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

Janhavi Nilekani, Harvard University *

Job Market Paper

January 6, 2018

Abstract

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 cost-effectiveness of 3.12. 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 cost-effectiveness of 4.22. Along with the high return on investment from fuel savings, the interventions generated positive externalities from reduced vehicle emissions.

*I would like to thank Rohini Pande, Rema Hanna, and Rob Stavins for their invaluable guidance. Participants of the Harvard Development Economics lunch, Environmental Economics lunch, and Development Economics tea provided numerous helpful suggestions, as did fellow Public Policy graduate students and anonymous reviewers for the JPAL Governance Initiative and Foundations for Human Behavior grants. I would also like to thank the Karnataka State Road Transport Corporation for their support and cooperation, J-PAL SA for excellent field research support, 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.

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 (PM2.5) 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, better only 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 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

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

(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 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

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

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 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 find 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 transient, 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 paper is organized as follows. Section 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 3 describes the experimental design, including the study area, recruitment, randomization, timeline, data, and empirical strategy. Section 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 5 estimates the cost-effectiveness of the interventions and Section 6 concludes.

2 Context

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.

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). 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).

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 shows the overall distribution of fuel efficiency in kilometers per liter (KMPL) in the dataset of KSRTC shift-level data.

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 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 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 entirely 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

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.

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.

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.

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]

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 four 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 re-

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

lated to fuel efficiency, which itself might 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

13. see Section 3.2 for details.

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 Hanna and Wang (2014). 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 had 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 (2014).

3 Experimental Design

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 A1 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).

	Control	Financial Incentives
Control	C: Control=381	T2: Incentives Only=381
Training	T1:Training Only=380	T3: Training and Incentives=380

I use the 2015 baseline survey to check the randomization. Table 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.

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 A2 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 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 A4 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.

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

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.

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, 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.

$$
\begin{aligned}
Y_{imd} = {} & \beta_1 Training_d \times Post_i + \beta_2 Incentives_d \times Post_i \\
& + \beta_3 Training_d \times Incentives_d \times Post_i + \alpha_d + \delta_m + Controls_{imd} + \epsilon_{imd}
\end{aligned}
\tag{1}
$$

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 2 I estimate monthly treatment effects.

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

Where I_m is a set of indicator variables for month-year m and the rest is as in Equation 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 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 2.

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

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 $Controls_{imd}$.

4 Results

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.

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 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 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 A1 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 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 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.

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 3 I verify there is balance on baseline fuel efficiency (KMPL) variables in the post-attrition subsample. Appendix Table A2 shows that the post-attrition subsample is balanced on baseline survey characteristics. Appendix Table A3 demonstrates that take-up and compliance was qualitatively similar in the post-attrition subsample and the complete sample.

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 4 presents the results from Equation 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 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.

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.

Monthly treatment effects:

First, I analyze the monthly treatment effects based on Equation 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 2, 3, and 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 A5.

The training program had a positive effect in the four months of implementation and no effect thereafter. Figure 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 A3, 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 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 A4.

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 3 studies treatment effects for 4 different experiment phases. Complete results are presented in Table 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

15. Appendix Figures A3, A4, and A5 reproduce these figures in Panel B, and compare them to the coefficients from the specification with no controls in Panel A.

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 3.2 and graphed in Appendix Figure A2.

During the Phase 1 training intervention implementation, from March-June 2016, the treatment effect point estimates in Table 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 5. The 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, Column 1 without controls and Column 4 with controls, 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 B in which this pattern is very noticeable. In the main approach in Table 5, the peak effect was in July-September 2016, when 69% of the observations are from the post-eligibility implementation phase[16].

4.4 Heterogeneity of the Treatment Effects

I analyze the heterogeneity of treatment effects based on the conceptual framework in Section 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

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.

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.

