs under 3.6% Scenario	Rs. 6,383 million	Rs. 6,383 million	Rs. 6,383 million
Costs under 4.8% Scenario	Rs. 4,788 million	Rs. 4,788 million	Rs. 4,788 million
Net Benefits under 3.6% Scenario	Rs. 584,262 million	Rs. 203,201 million	Rs. 1,060,589 million
Net Benefits under 4.8% Scenario	Rs. 585,857 million	Rs. 204,796 million	Rs. 1,062,184 million

3.6% Scenario: Surveyed Plants are 3.6% of Surat Industrial Emissions. 4.8% Scenario: Surveyed Plants are 4.8% of Surat Industrial Emissions.

3.7.3 Comparing Costs and Benefits

We combine the benefit estimates using Chen et al. (2013), Correia et al. (2013), and Laden et al. (2006) with the cost estimates for the 3.6% and 4.8% scenarios, as shown in Table 3.7. Our central estimates, based on Chen et al. (2013) and holding total emissions reduction fixed at 66%, are that the net benefits under a concentration standard are Rs. 580-583 billion, and the net benefits under an emissions tax are Rs. 584-586 billion. Table 3.7 presents alternate estimates of net benefits, which range from Rs. 199-1062 billion.

Estimated health benefits substantially exceed costs in every scenario we estimate, under every set of assumptions. As a comparison, we note that the United States Environmental Protection Agency retrospectively estimated that the original 1970 Clean Air Act had much larger benefits than costs over the period 1970-1990. Their central estimate of the benefit-cost ratio is over 42 to 1, and across all estimates, the ratio of benefits to costs ranges from 10.7 to 94.5 (EPA 1997). In Appendix Table C.1, we present estimated benefit to cost ratios. Under our central estimates, these are about 57 to 1 - 75 to 1 under a concentration standard, and about 93 to 1 - 123 to 1 under an emissions tax. We note that our central estimates under the emissions concentration standard are within the ratios for the original Clean Air Act, which also relied heavily on command and control regulation.

We also present net benefit estimates for different levels of regulation stringency. We use the same methods described above, which calculated net benefit estimates for a specific 66% reduction in baseline emissions, which corresponds to the expected reductions if status quo regulations were fully enforced. We assume the standards would only be tightened and not relaxed, so we consider emissions reductions of 66% or more. We conservatively assume that our surveyed plants are 3.6% of Surat's emissions, and not 4.8%. We use only the Chen et al. (2013) estimates, and present results for an emissions concentration standard and emissions tax policy. Appendix Table C.2 and Figure 3.11 present the results.

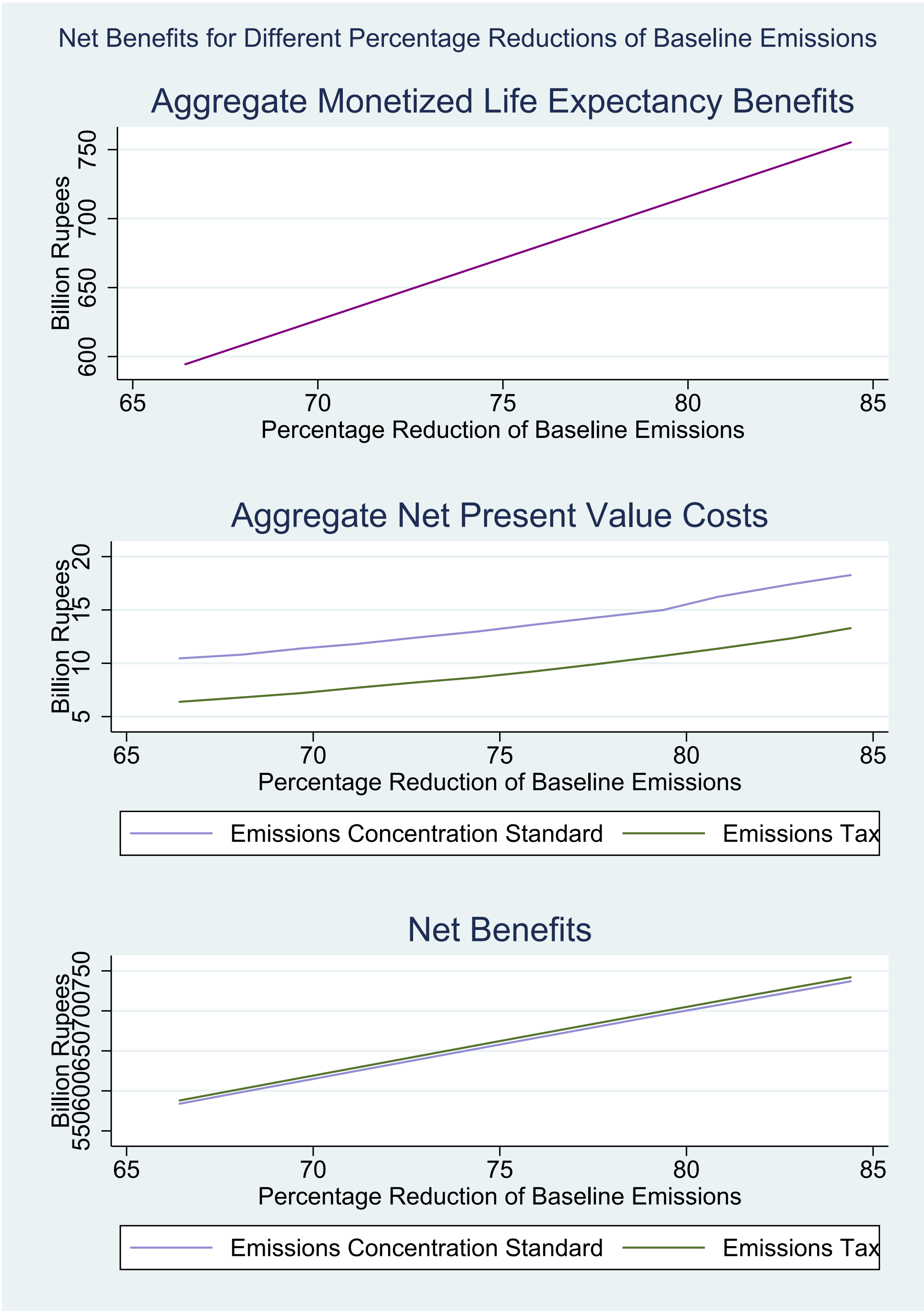

Figure 3.11: Net Benefit Estimates for Different Levels of Emissions Reduction

The monetary net benefits are increasing across the range we study, i.e. for reductions between 66% to 84% of baseline emissions. Our model cannot accurately estimate costs outside this range, as additional stringency would require different types of pollution technologies such as electrostatic precipitators (which are currently not incorporated in our model). However, the results suggest that the regulation stringency should be at least equal to an 84% reduction in baseline emissions.

3.8 Discussion and Conclusion

At current pollution levels, it is clear that the health benefits of particulate matter abatement greatly exceed the abatement costs under all the main policy options, despite the uncertainties in our estimates. Given the relatively low abatement costs, the key barriers to particulate matter abatement in Indian industrial plants are not costs to the polluter, but rather the regulator's constraints regarding bureaucratic capacity, monitoring, and enforcement. The two biggest challenges are monitoring and enforcement. The manpower of the State Pollution Control Boards (SPCBs) is very limited. Bhushan et al. (2009) describe how the number of approved employees at the SPCBs has decreased over time, although the number of industries they regulate has increased by two or three-fold. There are also broader challenges including the political difficulties of reforming the relevant environmental legislation, the slow judicial system, and so on.

The first key challenge is monitoring. Industrial plant pollution is monitored sporadically and manually. Officials from the State Pollution Control Boards (SPCBs) perform inspections of plants, which often but not always include collecting samples of emissions. These inspections could either be routine, or as a response to an industrial plant applying for consent to operate, or as a follow-up to a violation discovered in a previous inspection (Duflo et al. 2014). As our model indicates, plants can thus operate their pollution control equipment while being inspected, but 'bypass' the air pollution control system at other times, and our qualitative research indicates this takes place regularly. True annual emissions may thus be much higher than what is measured through the SPCB inspection process. Moreover, true emissions at baseline may be higher than the baseline emissions we calculate and report in this paper. The intermittent manual sampling regime has other problems as well, due to factors like corruption. The inspection reports from the SPCBs are periodically cross checked with third party audits for a subset of plants. As Duflo et al. (2013) show, the auditing system was corrupted, with 29 percent of audits falsely reporting compliance. After a reform of the audit process, industries reduced emissions by 0.2 standard deviations, with reductions highest among plants with the highest concentrations.

Ongoing reforms should improve monitoring quality drastically. After successful piloting of Continuous Emissions Monitoring Systems (CEMS), the Central Pollution Control Board mandated in 2014 that all plants that fit certain criteria must install these devices, and issued guidelines in 2017[14]. CEMS devices measure pollution at high frequency and immediately transmit the data to the regulator, thus providing high quality real time data. CEMS also allow for appropriate calculation of the emissions load emitted by the polluter over a period of time, as emissions is measured continuously and can be summed over a time period (EPOD, Mcintyre, and Ambroz 2016). The CEMS roll out will take a number of years. Problems are likely to persist, such as internet connectivity problems, device failures, and so on. However, overall, the monitoring challenges are reducing. However, CEMS installations are expensive, and incorporating CEMS costs would substantially add to our estimates of costs faced by plants.

The second key challenge is enforcement. Even under the current imperfect monitoring system, it is evident in our data and previous research (Duflo et al. 2013, 2014) that numerous industrial plants in Gujarat violate environmental regulations, with little consequence. Better identification of high polluters through improved monitoring is insufficient if no action can be taken against them. Some past efforts to improve monitoring in India have been insufficient because of weak enforcement capacity. For instance, although a large literature shows that surprise plant inspections in the USA reduce emissions (Hanna and Oliva 2010; Laplante and Rilstone 1996), Duflo et al. (2014) find that in Gujarat, when the number of regulatory inspections in a randomly selected treatment group of industrial plants was raised to the prescribed minimum— about twice that of status quo - compliance only improved marginally, and there was no effect on average emissions. The increase in inspections did not increase

14. Central Pollution Control Board, Government of India. "Online Monitoring Industrial Emission Effluent". Updated 7 February 2018. http://cpcb.nic.in/Online-Monitoring-Industrial-Emission-Effluent/. Accessed on 26th February 2018.

the detection of extreme polluters or the number of instances of costly penalties on industries, suggesting that regulators are typically aware of the high polluters and inspections in the status quo already target these.

As Ghosh (2015) describes, the current regulatory regime has a number of problems. The environmental regulator - the State Pollution Control Board (SPCB) - cannot levy penalties on non-compliant plants. The power to levy penalties lies with criminal courts. The SPCB must win a criminal case against the non-compliant plant so that the court can levy a penalty. However, criminal prosecution is an ineffective solution as the litigation is a slow process. The SPCBs do have other powers, such as the ability to shut down plant operations or electricity and water access. These too are not easy to implement. As a result, the SPCBs primarily target the highest polluters and focus enforcement resources on them (Duflo et al. 2014). Some reforms to enforcement are being piloted. The Maharashtra Star Ratings Programme, an information disclosure program, publicly reveals the relative pollution emissions of different industrial plants, as measured through the SPCB's manual inspections. Such information disclosure programs have been successful in other settings (Greenstone et al. 2017).

Given the scale of India's air pollution crisis, a dramatic improvement in enforcement is necessary. Even with no other changes to the policy instruments, our estimates indicate that enforcing the status quo regulation would lead to a 66% reduction in aggregate emissions in our Surat sample. If the Surat plants are representative of industrial plants across India, we can make the following rough estimates. Across India, industrial emissions contribute 7.7% percent of fine particulate matter pollution, and the population-weighted annual average $PM_{2.5}$ concentration is 74.3 microgram/m3 (GBD MAPS 2018). A 66% reduction in industrial emissions would thus lead to a $PM_{2.5}$ reduction of about 3.78 microgram/m3, if abatement was implemented proportionate to population. As our cost benefit analysis section shows, the benefits from improved life expectancy are very large for a reduction in $PM_{2.5}$ of this magnitude.

Improving enforcement should be the regulator's key priority, alongside implementation of continuous emissions monitoring systems. However, since improved enforcement will require a political process and new legislation, we argue that a shift to a market based instrument should occur simultaneously as part of the reforms. Evaluating alternate policies, it is clear that moving to market based instruments does indeed offer meaningful cost reductions. Our estimates suggest a 66% reduction in baseline emissions is achievable with a 39% reduction in costs if regulators switch from a concentration standard to a market based instrument. Alternatively, for the same abatement costs of a 66% reduction in emissions under a concentration standard, an 84% reduction in emissions is possible under a market based instrument. Shifting to a different type of command and control regulation, the quantity-based load standard, would increase costs and has no benefit.

Of the two potential market based instruments, we argue that emissions taxes outperform a cap and trade market in this setting, due to the weak enforcement environment with partial compliance. Theory and our empirical results indicate that abatement and costs would be similar under taxes or cap and trade. However, cap and trade depends on a functioning market for permits. In a setting where many plants are non-compliant, there would be minimal incentive to buy and sell permits, and the market might not reach any kind of equilibrium. For cap and trade, enforcement thus involves a specific plant, the regulator, and appropriate enforcement for all other plants as well. While improving enforcement is necessary, this will be a gradual process over a number of years, and partial compliance is likely to persist for a long period. Emissions taxes do not require all plants to be compliant to function, and enforcement only involves a specific plant and the regulator. In our data, many plants are indeed compliant at baseline. Even with no improvement in enforcement, an emissions tax would reduce abatement costs for compliers, while a cap and trade might not do so in a world of imperfect compliance.

The potential for reform is promising. Over the last two decades, the Indian government has reviewed environmental regulation through the appointment of multiple task forces, high-level committees, and external consultants. Many of these have emphasized the need to use market-based regulation and fiscal instruments that align incentives and reduce costs of complying with regulations, following the "polluter pays" principle (MoEFCC 2014). The High Level Committee on Forest and Environment Related Laws observes that "the legislations are weak, monitoring is weaker, and enforcement is weakest". Among the many recommendations in this report, the Committee suggests "reworking standard-setting and revising a system of financial penalties and rewards to proceed to a market-related incentive system" (MoEFCC 2014).

References

Adhvaryu, Achyuta, Namrata Kala, and Anant Nyshadham. 2016. "The Skills to Pay the Bills: Returns to On-the-job Soft Skills Training." *Working Paper.* https://namratakala.files.wordpress.com/2014/01/pace%7B%5C_%7Dnov2016.pdf.

Allcott, Hunt. 2011. "Social norms and energy conservation." *Journal of Public Economics* 95, nos. 9-10 (October): 1082–1095. doi:10.1016/j.jpubeco.2011.03.003.

Allcott, Hunt, and Todd Rogers. 2014. "The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation." *American Economic Review* 104, no. 10 (October): 3003–3037. doi:10.1257/aer.104.10.3003.

André, Michel, and Mario Rapone. 2009. "Analysis and modelling of the pollutant emissions from European cars regarding the driving characteristics and test cycles." *Atmospheric Environment* 43, no. 5 (February): 986–995. doi:10.1016/j.atmosenv.2008.03.013.

Ashraf, Nava, Oriana Bandiera, and Kelsey Jack. 2014. "No margin, no mission? A field experiment on incentives for public service delivery." *Journal of Public Economics* 120 (December): 1–17. doi:10.1016/j.jpubeco.2014.06.014.

Ashraf, Nava, Oriana Bandiera, and Scott Lee. 2016. "Do-gooders and Go-getters: Selection and Performance in Public Service Delivery." *Working Paper.* https://scholar.harvard.edu/files/lee/files/scott%7B%5C_%7Dlee%7B%5C_%7Ddo-gooders%7B%5C_%7Dlatest.pdf.

———. 2018. "Losing Prosociality in the Quest for Talent? Sorting, Selection, and Productivity in the Delivery of Public Services." *Working Paper - STICERD - Economic Organisation and Public Policy Discussion Papers Series, Suntory and Toyota International Centres for Economics and Related Disciplines, LSE.* https://econpapers.repec.org/paper/cepstieop/065.htm.

Ashraf, Nava, Dean Karlan, and Wesley Yin. 2006. "Tying Odysseus to the Mast: Evidence From a Commitment Savings Product in the Philippines." *The Quarterly Journal of Economics* 121, no. 2 (May): 635–672. doi:10.1162/qjec.2006.121.2.635.

Autor, David. 2003. "Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing." *Journal of Labor Economics* 21, no. 1 (January): 1–42. doi:10.1086/344122.

Banerjee, Abhijit V, Rachel Glennerster, and Esther Duflo. 2008. “Putting a Band-Aid on a Corpse: Incentives for Nurses in the Indian Public Health Care System.” *Journal of the European Economic Association* 6 (2-3): 487–500. doi:10.1162/JEEA.2008.6.2-3.487.

Banerjee, Abhijit, Raghabendra Chattopadhyay, Esther Duflo, Daniel Keniston, and Nina Singh. 2012. “Improving Police Performance in Rajasthan, India: Experimental Evidence on Incentives, Managerial Autonomy and Training.” *NBER Working Paper # 17912* (Cambridge, MA) (March). doi:10.3386/w17912.

Banerjee, Abhijit, Rema Hanna, Jordan Kyle, Benjamin Olken, and Sudarno Sumarto. 2016. “Contracting out the Last-Mile of Service Delivery: Subsidized Food Distribution in Indonesia.” *Working Paper.* http://scholar.harvard.edu/files/remahanna/files/160819%7B%5C_%7Dcompetition%7B%5C_%7Dpaper%7B%5C_%7Dms.pdf.

Bank, World. 2013. “An analysis of physical and monetary losses of environmental health and natural resources. Vol. 1 of India - Diagnostic assessment of select environmental challenges.” Washington DC: World Bank. http://documents.worldbank.org/curated/en/2013/06/18009327/india-diagnostic-assessment-select-environmental-challenges-vol-1-3-analysis-physical-monetary-losses-environmental-health-natural-resources.

Barfort, Sebastian, Nikolaj Harmon, Frederik Hjorth, and Asmus Leth Olsen. 2015. “Dishonest and Selection into Public Service in Denmark: Who Runs the World’s Least Public Sector?” *Discussion Papers 15-12, University of Copenhagen, Department of Economics.* http://www.economics.ku.dk/research/publications/wp/dp%7B%5C_%7D2015/1512.pdf/.

Barkenbus, Jack N. 2010. “Eco-driving: An overlooked climate change initiative.” *Energy Policy* 38, no. 2 (February): 762–769. doi:10.1016/j.enpol.2009.10.021.

Bénabou, Roland, and Jean Tirole. 2006. “Incentives and Prosocial Behavior.” *American Economic Review* 96, no. 5 (November): 1652–1678. doi:10.1257/aer.96.5.1652.

Bhattacharya, Soma, Anna Alberini, and Maureen L. Cropper. 2007. “The value of mortality risk reductions in Delhi, India.” *Journal of Risk and Uncertainty* 34, no. 1 (January): 21–47. doi:10.1007/s11166-006-9002-5.

Bhushan, Chandra, Nivit Kumar Yadav, Anil Roy, and Souparno Banerjee. 2009. *Turnaround: Reform agenda for India’s environmental regulators.* New Delhi: Centre for Science / Environment.

Blackman, a., and W. Harrington. 2000. “The Use of Economic Incentives in Developing Countries: Lessons from International Experience with Industrial Air Pollution.” *The Journal of Environment & Development* 9, no. 1 (March): 5–44. doi:10.1177/107049650000900102.

Bucciol, Alessandro, and Marco Piovesan. 2011. “Luck or cheating? A field experiment on honesty with children.” *Journal of Economic Psychology* 32, no. 1 (February): 73–78. doi:10.1016/j.joep.2010.12.001.

Burtraw, Dallas, Alan Krupnick, Erin Mansur, David Austin, and Deirdre Farrell. 1998. "Costs and Benefits of Reducing Air Pollutants Related to Acid Rain." *Contemporary Economic Policy* 16, no. 4 (October): 379–400. doi:10.1111/j.1465-7287.1998.tb00527.x.

Callen, Michael, Saad Gulzar, Ali Hasanain, Yasir Khan, and Arman Rezaee. 2015. "Personalities and Public Sector Performance: Evidence from a Health Experiment in Pakistan." *NBER Working Paper # 21180* (Cambridge, MA) (May). doi:10.3386/w21180.

Carlson, Curtis, Dallas Burtraw, Maureen Cropper, and Karen L. Palmer. 2000. "Sulfur Dioxide Control by Electric Utilities: What Are the Gains from Trade?" *Journal of Political Economy* 108, no. 6 (December): 1292–1326. doi:10.1086/317681.

Chaudhury, Nazmul, Jeffrey Hammer, Michael Kremer, Karthik Muralidharan, and F Halsey Rogers. 2006. "Missing in action: teacher and health worker absence in developing countries." *The journal of economic perspectives : a journal of the American Economic Association* 20 (1): 91–116. doi:10.1257/089533006776526058.

Chen, Yuyu, Avraham Ebenstein, Michael Greenstone, and Hongbin Li. 2013. "Evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River policy." *Proceedings of the National Academy of Sciences of the United States of America* 110, no. 32 (August): 12936–41. doi:10.1073/pnas.1300018110.

Coffey, Diane, and Dean Spears. 2017. *Where India Goes: Abandoned Toilets, Stunted Development, and the Costs of Caste.* India: Harper Collins.

Coria, Jessica, and Thomas Sterner. 2008. "Tradable Permits in Developing Countries: Evidence from Air Pollution in Santiago, Chile." *Resources for The Future Discussion Paper 08-51.*

Correia, Andrew W, C Arden Pope, Douglas W Dockery, Yun Wang, Majid Ezzati, and Francesca Dominici. 2013. "Effect of air pollution control on life expectancy in the United States: an analysis of 545 U.S. counties for the period from 2000 to 2007." *Epidemiology (Cambridge, Mass.)* 24, no. 1 (January): 23–31. doi:10.1097/EDE.0b013e3182770237.

Correia, Sergio. 2015. "Singletons, Cluster-Robust Standard Errors and Fixed Effects: A Bad Mix." *Working Paper,* no. November: 1–5. http://scorreia.com/research/singletons.pdf.

———. 2016. "A Feasible Estimator for Linear Models with Multi-Way Fixed Effects *." *Working Paper,* no. March. http://scorreia.com/research/hdfe.pdf.

Cropper, Maureen, Sarath Guttikunda, Puja Jawahar, Kabir Malik, and Ian Partridge. 2017. "Costs and Benefits of Installing Flue-Gas Desulfurization Units at Coal-Fired Power Plants in India." Chap. 13 in *Disease Control Priorities, Third Edition (Volume 7): Injury Prevention and Environmental Health,* edited by Charles N. Mock, Rachel Nugent, Olive Kobusingye, and Kirk R. Smith, 239–248. The World Bank, October. doi:10.1596/978-1-4648-0522-6.

Dal Bó, Ernesto, Frederico Finan, and Martín A. Rossi. 2013. "Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service*." *The Quarterly Journal of Economics* 128, no. 3 (August): 1169–1218. doi:10.1093/qje/qjt008.

Das, Jishnu, Abhijit Chowdhury, Reshmaan Hussam, and Abhijit V. Banerjee. 2016. "The impact of training informal health care providers in India: A randomized controlled trial." *Science* 354, no. 6308 (October): 7384–7384. doi:10.1126/science.aaf7384.

Das, Jishnu, Jeffrey Hammer, and Kenneth Leonard. 2008. "The Quality of Medical Advice in Low-Income Countries." *Journal of Economic Perspectives* 22, no. 2 (March): 93–114. doi:10.1257/jep.22.2.93.

Das, Jishnu, Alaka Holla, Veena Das, Manoj Mohanan, Diana Tabak, and Brian Chan. 2012. "In Urban And Rural India, A Standardized Patient Study Showed Low Levels Of Provider Training And Huge Quality Gaps." *Health Affairs* 31, no. 12 (December): 2774–2784. doi:10.1377/hlthaff.2011.1356.

Das, Jishnu, Alaka Holla, Aakash Mohpal, and Karthik Muralidharan. 2016. "Quality and Accountability in Health Care Delivery: Audit-Study Evidence from Primary Care in India." *American Economic Review* 106, no. 12 (December): 3765–3799. doi:10.1257/aer.20151138.

Dasgupta, Utteeyo, Subha Mani, Smriti Sharma, and Saurabh Singhal. 2016. "Eliciting Risk Preferences: Fire fighting in the Field." *MPRA Paper No. 69674 Munich Personal RePEc Archive.* https://mpra.ub.uni-muenchen.de/69674/.

De Vlieger, I. 1997. "On board emission and fuel consumption measurement campaign on petrol-driven passenger cars." *Atmospheric Environment* 31, no. 22 (November): 3753–3761. doi:10.1016/S1352-2310(97)00212-4.

Deserranno, Erika. 2017. "Financial Incentives as Signals: Experimental Evidence from the Recruitment of Village Promoters in Uganda." *Working Paper.* https://docs.google.com/viewer?a=v%7B%5C&%7Dpid=sites%7B%5C&%7Dsrcid=ZGVmYXVsdGRvbWFpbnxlcmlrYWRlc2VycmFubm98Z3g6NmYzOTAyNDE1NmUzZTIzNw.

Dhaliwal, Iqbal, and Rema Hanna. 2017. "The devil is in the details: The successes and limitations of bureaucratic reform in India." *Journal of Development Economics* 124 (January): 1–21. doi:10.1016/j.jdeveco.2016.08.008.

Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan. 2013. "Truth-telling by Third-party Auditors and the Response of Polluting Firms: Experimental Evidence from India." *The Quarterly Journal of Economics* 128, no. 4 (October): 1499–1545. doi:10.1093/qje/qjt024.

———. 2014. "The Value of Regulatory Discretion: Estimates from Environmental Inspections in India." *NBER Working Paper # 20590* (Cambridge, MA) (October). doi:10.3386/w20590.

Duflo, Esther, Rema Hanna, and Stephen P Ryan. 2012. "Incentives Work: Getting Teachers to Come to School." *American Economic Review* 102, no. 4 (June): 1241–1278. doi:10.1257/aer.102.4.1241.

Eckel, Catherine C., and Philip J. Grossman. 1996. "Altruism in Anonymous Dictator Games." *Games and Economic Behavior* 16, no. 2 (October): 181–191. doi:10.1006/game.1996.0081.

Eckel, Catherine C., and Philip J. Grossman. 2002. "Sex differences and statistical stereotyping in attitudes toward financial risk." *Evolution and Human Behavior* 23 (4): 281–295. doi:10.1016/S1090-5138(02)00097-1.

———. 2008. "Forecasting risk attitudes: An experimental study using actual and forecast gamble choices." *Journal of Economic Behavior and Organization* 68 (1): 1–17. doi:10.1016/j.jebo.2008.04.006.

ENVIS, ENVIS Centre Gujarat Ecological Commission. n.d. "State of Industrial Environment in Gujarat." http://www.gujenvis.nic.in/PDF/soe-industrial.pdf.

EPA, U.S. Environmental Protection Agency. 1997. *The Benefits and Costs of the Clean Air Act, 1970 to 1990.* https://www.epa.gov/sites/production/files/2015-06/documents/contsetc.pdf.

EPOD, Evidence for Policy Design, Vestal Mcintyre, and Angela Ambroz. 2016. *BCURE Case: Improving industrial monitoring to cut air pollution in India.* Accessed February 20, 2018. https://cems-case.herokuapp.com/index.html%7B%5C#%7Dtop.

Ericsson, Eva. 2001. "Independent driving pattern factors and their influence on fuel-use and exhaust emission factors." *Transportation Research Part D: Transport and Environment* 6, no. 5 (September): 325–345. doi:10.1016/S1361-9209(01)00003-7.

Finan, Frederico, Benjamin A. Olken, and Rohini Pande. 2017. "The Personnel Economics of the Developing State." In *Handbook of Economic Field Experiments,* 2:467–514. Elsevier Ltd. doi:10.1016/bs.hefe.2016.08.001.

Fischbacher, Urs, and Franziska Föllmi-Heusi. 2013. "Lies in Disguise- An Experimental Study on Cheating." *Journal of the European Economic Association* 11, no. 3 (June): 525–547. doi:10.1111/jeea.12014.

Fowlie, Meredith, Stephen P Holland, and Erin T Mansur. 2012. "What Do Emissions Markets Deliver and to Whom? Evidence from Southern California's NO x Trading Program." *American Economic Review* 102, no. 2 (April): 965–993. doi:10.1257/aer.102.2.965.

GBD MAPS, Working Group. 2018. *Burden of Disease Attributable to Major Air Pollution Sources in India.* Boston, MA: Health Effects Institute. https://www.healtheffects.org/system/files/GBD-MAPS-SpecRep21-India-revised%7B%5C_%7D0.pdf.

Ghosh, Shibani. 2015. "Reforming the Liability Regime for Air Pollution in India." *Environmental Law & Practice Review* 4:125–46. http://www.cprindia.org/research/papers/reforming-liability-regime-air-pollution-india.

Goel, Deepti, and Sonam Gupta. 2015. "Delhi Metro and Air Pollution." *Working Paper.* http://www.cdedse.org/pdf/work229.pdf.

GoI, Ministry of Statistics & Programme Implementation. 2018. *First Advance Estimates of National Income, 2017-18.* http://pib.nic.in/newsite/PrintRelease.aspx?relid=175399.

Greenstone, Michael, and Rema Hanna. 2014. "Environmental Regulations, Air and Water Pollution, and Infant Mortality in India." *American Economic Review* 104, no. 10 (October): 3038–3072. doi:10.1257/aer.104.10.3038.

Greenstone, Michael, Janhavi Nilekani, Rohini Pande, Nicholas Ryan, Anant Sudarshan, and Anish Sugathan. 2015. "Lower Pollution, Longer Lives: Life Expectancy Gains if India Reduced Particulate Matter Pollution." *Economic and Political Weekly* 50 (8): 40–46. http://www.epw.in/special-articles/lower-pollution-longer-lives.html.

Greenstone, Michael, Rohini Pande, Nicholas Ryan, and Anant Sudarshan. 2017. "Star Power: Maharashtra Starts Rating Industries by Emissions." *India Spend.* http://www.indiaspend.com/cover-story/star-power-maharashtra-starts-rating-industries-by-emissions-75962.

Gupta, Shreekanth. 2002. "Environmental Benefits and Cost Savings through Market-Based Instruments : An Application Using State-Level Data from India." In *Center for Energy and Environmental Policy Research Working Papers.* Massachusetts Institute of Technology. https://dspace.mit.edu/handle/1721.1/44986?show=full.

Guttikunda, Sarath K., and Puja Jawahar. 2012. "Application of SIM-air modeling tools to assess air quality in Indian cities." *Atmospheric Environment* 62 (December): 551–561. doi:10.1016/j.atmosenv.2012.08.074.

———. 2014. "Atmospheric emissions and pollution from the coal-fired thermal power plants in India." *Atmospheric Environment* 92 (August): 449–460. doi:10.1016/j.atmosenv.2014.04.057.

Guttikunda, Sarath K., and Dinesh Mohan. 2014. "Re-fueling road transport for better air quality in India." *Energy Policy* 68 (May): 556–561. doi:10.1016/j.enpol.2013.12.067.

Guttikunda, Sarath, and Pooja Jawahar. 2011. "Urban Air Pollution Analysis in India." *Simple Interactive Models for Better Air Quality Working Paper.* http://shaktifoundation.in/wp-content/uploads/2017/06/Urban-Air-Pollution-Analysis-in-India.pdf.

Hanna, Rema Nadeem, and Paulina Oliva. 2010. "The Impact of Inspections on Plant-Level Air Emissions." *The B.E. Journal of Economic Analysis & Policy* 10, no. 1 (January). doi:10.2202/1935-1682.1971.

Hanna, Rema, Esther Duflo, and Michael Greenstone. 2016. "Up in Smoke: The Influence of Household Behavior on the Long-Run Impact of Improved Cooking Stoves." *American Economic Journal: Economic Policy* 8, no. 1 (February): 80–114. doi:10.1257/pol.20140008.

Hanna, Rema, and Shing-Yi Wang. 2017. "Dishonesty and Selection into Public Service: Evidence from India." *American Economic Journal: Economic Policy* 9, no. 3 (August): 262–290. doi:10.1257/pol.20150029.

Harrison, Ann, Benjamin Hyman, Leslie Martin, and Shanthi Nataraj. 2015. *When do Firms Go Green? Comparing Price Incentives with Command and Control Regulations in India.* Technical report. Cambridge, MA: National Bureau of Economic Research, November. doi:10.3386/w21763.

HEI, Health Effects Institute. 2017. *State of Global Air 2017.* Technical report. Boston, MA: Health Effects Institute. https://www.stateofglobalair.org/sites/default/files/SoGA2017%7B%5C_%7Dreport.pdf.

Helfand, GE. 1991. "Standards versus standards: the effects of different pollution restrictions." *The American Economic Review* 81 (3): 622–634. http://www.jstor.org/stable/10.2307/2006524.

Houser, Daniel, Stefan Vetter, and Joachim Winter. 2012. "Fairness and cheating." *European Economic Review* 56, no. 8 (November): 1645–1655. doi:10.1016/j.euroecorev.2012.08.001.

IPCC. 2014. *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.* Edited by Core Writing Team R.K. Pachauri and L.A. Meyer (eds.) 1–151. Geneva, Switzerland: IPCC.

Jaffe, Adam B., and Robert N. Stavins. 1995. "Dynamic Incentives of Environmental Regulations: The Effects of Alternative Policy Instruments on Technology Diffusion." *Journal of Environmental Economics and Management* 29, no. 3 (November): S43–S63. doi:10.1006/jeem.1995.1060.

Jamison, Dean T, Lawrence H Summers, George Alleyne, Kenneth J Arrow, Seth Berkley, Agnes Binagwaho, Flavia Bustreo, et al. 2013. "Global health 2035: a world converging within a generation." *The Lancet* 382, no. 9908 (December): 1898–1955. doi:10.1016/S0140-6736(13)62105-4.

Jann, Ben. 2005. "Making regression tables from stored estimates." *The Stata Journal* 5 (3): 288–308. https://www.stata-journal.com/article.html?article=st0085.

———. 2007. "Making regression tables simplified." *The Stata Journal* 7 (2): 227–244. https://www.stata-journal.com/article.html?article=st0085%7B%5C_%7D1.

Kahneman, Daniel. 2003. "Maps of Bounded Rationality: Psychology for Behavioral Economics." *The American Economic Review* 93 (5): 1449–1475. www.jstor.org/stable/3132137.

Kandlikar, Milind, and Gurumurthy Ramachandran. 2000. "The Causes and Consequences of Particulate Air Pollution in Urban India: A Synthesis of the Science." *Annual Review of Energy and the Environment* 25, no. 1 (November): 629–684. doi:10.1146/annurev.energy.25.1.629.

Kautz, Tim, James Heckman, Ron Diris, Bas ter Weel, and Lex Borghans. 2014. "Fostering and Measuring Skills: Improving Cognitive and Non-Cognitive Skills to Promote Lifetime Success." *NBER Working Paper No. 20749* (Cambridge, MA) (December). doi:10.3386/w20749.

Kharola, Pradeep, Geetam Tiwari, and Dinesh Mohan. 2010. "Traffic Safety and City Public Transport System: Case Study of Bengaluru, India." *Journal of Public Transportation* 13, no. 4 (December): 63–91. doi:10.5038/2375-0901.13.4.4.

Kingdon, Geeta Gandhi. 2007. "The progress of school education in India." *Oxford Review of Economic Policy* 23 (2): 168–195. doi:10.1093/icb/grm015.

Kröll, Markus, and Devesh Rustagi. 2016. "Shades of dishonesty and cheating in informal milk markets in India." *SAFE Working Paper Series, No. 134.* http://nbn-resolving.de/urn:nbn:de:hebis:30:3-400667.

Laden, Francine, Joel Schwartz, Frank E Speizer, and Douglas W Dockery. 2006. "Reduction in fine particulate air pollution and mortality: Extended follow-up of the Harvard Six Cities study." *American journal of respiratory and critical care medicine* 173, no. 6 (March): 667–72. doi:10.1164/rccm.200503-443OC.

Lagarde, Mylene, and Duane Blaauw. 2014. "Pro-social preferences and self-selection into jobs: Evidence from South African nurses." *Journal of Economic Behavior & Organization* 107 (November): 136–152. doi:10.1016/j.jebo.2014.09.004.

Laplante, Benoıt, and Paul Rilstone. 1996. "Environmental Inspections and Emissions of the Pulp and Paper Industry in Quebec." *Journal of Environmental Economics and Management* 31, no. 1 (July): 19–36. doi:10.1006/jeem.1996.0029.

Madheswaran, S. 2007. "Measuring the value of statistical life: estimating compensating wage differentials among workers in India." *Social Indicators Research* 84, no. 1 (October): 83–96. doi:10.1007/s11205-006-9076-0.

Malik, Arun S. 1992. "Enforcement Costs and the Choice of Policy Instruments For Controlling Pollution." *Economic Inquiry* 30, no. 4 (October): 714–721. doi:10.1111/j.1465-7295.1992.tb01292.x.

Meier, Stephan. 2007. "A Survey of Economic Theories and Field Evidence on Pro-Social Behavior." In *Economics and Psychology: A Promising New Cross-Disciplinary Field,* edited by Bruno S. Frey and Alois Stutzer, 51–88. Cambridge, MA: MIT Press.

MoEFCC, Ministry of Environment Forests Climate Change. 2014. *Report of the High Level Committee on Forest and Environment Related Laws.* Government of India. http://envfor.nic.in/sites/default/files/press-releases/Final%7B%5C_%7DReport%7B%5C_%7Dof%7B%5C_%7DHLC.pdf.

Mohan, Dinesh, Omer Tsimhoni, Michael Sivak, and Michael J. Flannagan. 2009. *Road Safety in India : Challenges and Opportunities.* Technical report January. University of Michigan Transportation Research Institute.

MSME, Micro Small Medium Enterprises Development Institute. n.d. "Brief Industrial Profile of Surat District." *Government of India.* http://dcmsme.gov.in/dips/BIP%20SURAT%7B%5C_%7Dguj.pdf.

Muller, Nicholas Z., and Robert Mendelsohn. 2009. "Efficient Pollution Regulation: Getting the Prices Right." *American Economic Review* 99, no. 5 (December): 1714–1739. doi:10.1257/aer.99.5.1714.

Muralidharan, Karthik, and Venkatesh Sundararaman. 2011. "Teacher Performance Pay: Experimental Evidence from India." *The Journal of Political Economy* 119, no. 1 (February): 39–77. doi:10.1086/659655.

Narain, Urvashi, and Alan Krupnick. 2007. "The Impact of Delhi's CNG program on Air Quality." *Resources of for the future, Washington, DC,* no. February. http://www.rff.org/RFF/documents/RFF-DP-07-06.pdf.

Oates, WE, PR Portney, and AM McGartland. 1989. "The Net Benefits of Incentive-Based Regulation: A Case Study of Environmental Standard Setting." *The American Economic Review* 79 (5): 1233–1242. http://www.jstor.org/stable/10.2307/1831449.

Pope, C Arden, Majid Ezzati, and Douglas W Dockery. 2009. "Fine-particulate air pollution and life expectancy in the United States." *The New England journal of medicine* 360, no. 4 (January): 376–86. doi:10.1056/NEJMsa0805646.