To investigate heterogeneity, for the qualitative variable of time preference, I conduct a subgroup analysis of Equation 1. For all other variables, I adapt Equation 1 and add an interaction with the driver characteristic as follows

$$\begin{aligned} Y_{imd} = {} & \beta_1 Training_d \times Post_i + \beta_2 Incentives_d \times Post_i \\ & + \beta_3 Training_d \times Incentives_d \times Post_i + \beta_4 Training_d \times Post_i \times Char_d \\ & + \beta_5 Incentives_d \times Post_i \times Char_d + \beta_6 Training_d \times Incentives_d \times Post_i \times Char_d \\ & + \beta_7 Post_i \times Char_d + \alpha_d + \delta_m + Controls_{imd} + \epsilon_{imd} \end{aligned} \tag{4}$$

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 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 6 presents the results.

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 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 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 in detail in a forthcoming dissertation chapter (Nilekani 2018), and only provide a brief summary here. To measure dishonesty, I implemented a dice task that measures propensity to cheat in exchange for financial reward, closely following Hanna and Wang (2014). I measured time preferences through a series of hypothetical time discounting questions, closely following Ashraf, Karlan, and Yin (2006). Table 7 presents the results of the heterogeneity analysis and Table 9 presents the results of the subgroup analysis for drivers with different time preferences.

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.

Third, I test the importance of the salience and habit channels based on the conceptual framework in Section 2.5, and find no heterogeneous effects based on age, risk aversion, or pro-socialness, as seen in Table 8. 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 are intrinsically motivated to drive well and are easily activated by salient cues. I use behavioral measures of risk aversion and pro-socialness, which I describe in detail in a forthcoming dissertation chapter (Nilekani 2018), and briefly summarize here. 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 (2014).

4.5 Comparing the Conceptual Framework Predictions to the Results

In Section 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 4.2, 4.3, and 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 implementation 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 B. 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 no evidence that the interventions were more effective for risk averse or pro-social

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.

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 B, 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.

4.6 Robustness Checks

The results are robust to alternative specifications and methodologies.

Along with the difference in difference specification in Equation 1, I also estimate an analogous model using post-intervention data only, while controlling for drivers' baseline KMPL, and find similar results. I estimate

$$\begin{aligned} Y_{imd} &= \beta_1 Training_d + \beta_2 Incentives_d + \beta_3 Training_d \times Incentives_d \\ &\quad + \beta_4 BaselineKMPL_d + \gamma_s + \delta_m + Controls_{imd} + \epsilon_{imd}, \forall Y_{imd} \in [m >= Mar2016] \end{aligned} \tag{5}$$

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. Appendix Table A7 presents the results. Equations 1 and 5 yield similar results in the specifications without controls (Column 1 of Tables 4 and A7). Results differ significantly in the models with controls, as Equation 1 estimates vehicle, route, and other fixed effects using the full panel dataset, while Equation 5 estimates vehicle and other fixed effects using post-intervention data only.

Similarly, as an analogue to Equation 2, in Equation 6 I separately estimate treatment effects for each month-year m in September 2015-February 2017 and find similar time patterns:

$$\begin{aligned} Y_{imd} = {} & \beta_{1,m} Training_d + \beta_{2,m} Incentives_d + \beta_{3,m} Training_d \times Incentives_d \\ & + \beta_4 BaselineKMPL_d + \gamma_s + \epsilon_{imd} \end{aligned} \tag{6}$$

Where $BaselineKMPL_d$ is the average KMPL for shifts from January-August 2015, and the rest is as in Equation 5. Each $\beta_{1,m}, \beta_{2,m}$, and $\beta_{3,m}$ comes from separate regressions for each month-year m. Figure A6 plots the monthly treatment effects from the difference-in-difference specification without controls (Column 1 of Appendix Table A5), along with the separate estimates for each month from Equation 6. The results from both approaches are very consistent.

Results remain unchanged if I drop shifts with multiple drivers. As described in Section 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 A8 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.

5 Cost-Effectiveness of the Interventions

Both interventions were cost-effective. A complete 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 3.12 for training and 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 2.04 and the cost-effectiveness of incentives would be 1.46.

5.1 Cost-effectiveness analysis

I present approximate estimates of cost-effectiveness for the two interventions, both for the pilot and at scale. Figure 5 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 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.

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 2.04 and the cost-effectiveness of incentives would be 1.46.