Ramachandra, T.V., Bharath H. Aithal, and K. Sreejith. 2015. "GHG footprint of major cities in India." *Renewable and Sustainable Energy Reviews* 44, no. January (April): 473–495. doi:10.1016/j.rser.2014.12.036.

Rosenbaum, Stephen Mark, Stephan Billinger, and Nils Stieglitz. 2014. "Let's be honest: A review of experimental evidence of honesty and truth-telling." *Journal of Economic Psychology* 45 (December): 181–196. doi:10.1016/j.joep.2014.10.002.

Singh, Prakarsh, and William A. Masters. 2017. "Impact of caregiver incentives on child health: Evidence from an experiment with Anganwadi workers in India." *Journal of Health Economics* (August). doi:10.1016/j.jhealeco.2017.07.005.

Stavins, Robert N. 2003. "Chapter 9 Experience with market-based environmental policy." Chap. 9 in *Handbook of Environmental Economics Volume 1,* edited by Karl-Goran Maler and Jeffrey R. Vincent, 355–435. Elsevier. doi:10.1016/S1574-0099(03)01014-3.

The World Bank, Global Road Safety Facility, and Institute for Health Metrics & Evaluation IHME. 2014. *Transport for Health: The Global Burden of Disease from Motorized Road Transport.* Technical report. Seattle, WA, Washington DC: IHME, The World Bank. http://documents.worldbank.org/curated/en/984261468327002120/Transport-for-health-the-global-burden-of-disease-from-motorized-road-transport.

Walker, Dilys M., Susanna R. Cohen, Jimena Fritz, Marisela Olvera-García, Sarah T. Zelek, Jenifer O. Fahey, Martín Romero-Martínez, Alejandra Montoya-Rodríguez, and Héctor Lamadrid-Figueroa. 2016. "Impact Evaluation of PRONTO Mexico: A Simulation-Based Program in Obstetric and Neonatal Emergencies and Team Training." *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare* 11, no. 1 (February): 1–9. doi:10.1097/SIH.0000000000000106.

Weitzman, ML. 1974. “Prices vs. Quantities.” *The Review of Economic Studies* 41 (4): 477–491. http://www.jstor.org/stable/10.2307/2296698.

WHO. 2006. *WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide: Global update 2005: Summary of risk assessment.* Technical report. World Health Organization. http://apps.who.int/iris/bitstream/10665/69477/1/WHO%7B%5C_%7DSDE%7B%5C_%7DPHE%7B%5C_%7DOEH%7B%5C_%7D06.02%7B%5C_%7Deng.pdf.

Young, Mark S, Stewart A Birrell, and Neville A Stanton. 2011. “Safe driving in a green world: a review of driver performance benchmarks and technologies to support ’smart’ driving.” *Applied Ergonomics* 42, no. 4 (May): 533–9. doi:10.1016/j.apergo.2010.08.012.

Appendix A

Appendix to Chapter 1

A.1 Appendix Figures and Tables

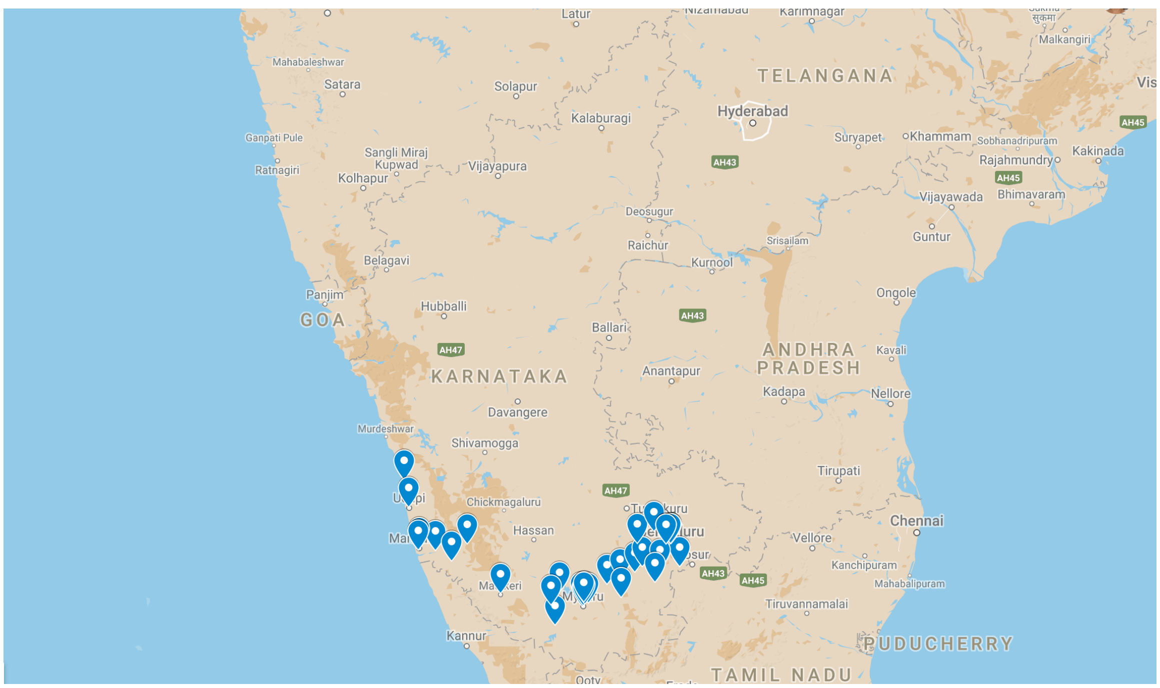

Figure A.1: Participating Bus Depots in Karnataka

Notes: This map was made manually using Google My Maps. The approximate locations of the 34 participating depots are marked.

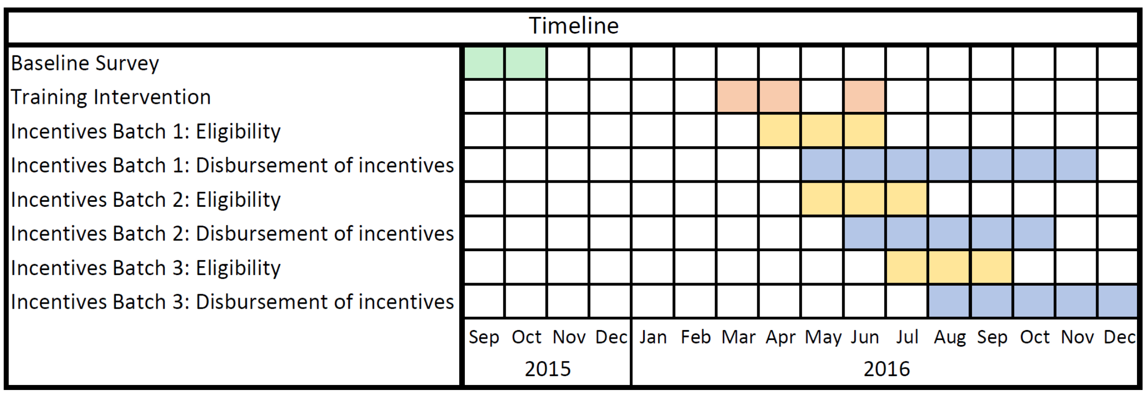

Figure A.2: Timeline of Interventions

Notes: Project Timeline: The baseline survey took place in September and October 2015. Training took place in March, April, and June 2016. There were three batches of incentive eligibility: April-May-June, May-June-July, and July-August-September 2016. The disbursement of incentives typically took place in the three months after incentive eligibility. Thus, a driver who achieved his April target typically received the incentive in May, June, or July.

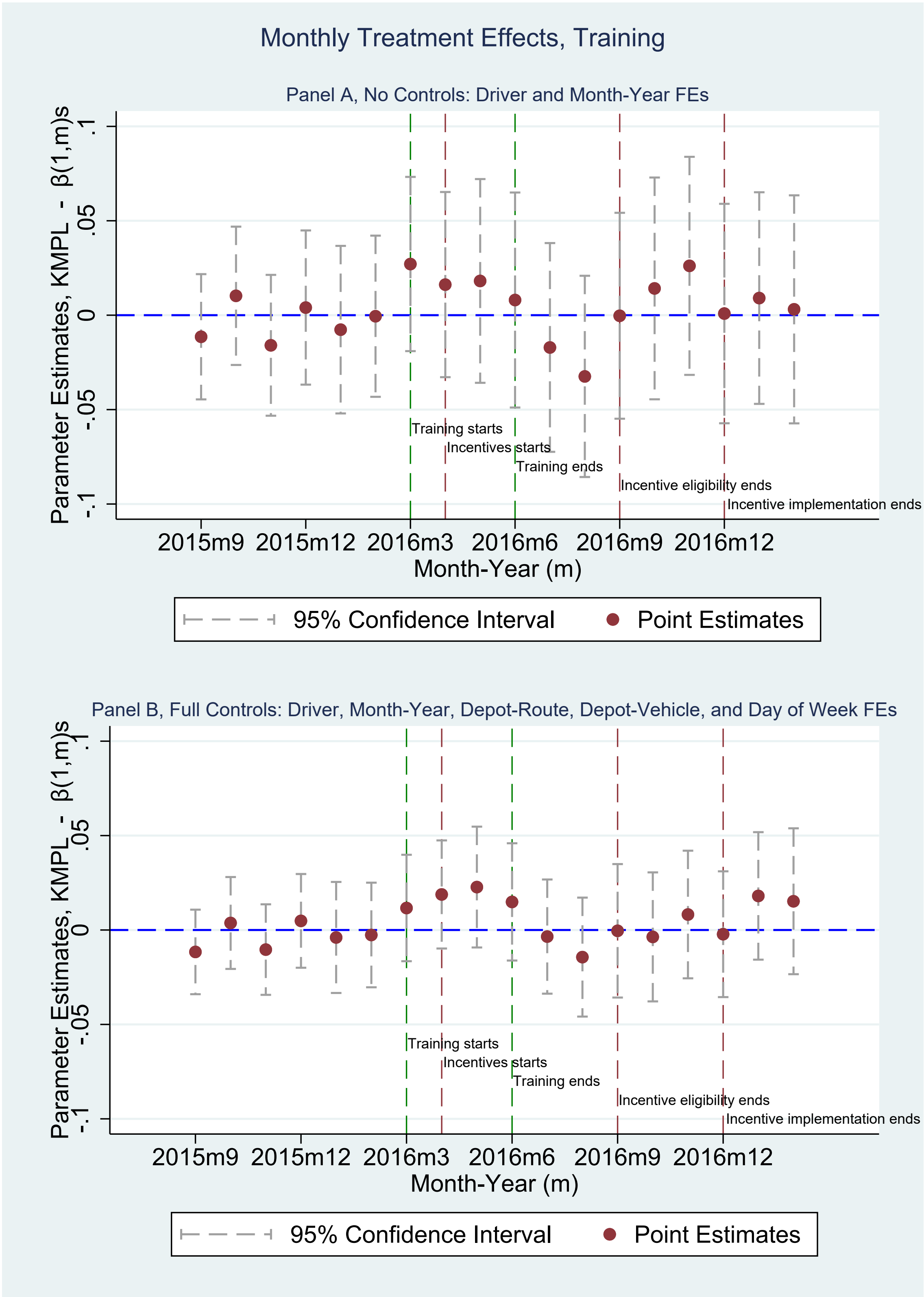

Figure A.3: Monthly Treatment Effects, Training ($\beta_{1,m}s$), With and Without Controls

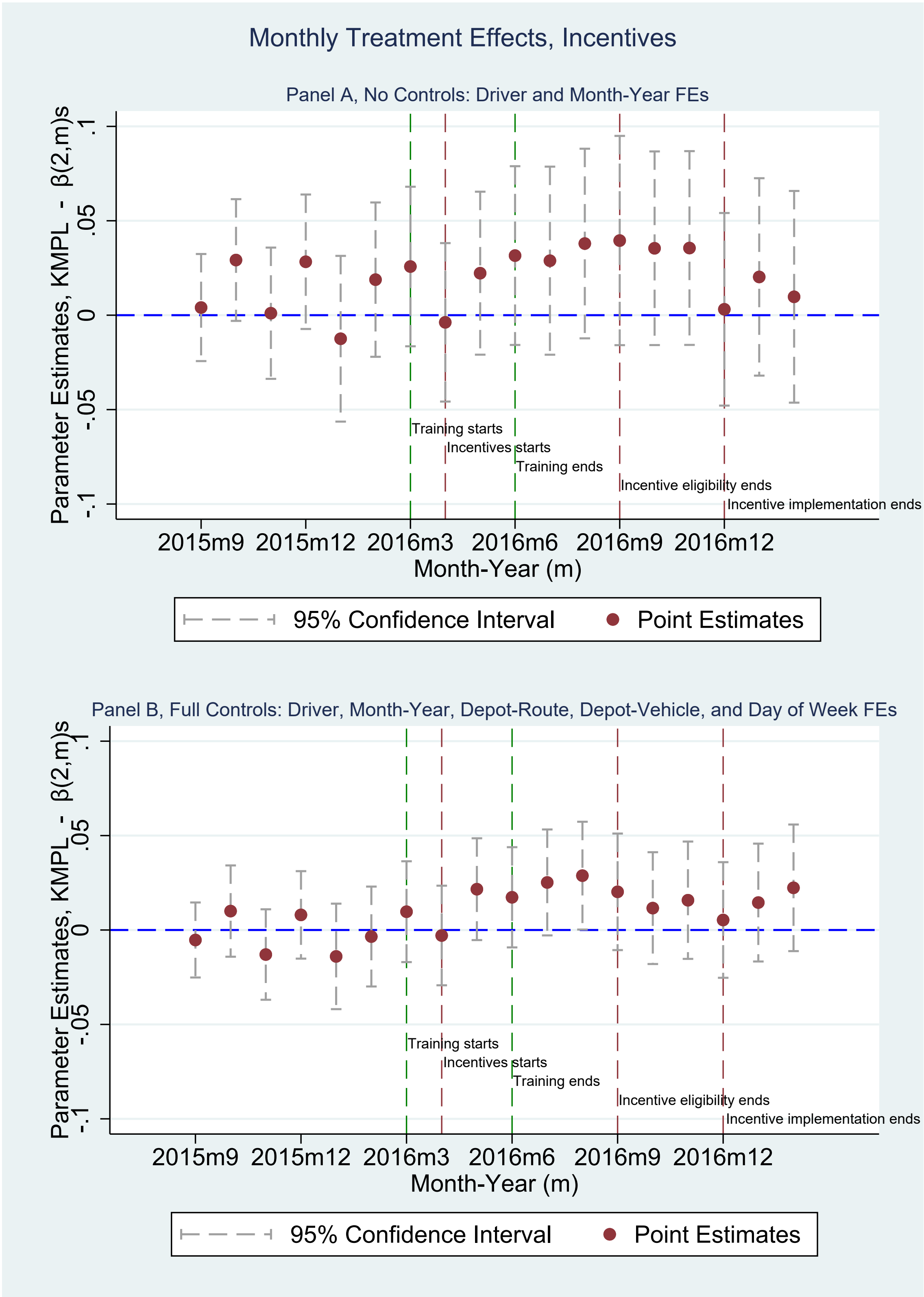

Figure A.4: Monthly Treatment Effects, Incentives ($\beta_{2,m}s$), With and Without Controls

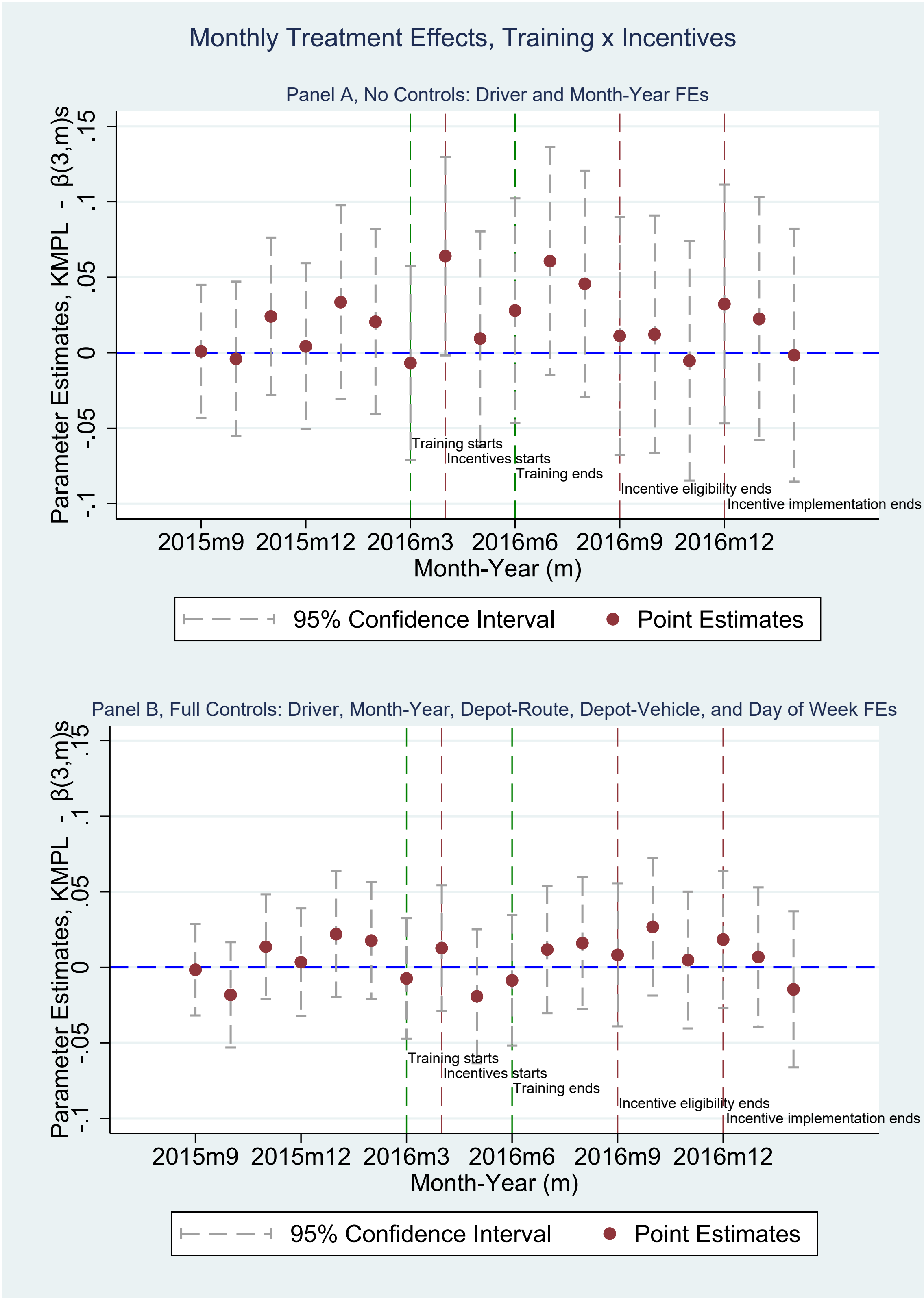

Figure A.5: Monthly Treatment Effects, Training × Incentives ($\beta_{3,m}s$), With and Without Controls

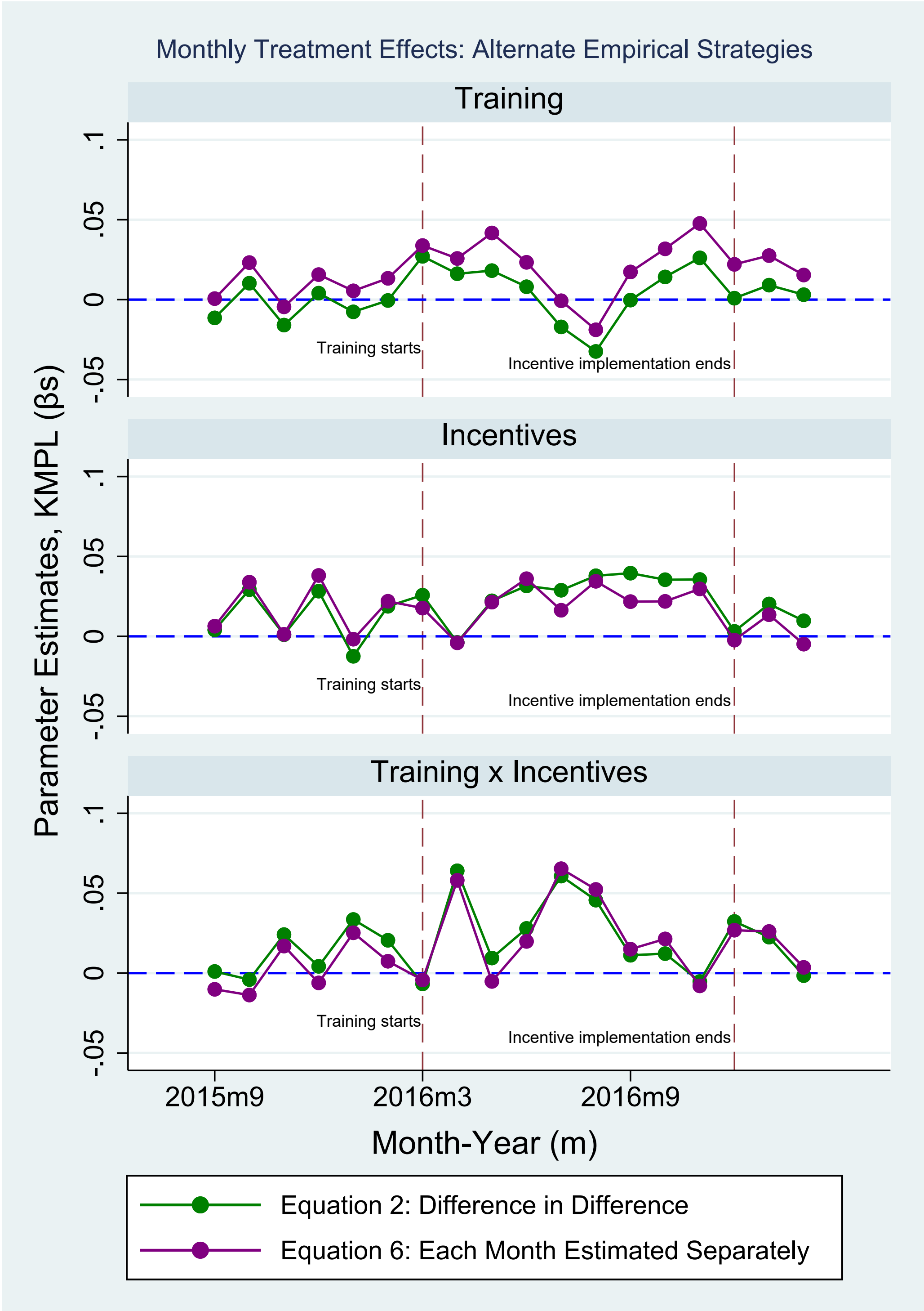

Figure A.6: Monthly Treatment Effects: Robustness Check

Notes: The difference in difference model does not include controls, i.e. Column 1 of Table A.5. Robust standard errors clustered by driver.

Table A.1: Incentives Intervention: Compliance Details

	T2: Incentives Only				T3: Training + Incentives			
	(1) Month1	(2) Month2	(3) Month3	(4) All	(5) Month1	(6) Month2	(7) Month3	(8) All
Individual Months:								
Compliance: Above target, received incentive	0.336	0.296	0.299		0.384	0.353	0.389	
Compliance: Below target, did not receive incentive	0.520	0.551	0.543		0.503	0.505	0.455	
Non-compliance: Above target, did not receive incentive	0.0262	0.0290	0.0289		0.0158	0.0158	0.0184	
Non-compliance: Below target, received incentive	0	0.00525	0		0	0	0	
Attrition: Driver not eligible this month	0.105	0.108	0.105		0.0947	0.116	0.121	
Attrition: Driver not located	0.0131	0.0158	0.0236		0.00263	0.0105	0.0158	
All Three Months:								
Attrition in all three months				0.0787				0.0737
Compliance in all three months				0.756				0.792
Mixture of compliance, non-compliance, and attrition				0.165				0.134
Number of drivers	381	381	381	381	380	380	380	380

Reported coefficients are means for each subgroup. The three incentives batches are pooled together so that months 1, 2, and 3 represent the first, second, and final month of incentive eligibility for each driver.

Table A.2: Post-Attrition Subsample: Balance on Baseline Characteristics

	Section I: Means					Section II: Differences		
	(1) C	(2) T1	(3) T2	(4) T3		(1) Training	(2) Incentives	(3) Tr. × Inc.
Age (years)	38.40	38.56	38.80	38.68	.	0.302	0.282	-0.324
	(8.201)	(8.540)	(8.194)	(8.303)		(0.329)	(0.328)	(0.459)
Education (years)	10.26	10.25	10.30	10.35	.	-0.0194	0.0524	0.0513
	(1.886)	(1.978)	(1.775)	(1.930)		(0.121)	(0.121)	(0.172)
KSRTC tenure (years)	9.563	9.824	9.719	9.662	.	0.352	0.0576	-0.318
	(7.664)	(8.231)	(7.688)	(7.855)		(0.375)	(0.372)	(0.520)
Job satisfaction (score)	4.588	4.475	4.594	4.457	.	-0.0988	0.0185	-0.0253
	(0.810)	(0.903)	(0.764)	(0.888)		(0.0645)	(0.0590)	(0.0891)
Risk aversion (score)	3.011	3.050	2.980	3.179	.	0.0521	-0.0187	0.144
	(1.430)	(1.504)	(1.463)	(1.484)		(0.108)	(0.109)	(0.157)
Pro-socialness (Rs.)	24.95	24.59	24.28	24.51	.	-0.465	-0.739	0.782
	(6.821)	(7.787)	(7.824)	(7.731)		(0.553)	(0.552)	(0.805)
Consistently patient	0.571	0.582	0.605	0.561	.	0.0128	0.0268	-0.0543
	(0.496)	(0.494)	(0.490)	(0.497)		(0.0367)	(0.0371)	(0.0523)
Dice points $>$ 99p	0.0838	0.0967	0.0943	0.0765	.	0.0128	0.00609	-0.0296
	(0.277)	(0.296)	(0.293)	(0.266)		(0.0217)	(0.0214)	(0.0303)
Dice points $>$ 95p	0.184	0.169	0.154	0.193	.	-0.0177	-0.0374	0.0613
	(0.388)	(0.375)	(0.362)	(0.395)		(0.0282)	(0.0281)	(0.0401)
Attended driving school	0.0390	0.0331	0.0227	0.0587	.	-0.00255	-0.0118	0.0393*
	(0.194)	(0.179)	(0.149)	(0.235)		(0.0139)	(0.0130)	(0.0203)
Had major accident	0.145	0.155	0.182	0.185	.	0.00380	0.0350	0.000185
	(0.352)	(0.362)	(0.386)	(0.389)		(0.0266)	(0.0282)	(0.0396)
Had minor accident	0.457	0.506	0.463	0.507	.	0.0545	0.0117	-0.00889
	(0.499)	(0.501)	(0.499)	(0.501)		(0.0375)	(0.0376)	(0.0530)
Attended KSRTC training	0.852	0.857	0.872	0.855	.	0.00438	0.0198	-0.0218
	(0.355)	(0.351)	(0.334)	(0.353)		(0.0263)	(0.0259)	(0.0368)
Number of drivers	359	363	352	358	.			
.					.			
Joint F-Test Statistic	.	0.688	0.915	1.210	.			
P-Value	.	0.776	0.537	0.267	.			

This table reproduces the balance checks on baseline characteristics in Table 1.1 for the post-attrition subsample used for analysis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Section I: reported coefficients are means for groups C(Control), T1(Training Only), T2(Incentives Only), and T3(Training+Incentives). Standard deviations are in parentheses. Section II: reported coefficients are from regressions of each variable on Training, Incentives, Training × Incentives (Tr × Inc), and strata fixed effects. Robust standard errors are in parentheses. For 'Attended KSRTC training' strata fixed effects are not included. The joint hypothesis tests are from separate regressions of each treatment group compared to the control group. The treatment indicator is regressed on all baseline characteristics. I test the joint hypothesis that all coefficients are equal to zero.

Table A.3: Post-Attrition Subsample: Take-up and Compliance

	Section I: Means					Section II: Differences		
	(1)	(2)	(3)	(4)		(1)	(2)	(3)
	C	T1	T2	T3		Training	Incentives	Tr. × Inc.
Panel A: Take-Up:					.			
Took up training	0.00557	0.945	0.0142	0.958	.	0.937***	0.00963	0.00505
	(0.0745)	(0.228)	(0.119)	(0.201)		(0.0129)	(0.00756)	(0.0178)
Took up incentives	0	0	0.932	0.939	.	0.0000689	0.935***	0.00427
	(0)	(0)	(0.252)	(0.240)		(0.00395)	(0.0129)	(0.0182)
Panel B: Compliance:					.			
Complied with training	0.994	0.945	0.986	0.958	.	-0.0496***	-0.00835	0.0220
	(0.0745)	(0.228)	(0.119)	(0.201)		(0.0129)	(0.00756)	(0.0178)
Complied with incentives	1	1	0.932	0.939	.	0.0000689	-0.0648***	0.00427
	(0)	(0)	(0.252)	(0.240)		(0.00395)	(0.0129)	(0.0182)
Number of drivers	359	363	352	358	.			

This table reproduces the balance checks on take-up and compliance in Table 1.2 for the post-attrition subsample used for analysis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Section I: reported coefficients are means for groups C (Control), T1 (Training Only), T2 (Incentives Only), and T3 (Training+Incentives). Standard deviations are in parentheses. Section II: reported coefficients are from regressions of each variable on Training, Incentives, Training × Incentives (Tr × Inc), and strata fixed effects. Robust standard errors are in parentheses.

Table A.4: Treatment Groups: Balance on Batch Assignment and Training Type

	Section I: Means				Section II: P-Values		
	(1) T1	(2) T2	(3) T3	(4) Overall	(5) (1) vs. (2)	(6) (1) vs. (3)	(7) (2) vs. (3)
Batch:							
In Batch 1	0.48	0.43	0.43	0.45	0.16	0.17	0.97
	(0.03)	(0.03)	(0.03)	(0.01)			
In Batch 2	0.34	0.39	0.39	0.38	0.19	0.18	0.98
	(0.02)	(0.03)	(0.03)	(0.01)			
In Batch 3	0.17	0.18	0.18	0.18	0.86	0.92	0.94
	(0.02)	(0.02)	(0.02)	(0.01)			
N	380	381	380	1141			
Training:							
Haneef	0.75		0.81	0.78		0.06	
	(0.02)		(0.02)	(0.02)			
PCRA	0.19		0.15	0.17		0.15	
	(0.02)		(0.02)	(0.01)			
NA	0.06		0.04	0.05		0.26	
	(0.01)		(0.01)	(0.01)			
N	380		380	760			

Columns 1-4 present means, with standard errors in parentheses. Columns 5-7 present the p-values from a t-test of the differences of the means of each indicator variable (Batch 1, Haneef, etc) across each treatment group (T1, T2, T3). NA indicates the driver didn't attend training.

Table A.5: Impact on Fuel Efficiency in Each Month

	(1) KMPL	(2) KMPL	(3) KMPL	(4) KMPL
Training $\times 1(m = Sep2015)$	-0.0114	-0.00738	-0.0118	-0.0116
	(0.0169)	(0.0128)	(0.0114)	(0.0114)
Training $\times 1(m = Oct2015)$	0.0103	0.00378	0.00377	0.00372
	(0.0187)	(0.0147)	(0.0124)	(0.0124)
Training $\times 1(m = Nov2015)$	-0.0159	-0.0117	-0.0106	-0.0104
	(0.0190)	(0.0151)	(0.0122)	(0.0122)
Training $\times 1(m = Dec2015)$	0.00404	0.00606	0.00503	0.00484
	(0.0208)	(0.0161)	(0.0127)	(0.0127)
Training $\times 1(m = Jan2016)$	-0.00766	-0.00788	-0.00402	-0.00395
	(0.0226)	(0.0179)	(0.0150)	(0.0150)
Training $\times 1(m = Feb2016)$	-0.000546	-0.00172	-0.00291	-0.00264
	(0.0218)	(0.0168)	(0.0141)	(0.0141)
Training $\times 1(m = Mar2016)$	0.0271	0.0186	0.0114	0.0117
	(0.0235)	(0.0187)	(0.0144)	(0.0144)
Training $\times 1(m = Apr2016)$	0.0162	0.0251	0.0188	0.0188
	(0.0250)	(0.0189)	(0.0146)	(0.0146)
Training $\times 1(m = May2016)$	0.0181	0.0275	0.0226	0.0227
	(0.0275)	(0.0208)	(0.0163)	(0.0163)
Training $\times 1(m = Jun2016)$	0.00804	0.0308	0.0144	0.0149
	(0.0290)	(0.0195)	(0.0158)	(0.0158)
Training $\times 1(m = Jul2016)$	-0.0171	0.0170	-0.00366	-0.00347
	(0.0282)	(0.0196)	(0.0154)	(0.0154)
Training $\times 1(m = Aug2016)$	-0.0324	0.00295	-0.0146	-0.0144
	(0.0272)	(0.0197)	(0.0161)	(0.0161)
Training $\times 1(m = Sep2016)$	-0.000300	0.00606	-0.000569	-0.000433
	(0.0278)	(0.0219)	(0.0180)	(0.0180)
Training $\times 1(m = Oct2016)$	0.0142	0.0138	-0.00399	-0.00365
	(0.0299)	(0.0211)	(0.0174)	(0.0174)
Training $\times 1(m = Nov2016)$	0.0261	0.0213	0.00805	0.00821
	(0.0294)	(0.0208)	(0.0172)	(0.0172)
Training $\times 1(m = Dec2016)$	0.000877	0.00641	-0.00251	-0.00222
	(0.0296)	(0.0206)	(0.0170)	(0.0170)
Training $\times 1(m = Jan2017)$	0.00907	0.0253	0.0179	0.0181
	(0.0286)	(0.0204)	(0.0172)	(0.0172)
Training $\times 1(m = Feb2017)$	0.00310	0.0100	0.0150	0.0152
	(0.0308)	(0.0235)	(0.0197)	(0.0197)
Incentives $\times 1(m = Sep2015)$	0.00405	-0.000210	-0.00543	-0.00529
	(0.0145)	(0.0110)	(0.0101)	(0.0101)

Table continued on next page

Table A.5: (Continued)

	(1) KMPL	(2) KMPL	(3) KMPL	(4) KMPL
Incentives $\times 1(m = Oct2015)$	0.0292*	0.0164	0.00984	0.0100
	(0.0164)	(0.0138)	(0.0123)	(0.0123)
Incentives $\times 1(m = Nov2015)$	0.00104	-0.00176	-0.0133	-0.0130
	(0.0177)	(0.0145)	(0.0122)	(0.0122)
Incentives $\times 1(m = Dec2015)$	0.0283	0.0193	0.00796	0.00801
	(0.0182)	(0.0140)	(0.0118)	(0.0118)
Incentives $\times 1(m = Jan2016)$	-0.0125	-0.00554	-0.0140	-0.0140
	(0.0224)	(0.0191)	(0.0142)	(0.0142)
Incentives $\times 1(m = Feb2016)$	0.0188	0.0143	-0.00343	-0.00347
	(0.0208)	(0.0151)	(0.0135)	(0.0135)
Incentives $\times 1(m = Mar2016)$	0.0258	0.0273*	0.00978	0.00969
	(0.0215)	(0.0165)	(0.0136)	(0.0136)
Incentives $\times 1(m = Apr2016)$	-0.00378	0.0123	-0.00231	-0.00291
	(0.0214)	(0.0161)	(0.0134)	(0.0134)
Incentives $\times 1(m = May2016)$	0.0222	0.0367**	0.0216	0.0216
	(0.0220)	(0.0167)	(0.0137)	(0.0137)
Incentives $\times 1(m = Jun2016)$	0.0316	0.0450***	0.0170	0.0173
	(0.0241)	(0.0175)	(0.0135)	(0.0135)
Incentives $\times 1(m = Jul2016)$	0.0289	0.0520***	0.0250*	0.0252*
	(0.0254)	(0.0182)	(0.0143)	(0.0143)
Incentives $\times 1(m = Aug2016)$	0.0379	0.0616***	0.0288**	0.0288**
	(0.0256)	(0.0193)	(0.0145)	(0.0146)
Incentives $\times 1(m = Sep2016)$	0.0395	0.0557**	0.0202	0.0202
	(0.0283)	(0.0216)	(0.0157)	(0.0157)
Incentives $\times 1(m = Oct2016)$	0.0354	0.0445**	0.0115	0.0116
	(0.0261)	(0.0190)	(0.0151)	(0.0151)
Incentives $\times 1(m = Nov2016)$	0.0356	0.0495**	0.0162	0.0158
	(0.0262)	(0.0195)	(0.0158)	(0.0158)
Incentives $\times 1(m = Dec2016)$	0.00312	0.0325*	0.00535	0.00531
	(0.0260)	(0.0193)	(0.0156)	(0.0156)
Incentives $\times 1(m = Jan2017)$	0.0203	0.0539***	0.0148	0.0145
	(0.0266)	(0.0198)	(0.0159)	(0.0159)
Incentives $\times 1(m = Feb2017)$	0.00974	0.0490**	0.0226	0.0223
	(0.0286)	(0.0223)	(0.0171)	(0.0171)
Training $\times$ Incentives $\times 1(m = Sep2015)$	0.00100	-0.00522	-0.00154	-0.00164
	(0.0225)	(0.0171)	(0.0154)	(0.0154)
Training $\times$ Incentives $\times 1(m = Oct2015)$	-0.00406	-0.0108	-0.0181	-0.0183
	(0.0261)	(0.0204)	(0.0178)	(0.0178)

Table continued on next page

Table A.5: (Continued)

	(1) KMPL	(2) KMPL	(3) KMPL	(4) KMPL
Training × Incentives ×1($m = Nov2015$)	0.0241	0.0174	0.0141	0.0136
	(0.0266)	(0.0210)	(0.0177)	(0.0177)
Training × Incentives ×1($m = Dec2015$)	0.00425	0.00526	0.00334	0.00345
	(0.0281)	(0.0218)	(0.0181)	(0.0181)
Training × Incentives ×1($m = Jan2016$)	0.0336	0.0194	0.0219	0.0220
	(0.0327)	(0.0268)	(0.0213)	(0.0213)
Training × Incentives ×1($m = Feb2016$)	0.0205	0.0104	0.0179	0.0176
	(0.0313)	(0.0235)	(0.0198)	(0.0198)
Training × Incentives ×1($m = Mar2016$)	-0.00674	-0.0175	-0.00732	-0.00738
	(0.0326)	(0.0248)	(0.0204)	(0.0204)
Training × Incentives ×1($m = Apr2016$)	0.0640*	0.0189	0.0122	0.0127
	(0.0335)	(0.0261)	(0.0212)	(0.0212)
Training × Incentives ×1($m = May2016$)	0.00945	-0.0138	-0.0192	-0.0192
	(0.0362)	(0.0278)	(0.0226)	(0.0226)
Training × Incentives ×1($m = Jun2016$)	0.0280	-0.0189	-0.00838	-0.00867
	(0.0379)	(0.0275)	(0.0220)	(0.0220)
Training × Incentives ×1($m = Jul2016$)	0.0607	0.00192	0.0119	0.0118
	(0.0385)	(0.0272)	(0.0215)	(0.0215)
Training × Incentives ×1($m = Aug2016$)	0.0457	-0.00358	0.0164	0.0160
	(0.0383)	(0.0280)	(0.0223)	(0.0223)
Training × Incentives ×1($m = Sep2016$)	0.0112	-0.0142	0.00813	0.00823
	(0.0401)	(0.0312)	(0.0241)	(0.0242)
Training × Incentives ×1($m = Oct2016$)	0.0122	0.00429	0.0266	0.0268
	(0.0401)	(0.0285)	(0.0232)	(0.0232)
Training × Incentives ×1($m = Nov2016$)	-0.00527	-0.0194	0.00436	0.00479
	(0.0405)	(0.0283)	(0.0231)	(0.0231)
Training × Incentives ×1($m = Dec2016$)	0.0323	0.0000354	0.0182	0.0184
	(0.0403)	(0.0284)	(0.0232)	(0.0233)
Training × Incentives ×1($m = Jan2017$)	0.0225	-0.0212	0.00626	0.00682
	(0.0411)	(0.0290)	(0.0235)	(0.0235)
Training × Incentives ×1($m = Feb2017$)	-0.00159	-0.0283	-0.0149	-0.0146
	(0.0427)	(0.0325)	(0.0264)	(0.0263)
Observations (shifts)	384241	381612	381375	381375
Number of clusters (drivers)	1432	1432	1432	1432
Driver fixed effects	Yes	Yes	Yes	Yes
Month × year fixed effects	Yes	Yes	Yes	Yes
Depot × route fixed effects		Yes	Yes	Yes
Depot × vehicle fixed effects			Yes	Yes
Day of week fixed effects				Yes

Table continued on next page

Table A.5: (Continued)

	(1) KMPL	(2) KMPL	(3) KMPL	(4) KMPL

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered by driver in parentheses. KMPL is winsorized at 1% and 99%. The reference group is all shifts from January-August 2015. Reported coefficients are interactions between the treatments and indicators for month $\times$ year.