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

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 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 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.

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.

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.

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, D. Karlan, and W. 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.

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.

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.

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

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.

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.

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.

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. 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.

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.

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

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.

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.

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. 2014. "Dishonesty and Selection into Public Service: Evidence from India." *NBER Working Paper #19649, forthcoming at American Economic Journal: Economic Policy,* no. October. http://scholar.harvard.edu/files/remahanna/files/4%7B%5C_%7Dselection%7B%5C_%7Dpaper%7B%5C_%7Dtables%7B%5C_%7Dappendix.pdf.

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.

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.

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.

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.

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.

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.

Nilekani, Janhavi. 2018. "Personality Traits and Absenteeism: An Analysis of Public Sector Bus Drivers in India." *Working Paper.*

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.

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.

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.

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.

Figures

Figure 1: Distribution of Kilometers Per Liter

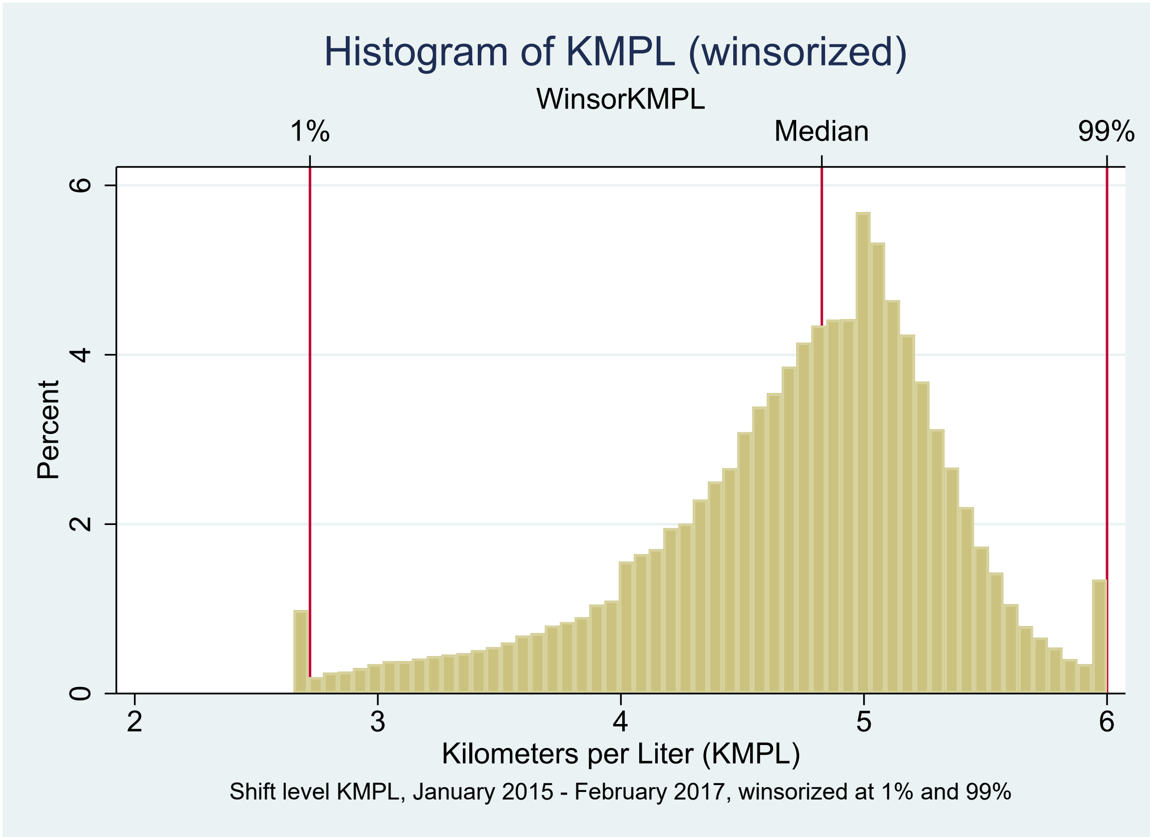


Figure 2: Monthly Treatment Effects, Training ($\beta_{1,m}s$)