Table A.6: Robustness Check: Impact using Post-Intervention Data Only

	(1) KMPL	(2) KMPL	(3) KMPL	(4) KMPL
Training	0.0185	0.0103	0.00572	0.00549
	(0.0161)	(0.00732)	(0.00716)	(0.00717)
Incentives	0.0141	0.00342	0.000936	0.000742
	(0.0151)	(0.00663)	(0.00649)	(0.00650)
Training × Incentives	0.0203	0.0103	0.0166*	0.0169*
	(0.0228)	(0.0100)	(0.00982)	(0.00983)
Baseline KMPL	0.694***	0.157***	0.129***	0.129***
	(0.0280)	(0.0124)	(0.0125)	(0.0125)
Observations (shifts)	167948	166579	166361	166361
Clusters (drivers)	1359	1356	1356	1356
Strata fixed effects	Yes	Yes	Yes	Yes
Month × year fixed effects	Yes	Yes	Yes	Yes
Depot × route fixed effects		Yes	Yes	Yes
Depot × vehicle fixed effects			Yes	Yes
Day of week fixed effects				Yes

$p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered by driver in parentheses. KMPL is winsorized at 1% and 99%. Only post-intervention shifts ($m >= March 2016$) are used. Baseline KMPL is the average of KMPL in all pre-intervention shifts ($m <= February 2016$).

Table A.7: Robustness Check: Dropping Shifts with Multiple Drivers

	All Shifts	Shifts with single:	
	(1)	(2)	(3)
	-	RCT driver	driver
Panel A: Without Controls			
Training × Post	0.00787	0.00850	0.00871
	(0.0189)	(0.0190)	(0.0198)
Incentives × Post	0.0190	0.0209	0.0201
	(0.0163)	(0.0164)	(0.0170)
Training × Incentives × Post	0.0176	0.0171	0.0191
	(0.0255)	(0.0257)	(0.0267)
Observations (shifts)	384241	379420	362260
Number of clusters (drivers)	1432	1431	1411
Panel B: With Controls			
Training × Post	0.00906	0.00898	0.00744
	(0.0102)	(0.0103)	(0.0107)
Incentives × Post	0.0168**	0.0174**	0.0178**
	(0.00829)	(0.00840)	(0.00878)
Training × Incentives × Post	0.000948	0.000421	0.00234
	(0.0136)	(0.0137)	(0.0142)
Observations (shifts)	381375	376551	359370
Number of clusters (drivers)	1432	1431	1408

$p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered by driver in parentheses. KMPL is winsorized at 1% and 99%. Post is an indicator for all shifts from March 1 2016 onwards. In Column 2, all shifts with 2 RCT-participant drivers are dropped. In Column 3, all shifts with 2 or more drivers (including non-participants) are dropped. All specifications include driver and month × year fixed effects. The control variables in Panel B are depot × vehicle, depot × route, and day of week fixed effects.

A.2 Event Time Approach

The main paper uses a single treatment date for all drivers. In this appendix, I use a complementary approach to further explore the time patterns of the treatment effects. Since the treatments were implemented in three batches (see the timeline in Appendix Figure A.2 and Section 1.3.2), I also consider a difference in difference regression with multiple treatment dates. I include an event-time variable, following Autor 2003, Greenstone and Hanna 2014, and Hanna, Duflo, and Greenstone 2016. I estimate

$$\begin{aligned} Y_{imd\tau} &= \beta_1 Training_d \times Post_{i\tau} + \beta_2 Incentives_d \times Post_{i\tau} \\ &\quad + \beta_3 Training_d \times Incentives_d \times Post_{i\tau} + \alpha_d + \delta_m + Controls_{imd} + \epsilon_{imd\tau} \end{aligned} \tag{A.1}$$

Where $Y_{imd\tau}$ is kilometers per liter (KMPL) for shift i in month-year m for driver d in event-month τ. The event-time variable τ measures the number of months since the intervention, such that $\tau = 0$ in the month training is assigned (groups T1 and T3), $\tau = 1$ in the first month incentives are assigned (groups T2 and T3), and $\tau = 0$ for all months for the control group C. $Post_{i\tau}$ is an indicator for any shift i in event-month $\tau >= 0$, and the rest is as in Equation 1.1. Results are presented in Appendix Table A.8.

I similarly estimate monthly treatment effects in event-time as follows

$$\begin{aligned} Y_{imd\tau} &= \sum_{\tau=-6}^{\tau=11} \beta_{1,\tau}(Training_d \times I_\tau) + \sum_{\tau=-6}^{\tau=11} \beta_{2,\tau}(Incentives_d \times I_\tau) \\ &\quad + \sum_{\tau=-6}^{\tau=11} \beta_{3,\tau}(Training_d \times Incentives_d \times I_\tau) + \alpha_d + \delta_m + Controls_{imd} + \epsilon_{imd\tau} \end{aligned} \tag{A.2}$$

Where I_τ is a set of indicator variables for event-month τ, and the rest is as in Equation 1.2. I include the six months prior to the intervention as a placebo test. The $\beta_{1,\tau}$s, $\beta_{2,\tau}$s, and $\beta_{3,\tau}$s measure the impact of training, incentives, and the interaction respectively, for the 6 months leading up to training, the month of training, and the 11 subsequent months. The β's are the average difference between treatment and control groups in event-month τ, relative to the average difference 7 or more months before training (i.e. the reference period is $\tau \in [-17, -7]$). Results are graphed in Appendix Figures A.7, A.8, and A.9.

As described in Section 1.3.2, T2 drivers were randomly assigned to batches but T1 and

T3 (training group) drivers were not. Thus, there may be some selection bias in the batches. For instance, managers may have sent eager drivers first, or preferentially sent drivers they perceived to be more in need of training. Equation A.1 pools the three batches together, so that the comparison is between all treatment group drivers and the control group. As drivers were randomly assigned to treatments, the control group drivers provide an appropriate counter factual and the event time βs should provide an estimate of the causal effect of treatment. Appendix Figures A.7, A.8, and A.9 demonstrate that the parallel trend assumption is satisfied and there are no differential trends in the pre-intervention period. Due to the timing of the batches, while the $\beta_{1/2/3,\tau}$s are identified from all treatment group drivers for $\tau \in [-6, 8]$, the $\beta_{1/2/3,\tau}$s are identified from the first two batches for $\tau \in [9, 10]$, and $\beta_{1/2/3,\tau=11}$ is identified only from the first batch. As there may some selection bias for the batches, the βs for $\tau \in [9, 11]$ may be biased estimates.

Overall, results are qualitatively similar with both the single treatment date and multiple treatment date approaches. The time patterns seen in the main paper for training and incentives are more clear with the event-time approach. The training program has a small positive impact in the short term for 3-4 months after training, statistically significant with controls. There is a clear time pattern for the incentives scheme, showing that the magnitude of the treatment effect sets in gradually, continues while implementation is ongoing, and ends once the incentive implementation ends. In the event-time method the monthly treatment effects are statistically significant with controls.

Table A.8: Impact on Fuel Efficiency using a Difference in Difference in Event-Time

	(1)	(2)	(3)	(4)
	KMPL	KMPL	KMPL	KMPL
Training × Post	0.0133	0.0267**	0.0185*	0.0187*
	(0.0181)	(0.0118)	(0.0104)	(0.0104)
Incentives × Post	0.0246*	0.0449***	0.0273***	0.0272***
	(0.0146)	(0.00956)	(0.00788)	(0.00789)
Training × Incentives × Post	0.0143	-0.0173	-0.00911	-0.00888
	(0.0250)	(0.0164)	(0.0139)	(0.0139)
Observations (shifts)	384241	381612	381375	381375
Number of clusters (drivers)	1432	1432	1432	1432
Driver fixed effects	Yes	Yes	Yes	Yes
Month × year fixed effects	Yes	Yes	Yes	Yes
Depot × route fixed effects		Yes	Yes	Yes
Depot × vehicle fixed effects			Yes	Yes
Day of week fixed effects				Yes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Robust standard errors clustered by driver in parentheses

KMPL is winsorized at 1% and 99%

Post is an indicator for $Y_{imd\tau} \in [\tau >= 0]$

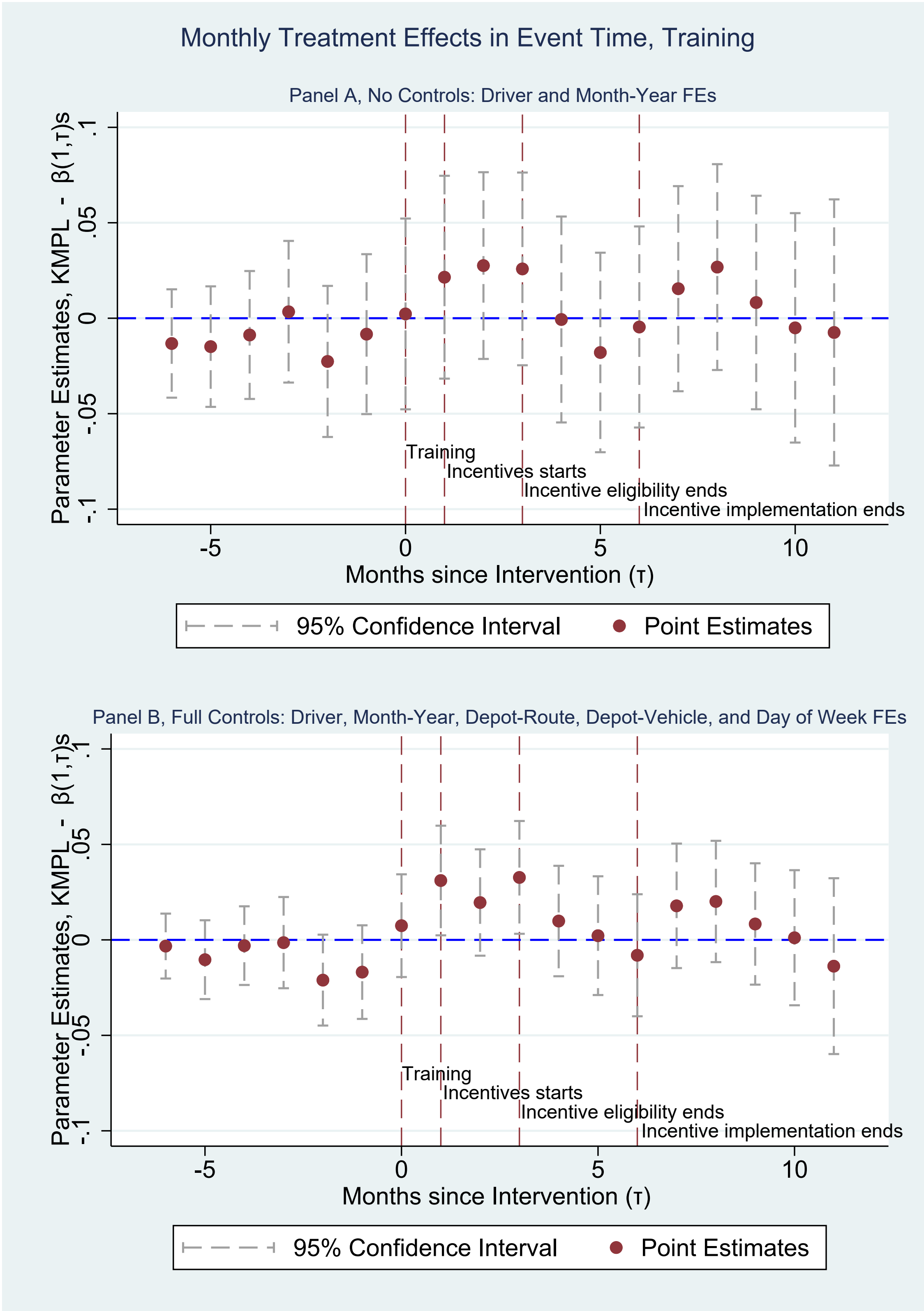

Figure A.7: Monthly Treatment Effects in Event Time, Training ($\beta_{1,\tau}s$)

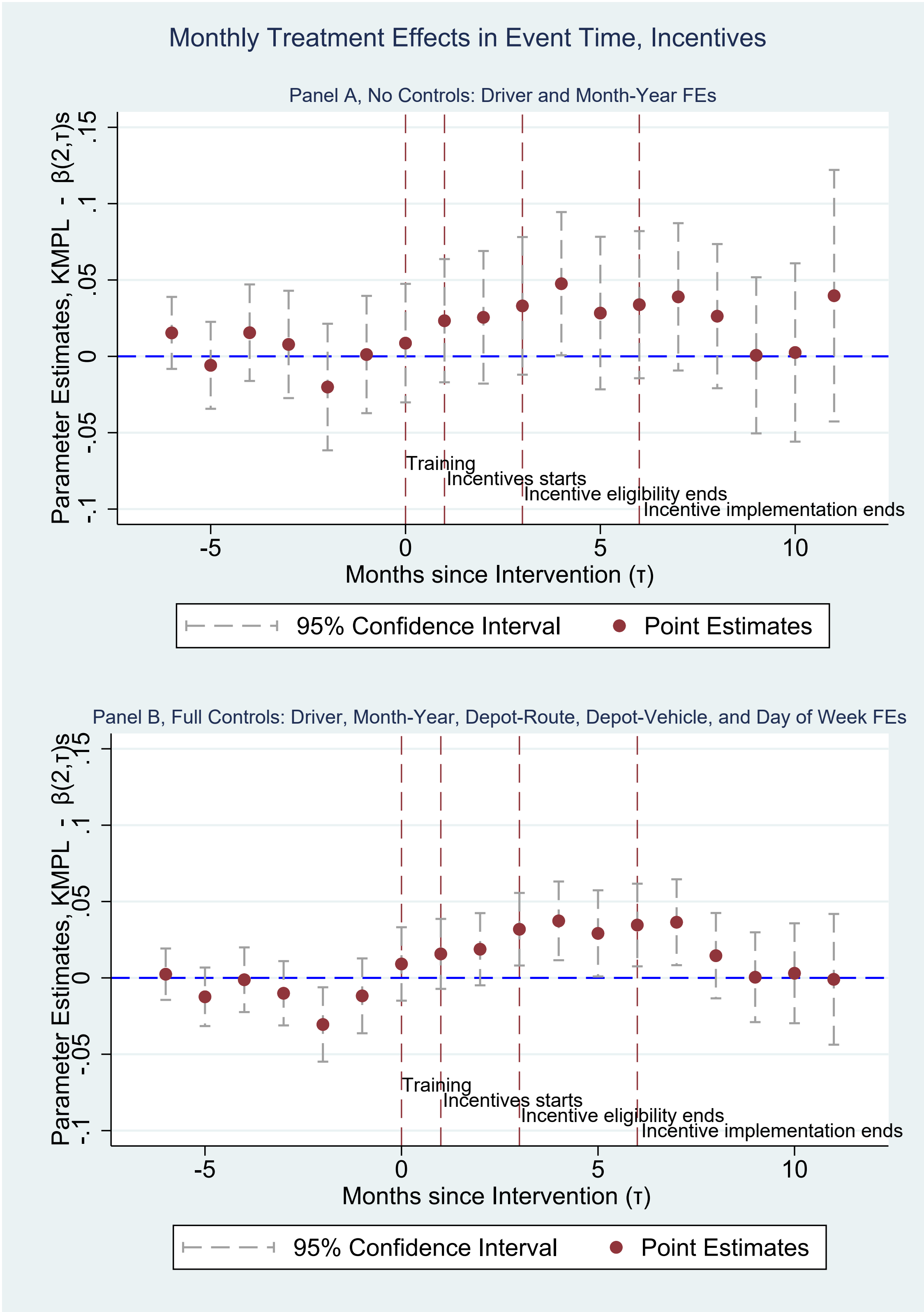

Figure A.8: Monthly Treatment Effects in Event Time, Incentives ($\beta_{2,\tau}s$)

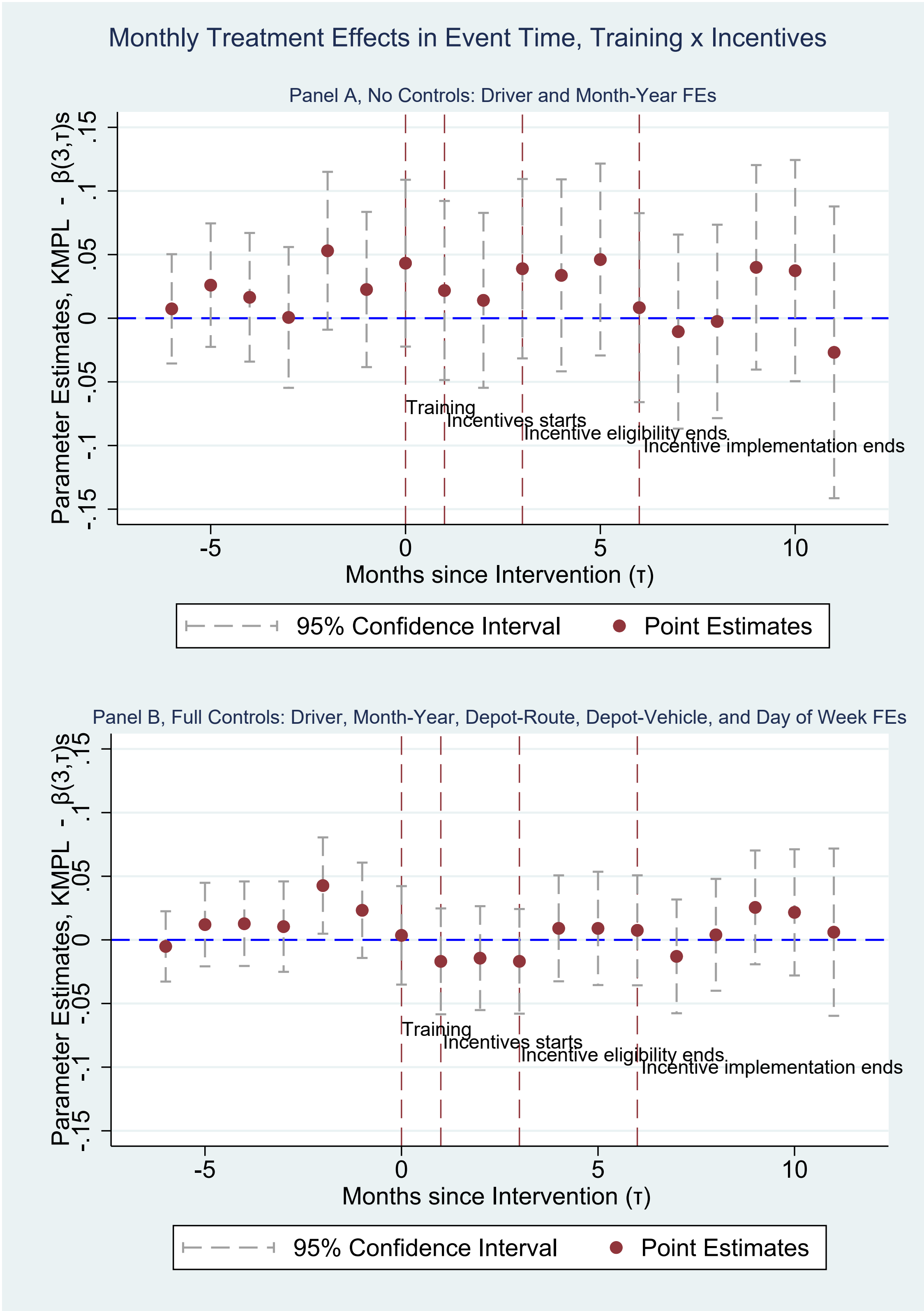

Figure A.9: Monthly Treatment Effects in Event Time, Training $\times$ Incentives ($\beta_{3,\tau}s$)

Appendix B

Appendix to Chapter 2

B.1 Appendix Figures and Tables

Section D: Lottery Game

We would now like to play a lottery game to help us understand how people make choices about risk. The game will take about 5 minutes, and you will win money in the lottery as per the instructions below. We will play the lottery with real money, and the surveyor will give you your winnings right after we play.

1. You will be asked to select one lottery to play out of 5 possible lotteries. You must select one and only one of these lotteries.
2. We will then play the lottery you select with a coin toss, so there are two possible outcomes ("Event 1: Coin is Heads Up" or "Event 2: Coin is Tails Up").
3. So there is a 50% chance of heads up and a 50% chance of tails up. That is, heads will be up half the time.
4. After you select a lottery, you may play the lottery.
5. The experimenter will toss the coin. You will win money as per the results of the coin toss.
6. Thus, your total earnings will be determined by: 1) which of the 5 lotteries you select; and 2) whether the coin is "Heads" or "Tails".
7. The five possible lotteries are below:

Lottery Number	Amount won if "Heads Up"	Amount won if "Tails Up"
1	Rs. 40	Rs. 40
2	Rs. 60	Rs. 30
3	Rs. 80	Rs. 20
4	Rs. 100	Rs. 10
5	Rs. 120	Rs. 0

8. Do you have any questions?

D1.	**Suppose Mr. Ramesh is playing this lottery. He selects Lottery 1. How much money will Mr. Ramesh win if the coin is "Heads"? How much will he win if the coin is "Tails"?** (***Surveyor instructions: If the respondent gives the wrong answer, ask him to look at the table again, and if necessary repeat the instructions.)***	If the coin is "Heads", Mr. Ramesh will win Rs._______ If the coin is "Tails", Mr. Ramesh will win Rs._______
D2.	**Suppose Mr. Satish is playing this lottery. He selects Lottery 5. How much money will Mr. Satish win if the coin is "Heads"? How much will he win if the coin is "Tails"?** (***Surveyor instructions: If the respondent gives the wrong answer, ask him to look at the table again, and if necessary repeat the instructions.)***	If the coin is "Heads", Mr. Satish will win Rs._______ If the coin is "Tails", Mr. Satish will win Rs._______
D3.	**Please tell us which lottery you would like to play.**	Lottery Number: __________________
D4.	**When the experimenter tossed the coin, what happened?**	1. ☐ Heads was up 2. ☐ Tails was up 3. ☐ Participant refuses to answer
D5.	**How much money did you win in total?**	I won Rs._________

Figure B.1: Instructions for Lottery Selection Task to Measure Risk Aversion

	Section E: Allocation Task	
	Next is a decision-making task that involves real money. In this task, you will allocate Rs.30 between yourself and a charity of your choice. You will decide how much of the Rs.30 you want us to give directly to you and how much you want us to send a charity of your choice from the list given below. You can keep all, donate all, or keep the amount you decide to. We will double what you decide to give to the charity. For example, if you decrease your own payment by Rs.10 and leave Rs.20 for yourself, the charity will receive Rs. 20.	

Charity's Name	Charity's goals:
1. UNICEF (United Nations Children's Fund)	Health and education of children
2. CRY (Child Rights and You)	Child rights
3. Being Human	Health and education for underprivileged
4. Help AGE INDIA	Care for disadvantaged older persons
5. CARE India	Empowerment of women and girls
6. Red Cross	Relief during disasters/emergencies
7. Save the Children	Health and education of children

E1.	**Out of the following options, please select the charity you would like to donate to. (*Surveyor instructions: Read out the options*).**	1. ☐ UNICEF (United Nations Children's Fund) 2. ☐ CRY (Child Rights and You) 3. ☐ Being Human 4. ☐ Help AGE INDIA 5. ☐ CARE India 6. ☐ Red Cross 7. ☐ Save the Children 8. ☐ Don't Know 9. ☐ Participant refuses to answer
E2.	**Please select how much you want to keep of the Rs. 30. The rest will be donated.**	I want to keep Rs. _________

Figure B.2: Instructions for Dictator Game to Measure Pro-socialness

	Section F: Dice Task	
F1.	Unique survey respondent ID:	

In this task, the surveyor will step out. We will ask you to privately throw a dice. Please write down the number outcome on the top of the dice after it lands. For example, the outcome is "4" if the top of the dice looks like:

We will ask you to do this 42 times, and write the outcome each time. You will receive Rs. 0.50 (50 paise) for each point rolled. Thus, you will receive money as per the table below.

Roll:	1	2	3	4	5	6
Rs. Received:	Rs. 0.50	Rs. 1.00	Rs. 1.50	Rs. 2.00	Rs. 2.50	Rs. 3.00

If the dice fell on "4", you would receive Rs. 2.00 for this round.

In total, you can earn between Rs. 21 and Rs. 126. Please write down the numbers that you rolled in the table.

F2.

Roll Number	Outcome (from 1-6)	Roll Number	Outcome (from 1-6)	Roll Number	Outcome (from 1-6)
1		16		31	
2		17		32	
3		18		33	
4		19		34	
5		20		35	
6		21		36	
7		22		37	
8		23		38	
9		24		39	
10		25		40	
11		26		41	
12		27		42	
13		28			
14		29			
15		30			
Total (A)		**Total (B)**		**Total (C)**	

F3. **Total Payment (A+B+C)x0.5= Rs.______________**

Figure B.3: Instructions for Dice Task to Measure Dishonesty

Table B.1: Descriptive Statistics by Time Preference Categories

	Section I: Means					Section II: Differences					
	(1) CoPa	(2) CoIm	(3) Hype	(4) PNIL		(1) 1 v. 2	(2) 1 v. 3	(3) 1 v. 4	(4) 2 v. 3	(5) 2 v. 4	(6) 3 v. 4
Behavioral Traits:					.						
Risk aversion (score)	3.082	3.059	2.901	3.263	.	0.0239	0.182^{*}	-0.181	0.158	-0.205	-0.363^{**}
	(1.484)	(1.463)	(1.458)	(1.376)		(0.119)	(0.0952)	(0.147)	(0.133)	(0.170)	(0.156)
Amount donated >= median	0.630	0.463	0.605	0.596	.	0.167^{***}	0.0243	0.0332	-0.143^{***}	-0.134^{**}	0.00893
	(0.483)	(0.500)	(0.489)	(0.493)		(0.0391)	(0.0313)	(0.0482)	(0.0450)	(0.0590)	(0.0532)
Dice points > 95p	0.178	0.176	0.182	0.133	.	0.00263	-0.00366	0.0454	-0.00629	0.0428	0.0491
	(0.383)	(0.381)	(0.386)	(0.341)		(0.0308)	(0.0248)	(0.0378)	(0.0351)	(0.0437)	(0.0409)
Demographic Characteristics:					.						
Age (years)	39.13	38.84	38.23	38.55	.	0.289	0.897	0.577	0.608	0.288	-0.321
	(8.522)	(8.144)	(8.355)	(8.660)		(0.680)	(0.546)	(0.850)	(0.756)	(0.990)	(0.915)
Education (years)	10.23	10.37	10.34	10.25	.	-0.146	-0.117	-0.0193	0.0290	0.127	0.0978
	(1.942)	(1.841)	(1.868)	(1.985)		(0.155)	(0.124)	(0.194)	(0.170)	(0.225)	(0.206)
Career Characteristics:					.						
KSRTC tenure (years)	10.20	10.37	9.331	9.544	.	-0.166	0.870^{*}	0.657	1.036	0.823	-0.213
	(8.137)	(8.317)	(8.117)	(7.708)		(0.657)	(0.524)	(0.805)	(0.748)	(0.961)	(0.870)
Job satisfaction (score)	4.593	4.495	4.488	4.351	.	0.0985	0.105^{**}	0.242^{***}	0.00673	0.144	0.137
	(0.756)	(0.989)	(0.874)	(1.022)		(0.0645)	(0.0509)	(0.0787)	(0.0837)	(0.119)	(0.0992)
Attended driving school	0.0297	0.0585	0.0602	0.0526	.	-0.0288^{*}	-0.0305^{**}	-0.0229	-0.00173	0.00588	0.00761
	(0.170)	(0.235)	(0.238)	(0.224)		(0.0147)	(0.0123)	(0.0176)	(0.0217)	(0.0274)	(0.0255)
Attended KSRTC training	0.859	0.883	0.843	0.825	.	-0.0236	0.0161	0.0349	0.0396	0.0584	0.0188
	(0.348)	(0.322)	(0.364)	(0.382)		(0.0276)	(0.0227)	(0.0350)	(0.0319)	(0.0411)	(0.0400)
Has had accident	0.585	0.601	0.530	0.623	.	-0.0159	0.0550^{*}	-0.0377	0.0709	-0.0217	-0.0927^{*}
	(0.493)	(0.491)	(0.500)	(0.487)		(0.0396)	(0.0319)	(0.0490)	(0.0453)	(0.0581)	(0.0539)
Number of drivers	875	188	332	114	.	1063	1207	989	520	302	446

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Section I reports means, with standard deviations in parentheses, for groups CoPa(Consistently Patient), CoIm(Consistently Impatient), Hype(Hyperbolic), and PNIL(Patient Now Impatient Later)(see Table **??** for details.). Section II reports differences from pair-wise mean comparison tests (ttests), with standard errors in parentheses. See Table 2.1 notes for variable descriptions.

Table B.2: Effect of Randomized Interventions on Attendance

	Overall Treatment Effects		Heterogeneous Treatment Effects			
			Dishonest	Pro-Social	Patient	Risk-Averse
	(1)	(2)	(3)	(4)	(5)	(6)
	Pre	Post	Post	Post	Post	Post
Training	-1.043	-6.298	-6.074	-9.513	-6.723	-12.28
	(4.757)	(4.751)	(5.238)	(7.805)	(7.419)	(11.58)
Incentives	-3.169	-3.405	-1.877	-3.751	-5.412	-11.17
	(4.866)	(4.775)	(5.181)	(7.875)	(7.807)	(12.01)
Training × Incentives	1.922	5.803	5.247	9.873	1.424	18.54
	(6.631)	(6.605)	(7.328)	(10.84)	(10.42)	(16.33)
Char			3.852	-3.747	-6.379	-0.905
			(9.806)	(7.436)	(7.389)	(2.585)
Training × Char			0.0252	5.394	0.173	2.011
			(13.10)	(10.18)	(9.914)	(3.459)
Incentives × Char			-8.479	0.627	2.898	2.570
			(13.94)	(10.29)	(10.06)	(3.479)
Training × Incentives × Char			2.532	-6.829	9.418	-4.227
			(18.26)	(14.21)	(13.80)	(4.758)
Number of Drivers	1508	1508	1497	1507	1496	1506
.						
Dependent Variable Mean	143.2	111.6	111.6	111.6	111.6	111.6

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Parentheses contain robust standard errors. All specifications include strata fixed effects. The outcome variable in Column 1 is attendance in the pre-intervention period between January 2015 and February 2016. The outcome variable in the remaining columns is attendance in the post-intervention period between March 2016 and February 2017. See Table 2.1 notes for detailed descriptions of 'Dishonest', 'Pro-Social', 'Patient', 'Risk-Averse'. Char is the driver characteristic specified in the column title ('Dishonest', etc.).

B.2 Measuring Dishonesty on the Dice Task

In this appendix, I contribute to the literature on using dice games to measure dishonesty at the individual level. I report individual-level dishonesty measures (Hanna and Wang 2017), group-level dishonesty measures (Fischbacher and Föllmi-Heusi 2013), and estimated Type I and Type II measurement errors for several possible dishonesty metrics (Kröll and Rustagi 2016). I find large measurement errors in this setting for all the commonly used metrics. I conclude that the common sample median metric (Hanna and Wang 2017; Banerjee et al. 2016) is a good metric in a dishonest population, but has excessive Type I error in a relatively honest population. The metrics of the 90th, 95th, or 99th percentiles of the theoretical minimize Type I error, and can detect 'strong cheating', but are low-powered to detect 'any type of cheating, partial or strong'. Since most studies indicate the widespread presence of partial cheaters (see Rosenbaum, Billinger, and Stieglitz (2014) for a review), and the theoretical percentile based metrics pool together honest and weakly dishonest respondents and compare them to the strongly dishonest, the downward bias noted by Kröll and Rustagi (2016) is likely in many settings.

To measure dishonesty, I use the dice task introduced by Fischbacher and Föllmi-Heusi (2013), as adapted by Hanna and Wang (2017). Fischbacher and Föllmi-Heusi (2013) introduce a task where respondents privately roll a die once and report the outcomes. This measure can capture propensity for dishonesty at the group level, as the distribution of responses can be compared to the theoretical distribution. Hanna and Wang (2017) adapt this task to obtain an individual level measure of dishonesty, by requesting respondents to roll the dice 42 times, and comparing the individual's distribution of responses to the theoretical distribution. They calculate that 37 dice rolls are needed to identify an individual's propensity for dishonesty using the effect sizes observed in Fischbacher and Föllmi-Heusi (2013), an alpha equal to 0.05, and a power level of 0.8. They round up to 42 dice rolls. They then estimate the respondent's likelihood of cheating through a number of metrics: the raw total dice points, whether the respondent reports above-median dice points, and whether the respondent reports total dice points above the 95th or 99th percentiles of the theoretical distribution.

Kröll and Rustagi (2016) note that statistical distribution based measures can have substantial measurement error, and poorly capture partial cheating. They build on Fischbacher and Föllmi-

Heusi (2013) and Hanna and Wang (2017) and use an innovation of the dice task among 72 informal milkmen in Delhi. Milkmen are asked to roll the die 40 times. They use a Bluetooth enabled dice which transmits the actual outcome of each roll. This enables precise measures of dishonesty as they can compare actual and reported outcomes for each roll of the dice. They demonstrate heterogeneity in who cheats, as measured by the number of over-reported rolls, with half their sample - 36 of 72 milkmen - reporting dishonestly at least once, and also show there is a large variation in the degree of dishonesty, as the number of over-reported rolls ranges from 0 to 27. They compare this Bluetooth enabled precise measure of dishonesty to the binary measures of dishonesty based on statistical distributions, as used by Hanna and Wang (2017). They find significant measurement error using the binary measures of how many drivers reported dice total above the 95th or 99th percentile of the theoretical distribution. Additionally, the binary measure does not capture the heterogeneity in magnitude of dishonesty.

I first report the individual-level dishonesty measures used by Hanna and Wang (2017), based on statistical distributions. In the dishonesty task, across 42 dice rolls, the theoretical mean is 147 points. Among the 1511 dice task respondents, the mean is 153 points, 9% of the respondents have scores above the theoretical 99th percentile, and 18% have scores above the theoretical 95th percentile. Figure 2.1 shows the distribution of dice points in the sample and compares it to the theoretical distribution.

I next report group-level measures of dishonesty, and present estimated frequency distributions for honest and dishonest respondents in Appendix Table B.3. I estimate that 67.5% of the respondents are unconditionally honest and do not cheat at all, 18.5% are partially dishonest, and 14% are strongly dishonest, i.e. they report dice points above the 95th percentile. Using the 99th percentile, 8% are strongly dishonest and 24.5% are partially dishonest[1]. I exploit my large sample of 1522 participants, of whom 1511 did the dice task. Following Fischbacher and Föllmi-Heusi (2013), I assume that individuals reporting the lowest possible outcomes are all

1. Fischbacher and Föllmi-Heusi (2013) categorize people into honest, partial liars, and 'income maximizers' who lie to the fullest extent possible. No one in my sample lies to the fullest possible extent, so even the 'strongly dishonest' people do not report 42 rolls of a 6, i.e. there are no 'income maximizers'. Kröll and Rustagi (2016) classify weakly dishonest people as those who exhibit any amount of cheating and strongly dishonest ones as those who score above the 95th and 99th percentile. My definition is similar, but unlike Kröll and Rustagi (2016) I cannot identify weakly dishonest people at the individual level.

honest, and use this to estimate how many are honest overall. In a large sample of 1511 drivers, I assume the distribution of outcomes among the unconditionally honest drivers is similar to the theoretical distribution. I find that 51 drivers, 103 drivers, and 243 drivers respectively report outcomes below the 5th, 10th, and 25th percentiles of the theoretical distribution. This corresponds to estimates of 1020, 1030, or 972 honest drivers overall[2]. I use the median estimate, that 1020 drivers out of 1511 are unconditionally honest, 67.5%. In Table B.3, I present the actual frequency distribution of all drivers, estimate the frequency distribution of honest drivers, and subtract the latter to estimate the frequency distribution of dishonest drivers. Partial cheating is noticeable in this sample of bus drivers, and no individual maximizes the reported points. Partial dishonesty is typical in the literatre - in a review of experimental evidence on honesty from economics and psychology, Rosenbaum, Billinger, and Stieglitz (2014) find robust evidence of unconditional cheaters and non-cheaters as well as situation dependent cheaters across settings.

Finally, I report the estimated measurement errors of different possible dishonesty metrics, as done by Kröll and Rustagi (2016). I find large measurement errors in all considered metrics. In the 42 rolls dice task, respondents are classified as dishonest if they fall above a particular cutoff, usually the sample median, 95th percentile, or 99th percentile. I consider four possible cutoffs - the sample median, and the 90th, 95th, and 99th percentiles of the theoretical distribution. The results are presented in Panel B of Table B.3. The 90th, 95th, and 99th percentiles have power levels of 0.56, 0.43, and 0.25 respectively. Thus Type II error is substantial - it is 75% at the 99th percentile cutoff, compared to 37.5% in Kröll and Rustagi (2016). The sample median, 153, has a power level of 0.90 and thus a low Type II error of 0.10, but has a Type I error of 0.29.

The sample median metric works well in a strongly dishonest population, but not in a relatively honest one. Many studies use the sample median as a cutoff (Hanna and Wang 2017; Banerjee et al. 2016). This measure has the most power in my population, but a large Type I error. In the original Hanna and Wang (2017) study, however, the sample median corresponds to the 95th percentile of the theoretical distribution, and thus has a small Type I error of 0.05.

2. If unconditionally honest respondents have outcomes similar to the theoretical distribution, then about 5% should report outcomes less than the 5th percentile of the theoretical distribution, and the total number of unconditionally honest respondents should equal $20 \times 51 = 1020$. With the 10th or 25th percentiles, the estimates are 1030 or 972 honest drivers.

On the other hand, the 90th, 95th, and 99th percentile measures are low powered to detect individuals who are partially dishonest. However, these cutoffs have higher power to measure the strongly dishonest. Since partial cheating is common, as Kröll and Rustagi (2016) points out, downward bias is expected. For the main paper, I select the 95th percentile cutoff as the best balance between Type I and Type II error.

Table B.3: Frequency Distribution of Dishonesty on the Dice Task

Panel A: Frequency Distributions of Honest and Dishonest Drivers:

Range		Frequency Distribution			Percentage
Dice Points	Theoretical Percentile	All Drivers - Observed	Honest - Estimated	Dishonest - Estimated	Dishonest - Estimated
42-152	0-67	725	688	37	5.1
153-158	68-84	270	169	101	37.4
159-165	85-95	251	110	141	56.2
166-172	96-99	132	41	91	68.9
173-252	>99	133	12	121	90.98
Total:		1511	1020	491	32.5

Panel B: Measurement Errors with Different Cutoffs:

	Drivers Above Cutoff		Measurement Errors Related to Cutoff		
Cutoff	All (Observed)	Honest/Dishonest (Estimated)	Type I Error (Alpha)	Power (1- Type II Error) (1-Beta)	Type I + Type II Errors
153 (Sample Median)	743	300/443	0.29	0.9	0.39
161 (90th Theoretical Percentile)	382	105/277	0.1	0.56	0.54
165 (95th Theoretical Percentile)	265	53/212	0.05	0.43	0.62
172 (99th Theoretical Percentile)	133	12/121	0.01	0.25	0.76

Appendix C

Appendix to Chapter 3

C.1 Appendix Figures and Tables

Table C.1: Alternate Benefit-Cost Ratios

	Alternative Benefit-Cost Ratios		
Increase in Life Expectancy (years)	1.00 (Chen et al. 2013)	0.35 (Correia et al. 2013)	1.80 (Laden et al. 2006)
per 10 microgram/m3 Decrease in PM2.5			
Total Monetized Life Expectancy Benefits in Surat District	Rs. 590,645 million	Rs. 209,584 million	Rs. 1,066,972 million
Concentration Standard:			
Costs under 3.6% Scenario	Rs. 10,452 million	Rs. 10,452 million	Rs. 10,452 million
Costs under 4.8% Scenario	Rs. 7,839 million	Rs. 7,839 million	Rs. 7,839 million
Benefit: Cost Ratio under 3.6% Scenario	57 to 1	20 to 1	102 to 1
Benefit: Cost Ratio under 4.8% Scenario	75 to 1	27 to 1	136 to 1
Emissions Tax:			
Costs under 3.6% Scenario	Rs. 6,383 million	Rs. 6,383 million	Rs. 6,383 million
Costs under 4.8% Scenario	Rs. 4,788 million	Rs. 4,788 million	Rs. 4,788 million
Benefit: Cost Ratio under 3.6% Scenario	93 to 1	33 to 1	167 to 1
Benefit: Cost Ratio under 4.8% Scenario	123 to 1	44 to 1	223 to 1

3.6% Scenario: Surveyed Plants are 3.6% of Surat Industrial Emissions. 4.8% Scenario: Surveyed Plants are 4.8% of Surat Industrial Emissions.

Table C.2: Net Benefits for Different Percentage Reductions of Baseline Emissions

Variable Name	Policy	Net Benefits for Different Percentage Reductions of Baseline Emissions											
Emissions Reduction (%)		66.42	68.08	69.69	71.14	72.84	74.37	75.84	77.49	79.36	80.85	82.84	84.4
Emissions Abated (Tons)		8625	8841	9050	9238	9459	9658	9849	10063	10306	10499	10758	10960
Average Abatement Cost (Rs./Ton)	E.C.S.	1310	1320	1360	1380	1420	1450	1490	1530	1570	1670	1750	1800
	E.T.	800	830	860	900	940	970	1010	1060	1120	1170	1240	1310
Total Abatement Cost for Surveyed Plants (Rs. Million)	E.C.S.	11.299	11.67	12.308	12.749	13.432	14.004	14.674	15.396	16.18	17.534	18.826	19.728
	E.T.	6.9	7.338	7.783	8.314	8.891	9.368	9.947	10.667	11.542	12.284	13.339	14.358
Aggregate Annual Abatement Costs for all Surat Plants (Rs. Million)	E.C.S.	314	324	342	354	373	389	408	428	449	487	523	548
	E.T.	192	204	216	231	247	260	276	296	321	341	371	399
Net Present Value of Abatement Costs for all Surat Plants (Rs. Million)	E.C.S.	10462	10806	11396	11804	12437	12966	13587	14256	14981	16235	17431	18267
	E.T.	6389	6794	7206	7699	8233	8674	9210	9877	10687	11374	12351	13294
Total Monetized Benefits (Rs. Million)		594404	609259	623668	636644	651858	665550	678705	693471	710206	723540	741349	755310
Net Benefit (Rs. Million)	E.C.S.	583942	598453	612272	624840	639421	652584	665118	679215	695225	707305	723918	737043
	E.T.	588015	602465	616462	628945	643625	656876	669495	683594	699519	712166	728998	742016

'E.C.S'=Emissions Concentration Standard. 'E.T.'=Emissions Tax. Row 2: Emissions Reduction % (Row 1) $\times$ 12986 Tons (Baseline Emissions). Row 5 = (Row 2 $\times$ Row 3)/(10^6), Row 6 = (Row 2 *times* Row 4)/(10^6). Row 7 = Row 5/0.036 (Surveyed Plants=0.036 of all Plants), Row 8 =Row 6/0.036. Row 9 = Row 7/0.03 (Net Present Value using 0.03 discount rate), Row 10 = Row 8/0.03. Row 11 = Row 1 $\times$ Rs. 894,917 million rupees (Monetized life expectancy gains for 100% reduction of emissions, based on Chen et al. (2013) and Jamison et al. (2013).). Row 12 = Row 11 - Row 9, Row 13 = Row 11 - Row 10.

C.2 Estimating Empirical Inputs for the Engineering Model

In this appendix, we describe the empirical inputs for the engineering model. The empirical inputs are based on the industrial plant survey, and also on expert inputs on abatement methods for industrial plants. We sought advice from consultants in the energy and environment department of the Federation of Indian Chambers of Commerce and Industry (FICCI). FICCI assisted with the survey questionnaire and with a method to analyze abatement options based on the survey data. The survey questionnaire was designed specifically to enable this analysis of each industrial plant's abatement options.

C.2.1 Data Sources

Inputs on Air Pollution Control Devices (APCDs) from FICCI

FICCI experts were consulted to prepare detailed energy and air pollution control audits for eight sample industrial plants, and then help in providing a detailed methodology to enable automating the audit analysis for all surveyed plants. FICCI provided inputs on capital, operations, and maintenance costs for Air Pollution Control Device (APCDs), their best collection efficiencies, and typical lifetimes. A summary of these inputs is provided in Tables C.3 and C.4.

Table C.3: Capital Costs of APCDs by Boiler Capacity

Boiler Capacity (Tons Per Hour)	**Capital Costs** (Hundred Thousand Rupees)			
	Bag Filter	**Cyclone**	**Scrubber**	**Boiler**
2	7.1	1.2	2.8	35
5	10.3	1.7	3.3	82
10	14.7	2.1	4.3	136

After consulting FICCI, we estimate operations and maintenance costs as 3% and 6% respectively of the capital costs of APCDs.

Table C.4: APCD Lifetimes and Collection Efficiencies

APCD Type	Lifetime (years)	Efficiency
Cyclone	13	80%
Scrubber	10	94%
Bag Filter	20	99%

As Electrostatic Precipitators (ESPs) are very expensive, and are rare in our survey of industrial plants, we do not consider upgrades to existing ESPs or installation of new ESPs in this model.

Inputs from the Industrial Plant Survey

The survey of industrial plants provides a rich dataset to enable a systematic appraisal of each plant's current PM emissions, reduction potential, and costs of abatement measures. The survey questionnaire included questions on the fuel consumption and combustion sources, the parallel chain configurations attached to each stack, the technical specifications of each APCD, and most importantly, included gravimetric stack sampling to measure stack emissions concentrations and stack flow rates.

C.2.2 Estimating Plant Specific Baseline APCD Efficiencies

We empirically estimate the baseline collection efficiencies of each APCD in each plant using the survey stack sampling results and the FICCI-provided theoretical maximum efficiencies for each APCD type.

For any particular parallel chain (i.e. a combustion source followed by a system of APCDs), the overall chain APCD efficiency relates to the n individual APCD efficiencies as follows:

$$Eff_{chain} = 1 - [(1 - Eff_{APCD1}) \times (1 - Eff_{APCD2}) \times ... \times (1 - Eff_{APCDn})]$$

Where Eff_{APCD1} is the fraction of particulates captured by APCD 1, etc. Thus, $(1 - Eff_{APCD1})$ is the fraction of particulates not captured by APCD 1. $[(1 - Eff_{APCD1}) \times ... \times (1 - Eff_{APCDn})]$

is the fraction of total particulates that is not captured by any of the APCDs, and Eff_{chain} is thus the fraction of particulates captured by the overall APCD system.

We note that it is difficult to measure APCD efficiencies using our available dataset. We have made empirical estimates by combining certain assumptions with survey data on the parallel chain configurations of each stack and the PM concentrations at each stack, as follows.

Estimating Baseline Overall Chain Efficiency:

We group all the industrial plants in the survey by the parallel chain configurations attached to the stack. Within each group of plants with a common configuration, the following process is followed to determine efficiencies of each APCD chain in each plant. 'Inlet concentration' refers to the PM concentration that enters the APCD chain from the combustion source.'Outlet concentration' refers to the PM concentration released from the APCD system through the stack outlet.

- For each group of plants with the same parallel chain configuration, the plant with the lowest outlet PM concentration, the "best performer", is assumed to have a chain efficiency equal to the theoretical maximum efficiency.
- The chain efficiencies of all other plants with the same parallel chain configuration are estimated relative to the best performer, as follows
  - The inlet concentration of the best performer is calculated based on the outlet concentration and the theoretical maximum efficiency.
  - For any given parallel chain, inlet concentration, outlet concentration, and APCD chain efficiency relate as
    $OutletConcentration = InletConcentration \times (1 - Eff_{Chain})$
  - The inlet concentration of all the other plants are assumed to be the same as that of the best performer.
  - Using this estimate of inlet concentration, and the measured outlet concentration, the chain efficiencies of all the plants are estimated.

- If plants have multiple parallel chains, each chain is assigned the same inlet concentration and chain efficiency, but flow rates (i.e. the volume of air associated with that parallel chain) are allowed to differ based on the combustion source capacity within that chain. We then estimate chain efficiency as per the method above.

- For plants that are not part of any group of parallel chain configurations, we assume an inlet concentration of 2000 mg/Nm3. The average inlet concentration, as determined by the inlet sampling done in the survey for a subset of plants, was found to be 1750 mg/Nm3, justifying this choice.

Estimating Baseline Individual APCD Efficiencies:

Although we know the maximum efficiency possible with each APCD type, it is difficult to estimate the efficiency of the individual APCDs in every plant. As a result, we need to make a few assumptions in apportioning the chain efficiency among its constituent APCDs. As a general assumption, we assume that each constituent APCD in a given chain has the same efficiency, bound above by the theoretical maximum for its type. As the efficiencies are multiplicative, the common APCD efficiency would equal $1 - (1 - Effchain)^{1/n}$, where n is the number of APCDs in the parallel chain.

If the efficiency determined above exceeds the maximum theoretical efficiency of the APCD type, the efficiency for that APCD is set at the maximum and the efficiencies of the other APCDs are adjusted accordingly.

C.2.3 Plant-wise Abatement Costs for Each APCD

We combine the FICCI cost inputs with the estimated APCD efficiencies to estimate the capital, operations, and maintenance costs associated with retrofitting particular APCDs or installing new ones.

Costs to install new equipment: The costs of installing new APCDs to complement existing pollution control measures are estimated based on combustion source capacity in the chain. Using the FICCI cost inputs in Table C.3, we interpolate to estimate capital costs of new equipment for each plant. The capital costs are then annualized over the lifetime of the equipment. The annual operations and maintenance costs of the equipment are taken to be 3% and 6% of the upfront capital costs.

C.2.4 Summary of All Empirical Inputs to the Engineering Model

Table C.5: Empirical Inputs to the Engineering Model

Panel A: Direct Inputs	**Source:**
Plant working hours	Industrial plant survey
Stack outlet Suspended Particulate Matter concentration	Industrial plant survey
Stack outlet flow rate (volume of air from the stack)	Industrial plant survey
Parallel chain configuration	Industrial plant survey
Combustion source capacity for each parallel chain	Industrial plant survey
Capital costs of new APCDs by combustion source capacity	FICCI
Operation and maintenance costs as a percentage of capital costs	FICCI
APCD equipment lifetimes	FICCI
APCD maximum collection efficiencies	FICCI

Panel B: Estimated Inputs	**Estimation Method:**
Annual SPM emitted from the plant at baseline	Concentration*Flow Rate *Working Hours
Capital costs of installing or retrofitting APCDs	Described above
Operations costs of installing or retrofitting APCDs	Described above
Maintenance costs of installing or retrofitting APCDs	Described above
Baseline efficiency of APCDs	Described above
Chain specific flow rate	Stack flow rate * (Chain combustion source capacity/ Stack combustionsource capacity)
Chain specific inlet concentration	Described above

Costs to retrofit existing equipment: Capital costs of retrofitting equipment are computed in a similar manner as for the new equipment. Increase in operations and maintenance cost are estimated as 3% and 6% respectively of the retrofit capital costs. However, instead of the full capital costs, the retrofit capital cost is estimated as a function of the relative increase in efficiency from baseline to ideal. Thus,

$$CapitalCostRetrofit_{ik} = \frac{MaxEff_i - BaselineEff_{ik}}{MaxEff_i} \times CapitalCostInstall_{ik}$$

Where for each APCD i in stack k, $CapitalCostRetrofit$ is the capital cost of retrofitting the equipment, $MaxEff_i$ is the theoretical maximum collection efficiency, $BaselineEff_{ik}$ is the APCD efficiency estimated based on the survey data, and $CapitalCostInstall$ is the capital cost of installing new equipment for that particular APCD i in that particular stack k.

C.2.5 Estimating Technology-Specific Costs and Abatement

As described in the body of the paper, a key feature of the model is technology selection as follows:

Model Feature 1 - Selecting a Technology:

For each stack k the firm must choose technology t out of T_k choices ($t = 1, 2, 3, \ldots, T_k, t = 0$ at baseline).

The firm can invest in new air pollution control devices (APCDs) or can retrofit and upgrade existing APCDs. Each distinct and technologically feasible combination of APCD equipment and quality is treated as a unique technology choice t. Firms can select abatement technologies that are a combination of new APCDs and retrofits. Thus, if the firm choses to install a new bag filter, that is one potential technology. If the firm chooses to upgrade the existing cyclone and add a scrubber, that combination of abatement actions is a separate technology. Each technology t, i.e. each discrete combination of abatement actions, yields a specific PM outlet concentration.

The T_k choices for stack k depend on the existing stack configuration, provided by the baseline data. The model organizes this by letting every industry have combinations of APCDs for up to two chains attached to each stack. Each parallel chain can have upto two cyclones, one bag filter, one scrubber, and one electrostatic precipitator. Given baseline data on the existing parallel chain configuration and their efficiencies, the model calculates every combination of installing new APCDs or retrofitting existing APCDs, to form the set of potential technologies T_k.

Empirically Estimating Costs and Abatement for Each Technology

The number of technologies T_k available for each plant stack k depends on the parallel chain configuration of the stack. As explained above, we do not incorporate upgrades to electrostatic precipitators. If a stack already has an APCD installed, we assume that APCD can only be retrofitted, and a new one cannot be purchased instead.

Thus, each parallel chain can have upto four APCD upgrades for the first cyclone, the second cyclone, the bag filter, and the scrubber, where an 'upgrade' can be a new installation or a retrofit. A stack with a single parallel chain can make up to 4 APCD upgrades, and a stack with two parallel chains can make up to 8 APCD upgrades.

Every potential combination of APCD upgrades is an abatement technology. Thus single chain stacks have $T_k = 2^4 = 16$ possible technologies, and two chain stacks have $T_k = 2^8 = 256$ possible technologies. Each technology is a combination of individual APCD upgrades.

We use our empirical estimates of individual APCD upgrade costs and benefits, for every possible combination of APCD upgrades, to get costs and benefits for every possible technology.

C.3 Solutions to the Constrained Optimization Problems

C.3.1 Setting up the Solutions

The stack-level constrained optimization problems involve the marginal abatement cost function and the stack emissions functions.

The marginal abatement cost function is

$$\begin{aligned} C_k(t, Fr_k) &= CC_k(t) + MC_k(t, Fr_k) + OC_k(t, Fr_k) \\ &= CC_k(t) + 1(Fr_k = 0) * 0 + 1(Fr_k > 0) * APCDAnnualMaintenanceCost_k(t) \\ &+ Workinghours_k * Fr_k * APCDHourlyOperatingCosts_k(t) \end{aligned}$$

The stack outlet concentration function is $EmissionsConcentration_k(t)$

The stack annual emissions function is

$$
\begin{aligned}
&AnnualEmissions_k(t, Fr_k) \\
&\qquad = Workinghours_k * [Fr_k * HourlyEmissionsControlled_k(t) \\
&\qquad + (1 - Fr_k) * HourlyEmissionsGenerated_k] \\
&\qquad = Workinghours_k * Fr_k * HourlyEmissionsControlled_k(t) \\
&\qquad + Workinghours_k * (1 - Fr_k) * HourlyEmissionsGenerated_k \\
&where \\
&\qquad HourlyEmissionsControlled_k(t) \\
&\qquad = HourlyEmissionsGenerated_k Chain_1 \\
&\qquad * (1 - APCDefficiency_k Chain1(t)) \\
&\qquad + HourlyEmissionsGenerated_k Chain_2 \\
&\qquad * (1 - APCDefficiency_k Chain2(t))
\end{aligned}
$$

As technology t is a discrete variable, we empirically evaluate costs and emissions for every potential value of t.

Conditional on t being constant, i.e. for a specific technology, the marginal abatement cost function and stack emissions functions contain several functions of t that simplify to constants. Let these be represented by specific letters as follows:

Table C.6: Functions that are Constant Conditional on Technology

Functions that are Constant Conditional on t	Letter
$CC_k(t\|t)$	A
$APCDAnnualMaintenanceCost_k(t\|t)$	B
$APCDHourlyOperatingCosts_k(t\|t)$	C
$EmissionsConcentration_k(t\|t)$	D
$HourlyEmissionsControlled_k(t\|t)$	E

Additionally, certain variables are constant for a particular stack:

Table C.7: Functions that are Constant Conditional on Stack

Constants for a Given Stack k	Letter
$Workinghours_k$	F
$HourlyEmissionsGenerated_k$	G

Thus, the marginal abatement cost function simplifies to

$$\begin{aligned} C_k(t, Fr_k|t) &= A + 1(Fr_k = 0) * 0 + 1(Fr_k > 0) * B + F * Fr_k * C \\ &= A + 1(Fr_k > 0) * B + F * Fr_k * C \end{aligned}$$

The stack emissions functions simplify to

$$\begin{aligned} EmissionsConcentration_k(t|t) &= D \\ AnnualEmissions_k(t, Fr_k|t) &= F * Fr_k * E + F * (1 - Fr_k) * G \\ &= Fr_k * F * E + F * G - Fr_k * F * G \\ &= Fr_k * (F * E - F * G) + F * G \end{aligned}$$

C.3.2 Emissions Concentration Standard

Solving the emissions concentration standard problem:

$$Minimize_{t,Fr_k} C_k(t, Fr_k)$$

$$subject\ to$$

1. $EmissionConcentration_k(t) \leq ConcentrationStandard$
2. $Fr_k = 1$

replacing the constraint $Fr_k = 1$ into the cost function, this simplifies to

$$\begin{aligned} Minimize_{t,Fr_k|Fr_k=1} CC_k(t) &+ APCDAnnualMaintenanceCost_k(t) \\ &+ Workinghours_k * APCDHourlyOperatingCosts_k(t) \end{aligned}$$

$$subject\ to$$

1. $EmissionsConcentration_k(t) \leq ConcentrationStandard$

As t is an discrete variable, we solve this empirically by evaluating every potential value of t.

Thus, the solution to the emissions concentration standard optimization problem is $Fr_k^* = 1$, $t = t^*$, which is the empirical solution to the above cost minimization problem.

C.3.3 Emissions Load Standard

Solving the emissions load standard problem:

$$\begin{aligned}
&Minimize_{t,Fr_k} C_k(t, Fr_k)\\
&\quad subject\ to\\
&\qquad\qquad 1.\ AnnualEmissions_k(t, Fr_k) \leq EmissionsPermits_k
\end{aligned}$$

As this function is discontinuous at $Fr_k = 0$, we analyze two cases.

Case 1: $Fr_k = 0$ This simplifies to

$$\begin{aligned}
&Minimize_{t,Fr_k|Fr_k=0} CC_k(t)\\
&\quad subject\ to\\
&\qquad\qquad 1.\ Workinghours_k * HourlyEmissionsGenerated_k\\
&\qquad\qquad \leq EmissionsPermits_k
\end{aligned}$$

And we solve for the optimal discrete value of t as the empirical solution to the above cost minimization function.

Case 2: $Fr_k > 0$ Conditional on t, and thus replacing various functions of t with constants, and including the boundary conditions for Fr_k, this simplifies to

$$\begin{aligned}
&Minimize_{Fr_k} A + B + Fr_k * F * C\\
&\quad subject\ to\\
&\qquad\qquad 1.\ Fr_k * (F * E - F * G) + F * G \leq EmissionsPermits_k\\
&\qquad\qquad 2.\ Fr_k \epsilon (0, 1]
\end{aligned}$$

This problem can be solved using Lagrangian multipliers and Karush-Kuhn-Tucker conditions for a constrained optimization problem with inequality constraints. The constraints can be rewritten as three inequality constraints.

1. $EmissionsPermits_k - Fr_k * (F * E - F * G) - F * G \geq 0$

2. $Fr_k > 0$

3. $1 - Fr_k \geq 0$

Let the Lagrangian be written in a standard minimization form:

$$
\begin{aligned}
Minimize_{Fr_k} L(Fr_k, \lambda) = & \; A + B + Fr_k * F * C \\
& - \lambda_1 (EmissionsPermits_k - Fr_k * (F * E - F * G) - F * G) \\
& - \lambda_2 (Fr_k) - \lambda_3 (1 - Fr_k)
\end{aligned}
$$

The Karush-Kuhn-Tucker conditions are as follows:

$$
\begin{aligned}
1. \; \frac{\partial L(Fr_k, \lambda_1, \lambda_2, \lambda_3)}{\partial Fr_k} &= F * C - \lambda_1^*(-)(F * E - F * G) - \lambda_2^* - \lambda_3^*(-1) = 0 \\
&= F * C + \lambda_1^*(F * E - F * G) - \lambda_2^* + \lambda_3^* = 0 \\
& 2a. \; EmissionsPermits_k - Fr_k^* * (F * E - F * G) - F * G \geq 0 \\
& 2b. \; \lambda_1^* \geq 0 \\
& 2c. \; \lambda_1^*(EmissionsPermits_k - Fr_k^* * (F * E - F * G) - F * G) = 0 \\
& 3a. \; Fr_k^* \geq 0 \\
& 3b. \; \lambda_2^* \geq 0 \\
& 3c. \; \lambda_2^*(Fr_k^*) = 0 \\
& 4a. \; 1 - Fr_k^* \geq 0 \\
& 4b. \; \lambda_3^* \geq 0 \\
& 4c. \; \lambda_3^*(1 - Fr_k^*) = 0
\end{aligned}
$$

However, we know that $Fr_k > 0$, and thus 3a., 3b., and 3c. simplify to $\lambda_2^* = 0$.

λ_1^* and λ_3^* can each take on the value $= 0$ or > 0, so there are four subcases to check given $Fr_k > 0$ and $\lambda_2^* = 0$.

Let the four subcases of Case 2: $Fr_k > 0$ be Case 2.I., 2.II., 2.III., and 2.IV.

Case 2.I: $\lambda_1^* = 0$,$\lambda_2^* = 0$, $\lambda_3^* = 0$

Case 2.II: $\lambda_1^* > 0$,$\lambda_2^* = 0$, $\lambda_3^* = 0$

Case 2.III: $\lambda_1^* = 0$,$\lambda_2^* = 0$, $\lambda_3^* > 0$

Case 2.IV: $\lambda_1^* > 0$,$\lambda_2^* = 0$, $\lambda_3^* > 0$

Case 2.I: $\lambda_1^* = 0$,$\lambda_2^* = 0$, $\lambda_3^* = 0$, so the complementary slackness conditions are $EmissionsPermits_k - Fr_k^* * (F * E - F * G) - F * G \geq 0$, $Fr_k^* > 0$, and $1 - Fr_k^* \geq 0$.
The first order condition is $F * C = 0$, i.e.
$Workinghours_k * APCDHourlyOperatingCosts_k(t) = 0$

The inequality constraints are slack in this sub-case. Thus this sub-case can only be feasible if empirically there are any situations where
$Workinghours_k * APCDHourlyOperatingCosts_k(t) = 0$.

Solution: $Fr_k^* > 0$, $\lambda_1^* = 0$, $\lambda_2^* = 0$, $\lambda_3^* = 0$, feasible solution only if
$Workinghours_k * APCDHourlyOperatingCosts_k(t) = 0$.

Case 2.II. $\lambda_1^* > 0$,$\lambda_2^* = 0$, $\lambda_3^* = 0$, so the complementary slackness conditions are $EmissionsPermits_k - Fr_k^* * (F * E - F * G) - F * G = 0$, $Fr_k^* > 0$, and $1 - Fr_k^* \geq 0$, i.e. the load standard constraint is binding. The first order constraint is $F * C + \lambda_1^*(F * E - F * G) = 0$.

Solution: $Fr_k^* * = \frac{EmissionsPermits_k - F*G}{F*E - F*G}$, $\lambda_1^* = \frac{-F*C}{F*E-F*G}$, $\lambda_2^* = 0$, $\lambda_3^* = 0$.

Case 2.III. $\lambda_1^* = 0$,$\lambda_2^* = 0$, $\lambda_3^* > 0$, so the complementary slackness conditions are $EmissionsPermits_k - Fr_k^* * (F * E - F * G) - F * G \geq 0$, $Fr_k^* > 0$, and $1 - Fr_k^* = 0$. The first order condition is $F * C + \lambda_3^* = 0$, and $Fr_k^* = 1$.

Solution: $Fr_k^* * = 1$, $\lambda_1^* = 0$, $\lambda_2^* = 0$, $\lambda_3^* = -F * C$

Case 2.IV. $\lambda_1^* > 0$,$\lambda_2^* = 0$, $\lambda_3^* > 0$ so the complementary slackness conditions are $EmissionsPermits_k - Fr_k^* * (F * E - F * G) - F * G = 0$, $Fr_k^* > 0$, and $1 - Fr_k^* = 0$, i.e. the load standard constraint is binding and $Fr_k^* = 1$.

Solution: $Fr_k^* * = 1$, $\lambda_1^* > 0$,$\lambda_2^* = 0$, $\lambda_3^* > 0$, this is a feasible solution only if
$EmissionsPermits_k = F * E$

Returning to the emissions load standard problem, we have

$$Minimize_{t,Fr_k} C_k(t, Fr_k)$$

$$subject\ to$$

$$1.\ AnnualEmissions_k(t, Fr_k) \leq EmissionsPermits_k$$

Using Karush-Kuhn-Tucker conditions, we find that conditional on t, Fr_k can have up to 5 potential solutions, given the inequality constraints. These five cases are summarized in the tables below:

Table C.8: Karush-Kuhn-Tucker Solutions for the Emissions Load Standard Problem - Part I

Case	Fr_k^*	Constraint	Feasible if:
Case 1	0	Binding	$F * G \leq EmissionsPermits_k$
Case 2.I.	$0 < Fr_k^* \leq 1$	Not binding	$F * C = 0$
Case 2.II.	$Fr_k^* * = \frac{EmissionsPermits_k - F*G}{F*E - F*G}$	Binding	NA
Case 2.III.	$Fr_k^* = 1$	Not binding	NA
Case 2.IV.	$Fr_k^* = 1$	Binding	$EmissionsPermits_k = F * E$

Table C.9: Karush-Kuhn-Tucker Solutions for the Emissions Load Standard Problem - Part II

Case	Fr_k^*	Cost function
Case 1	0	A
Case 2.I.	$F * C = 0$	$A + B$
Case 2.II.	$Fr_k^* * = \frac{EmissionsPermits_k - F*G}{F*E - F*G}$	$A + B + \frac{(EmissionsPermits_k - F*G)*F*C}{F*E - F*G}$
Case 2.III.	$Fr_k^* = 1$	$A + B + F * C$
Case 2.IV.	$Fr_k^* = 1$	$A + B + F * C$

Note that since $B = APCDAnnualMaintenanceCost_k(t|t) \geq 0$, Case 1. weakly dominates Case 2.I.

To solve the emissions load standard problem, we empirically evaluate the costs and whether the constraint is satisfied for each of these 5 cases of Fr_k, for every value of t.

C.3.4 Cap and Trade

Solving the cap and trade problem:

$$Minimize_{t,Fr_k} C_k(t, Fr_k) + Pr * [AnnualEmissions_k(t, Fr_k) - EmissionsPermits_k]$$

For a given technology, conditional on t, this simplifies to:

$$\begin{aligned} Minimize_{Fr_k} A + 1(Fr_k > 0) * B + F * Fr_k * C \\ + Pr * [Fr_k * (F * E - F * G) + F * G - EmissionsPermits_k] \end{aligned}$$

Since this function is discontinuous at $Fr_k = 0$, we look at two cases.

Case 1. $Fr_k^* = 0$, Cost= $A + Pr * [F * G - EmissionsPermits_k]$

In this solution, the firm keeps its own emissions uncontrolled and only buys permits as needed.

Case 2. $Fr_k^* > 0$ and $Fr_k \leq 1$

$$\begin{aligned} Minimize_{Fr_k} A + B + F * Fr_k * C \\ + Pr * [Fr_k * (F * E - F * G) + F * G - EmissionsPermits_k] \\ subject\ to\ \ Fr_k \epsilon (0, 1] \end{aligned}$$

This can be rewritten as a Lagrangian problem with inequality constraints and solved via Karush-Kuhn-Tucker conditions. The two inequality constraints are 1. $Fr_k > 0$ and 2. $1 - Fr_k \geq 0$. The Lagrangian is written in a standard minimization form:

$$\begin{aligned} Minimize_{Fr_k} L(Fr_k, \lambda_1, \lambda_2) = {} & A + B + Fr_k * F * C \\ & + Pr * [Fr_k * (F * E - F * G) + F * G - EmissionsPermits_k] \\ & - \lambda_1(Fr_k) - \lambda_2(1 - Fr_k) \end{aligned}$$

The Karush-Kuhn-Tucker conditions are as follows:

$$1.\ \frac{\partial L(Fr_k, \lambda_1, \lambda_2)}{\partial Fr_k} = F * C + Pr * (F * E - F * G) - \lambda_1^* + \lambda_2^* = 0$$

$$2a.\ Fr_k^* \geq 0$$

$$2b.\ \lambda_1^* \geq 0$$

$$2c.\ \lambda_1^*(Fr_k^*) = 0$$

$$3a.\ 1 - Fr_k^* \geq 0$$

$$3b.\ \lambda_2^* \geq 0$$

$$3c.\ \lambda_2^*(1 - Fr_k^*) = 0$$

We know that $Fr_k > 0$, thus $\lambda_1^* = 0$. There are two remaining sub-cases.

Case 2.I. $\lambda_1^* = 0$, $\lambda_2^* = 0$, so the slack complementarity condition is $1 - Fr_k^* \geq 0$ and the first order constraint is $F * C + Pr * (F * E - F * G) = 0$. This solution is only valid if this equation holds empirically for a specific price Pr.

Solution:$Fr_k^* \epsilon (0, 1]$, valid only if $F * C + Pr * (F * E - F * G) = 0$

Case 2.II. $\lambda_1^* = 0$, $\lambda_2^* > 0$ so the complementarity condition is binding at $1 - Fr_k^* = 0$

Solution: $Fr_k^* = 1$

Thus, using Karush-Kuhn-Tucker conditions, we find that conditional on t, Fr_k can have up to 3 potential solutions, given the inequality constraints. These three cases are summarized in the table below:

Table C.10: Karush-Kuhn-Tucker Solutions for the Cap and Trade Problem - Part I

Case	Fr_k^*	Feasible solution if:
Case 1	0	NA
Case 2.I.	$0 < Fr_k^* \leq 1$	$F * C + Pr * (F * E - F * G) = 0$
Case 2.II.	$Fr_k^* = 1$	NA

Table C.11: Karush-Kuhn-Tucker Solutions for the Cap and Trade Problem - Part II

Case	Fr_k^*	Cost function
Case 1	0	$A + Pr * [F * G - EmissionsPermits_k]$
Case 2.I.	$0 < Fr_k^* \leq 1$	$A + B + Pr * [F * G - EmissionsPermits_k]$
Case 2.II.	$Fr_k^* = 1$	$A + B + F * C + Pr * [F * E - EmissionsPermits_k]$

To solve the emissions trading scheme problem, we empirically evaluate the costs and whether the constraint is satisfied for each of these three cases of Fr_k, for every value of t.

Market Equilibrium for Cap and Trade:

The aggregate emissions cap is $\sum_k^K EmissionsPermits_k$, where K is the number of stacks in the market. When the market is in equilibrium, aggregate supply of permits should equal aggregate demand for permits. Thus the net sum of permits purchased by all stacks should equal zero (given that it is a closed market and some stacks are sellers). Ideally in equilibrium

$$\sum_k^K [AnnualEmissions_k(t, Fr_k) - EmissionsPermits_k] = 0$$

In order to ensure that aggregate emissions are below the cap, it must be the case that the net sum of purchases is negative, i.e. that there are more permits sold than bought.

Due to discontinuities in the cost curves for market participants, the condition that aggregate supply of permits equals aggregate demand is rarely empirically satisfied, and thus the equilibrium permit price is defined as the minimum permit price such that net sum of permit purchases is negative. It is the solution to the following minimization problem.

$$Minimize_{Pr} Pr$$

$$subject\ to$$

$$1.\ \sum_k^K [AnnualEmissions_k(t, Fr_k) - EmissionsPermits_k] <= 0$$

Thus the equilibrium conditions for a cap and trade program are:

1. $Pr = Pr^*$, the solution to the above minimization problem

2. $t_k = t_k^*, Fr_k = Fr_k^* \; \forall k$, i.e. each firm is optimizing given the equilibrium permit price

3. $\sum_k^K AnnualEmissions_k(t, Fr_k) <= \sum_k^K EmissionsPermits_k$ i.e. aggregate emissions are below the aggregate emissions cap.

C.3.5 Emissions Tax

Solving the emissions tax problem: Each stack's constrained optimization problem is

$$Minimize_{t,Fr_k} C_k(t, Fr_k) + Tax * AnnualEmissions_k(t, Fr_k)$$

For a given technology, conditional on t, this simplifies to:

$$Minimize_{Fr_k} A + 1(Fr_k > 0) * B + Fr_k * F * C + Tax * [Fr_k * (F * E - F * G) + F * G]$$

Since this function is discontinuous at $Fr_k = 0$, we consider two cases.

Case 1. $Fr_k^* = 0$, Cost $= A + Tax * F * G$.

In this solution, the firm keeps its emissions uncontrolled and pays the resulting emissions tax.

Case 2. $0 < Fr_k \leq 1$

$$Minimize_{Fr_k} A + B + Fr_k * F * C + Tax * [Fr_k * (F * E - F * G) + F * G]$$
$$subject\ to\ 0 < Fr_k \leq 1$$

This can be rewritten as a Lagrangian with inequality constraints and solved via Karush-Kuhn-Tucker conditions. The two inequality constraints are 1. $Fr_k > 0$ and 2. $1 - Fr_k \geq 0$. The Lagrangian is written in a standard minimization form:

$$\begin{aligned} Minimize_{Fr_k} L(Fr_k, \lambda_1, \lambda_2) A + B + Fr_k * F * C \\ + Tax * [Fr_k * (F * E - F * G) + F * G] \\ - \lambda_1(Fr_k) - \lambda_2(1 - Fr_k) \end{aligned}$$

The Karush-Kuhn-Tucker conditions are as follows:

$$1.\ \frac{\partial L(Fr_k, \lambda_1, \lambda_2)}{\partial Fr_k} = F * C + Tax * (F * E - F * G) - \lambda_1^* + \lambda_2^* = 0$$

$$2a.\ Fr_k^* \geq 0$$

$$2b.\ \lambda_1^* \geq 0$$

$$2c.\ \lambda_1^*(Fr_k^*) = 0$$

$$3a.\ 1 - Fr_k^* \geq 0$$

$$3b.\ \lambda_2^* \geq 0$$

$$3c.\ \lambda_2^*(1 - Fr_k^*) = 0$$

We know that $Fr_k > 0$, thus $\lambda_1^* = 0$. There are two remaining sub-cases.

Case 2.I. $\lambda_1^* = 0$, $\lambda_2^* = 0$, so the slack complementarity condition is $1 - Fr_k^* \geq 0$ and the first order constraint is $F * C + Tax * (F * E - F * G) = 0$. This solution is only valid if this equation holds empirically for a specific tax Tax.

Solution:$Fr_k^* \epsilon (0, 1]$, valid only if $F * C + Tax * (F * E - F * G) = 0$

Case 2.II. $\lambda_1^* = 0$, $\lambda_2^* > 0$ so the complementarity condition is binding at $1 - Fr_k^* = 0$

Solution: $Fr_k^* = 1$

Thus, using Karush-Kuhn-Tucker conditions, we find that conditional on t, Fr_k can have up to 3 potential solutions, given the inequality constraints. These three cases are summarized in the table below. Note that since $B = APCDAnnualMaintenanceCost_k(t|t) \geq 0$, Case 1. weakly dominates Case 2.I.

Table C.12: Karush-Kuhn-Tucker Solutions for the Emissions Tax Problem

Case	Fr_k^*	Feasible solution if:	Cost function
Case 1	0	NA	$A + Tax * F * G$
Case 2.I.	$0 < Fr_k^* \leq 1$	$F * C + Tax * (F * E - F * G) = 0$	$A + B + Tax * F * G$
Case 2.II.	$Fr_k^* = 1$	NA	$A + B + F * C + Tax * F * E$

To solve the emissions tax problem, we empirically evaluate the costs and whether the constraint is satisfied for each of these 3 cases of Fr_k, for every value of t.