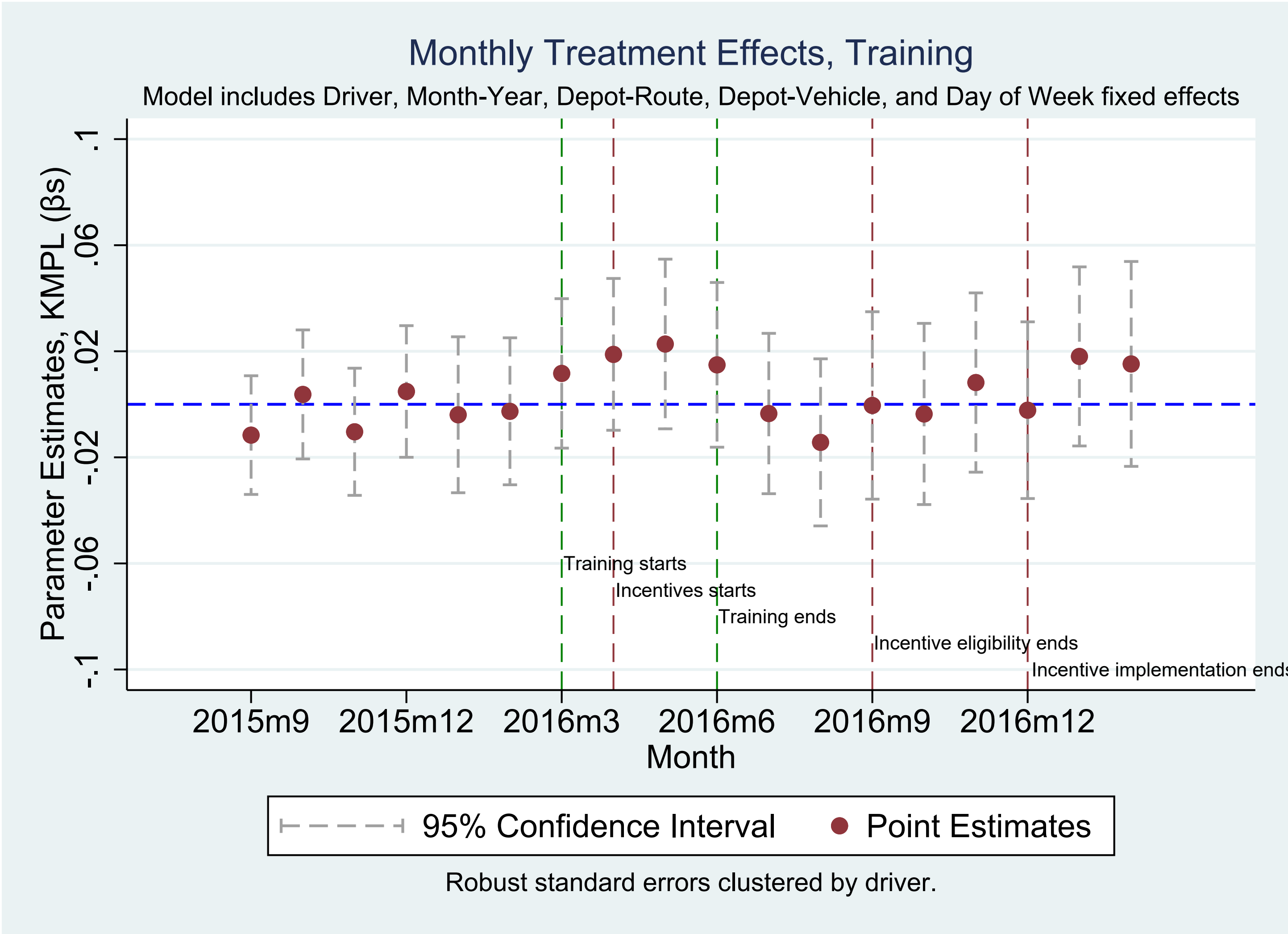


Figure 3: Monthly Treatment Effects, Incentives ($\beta_{2,m}s$)

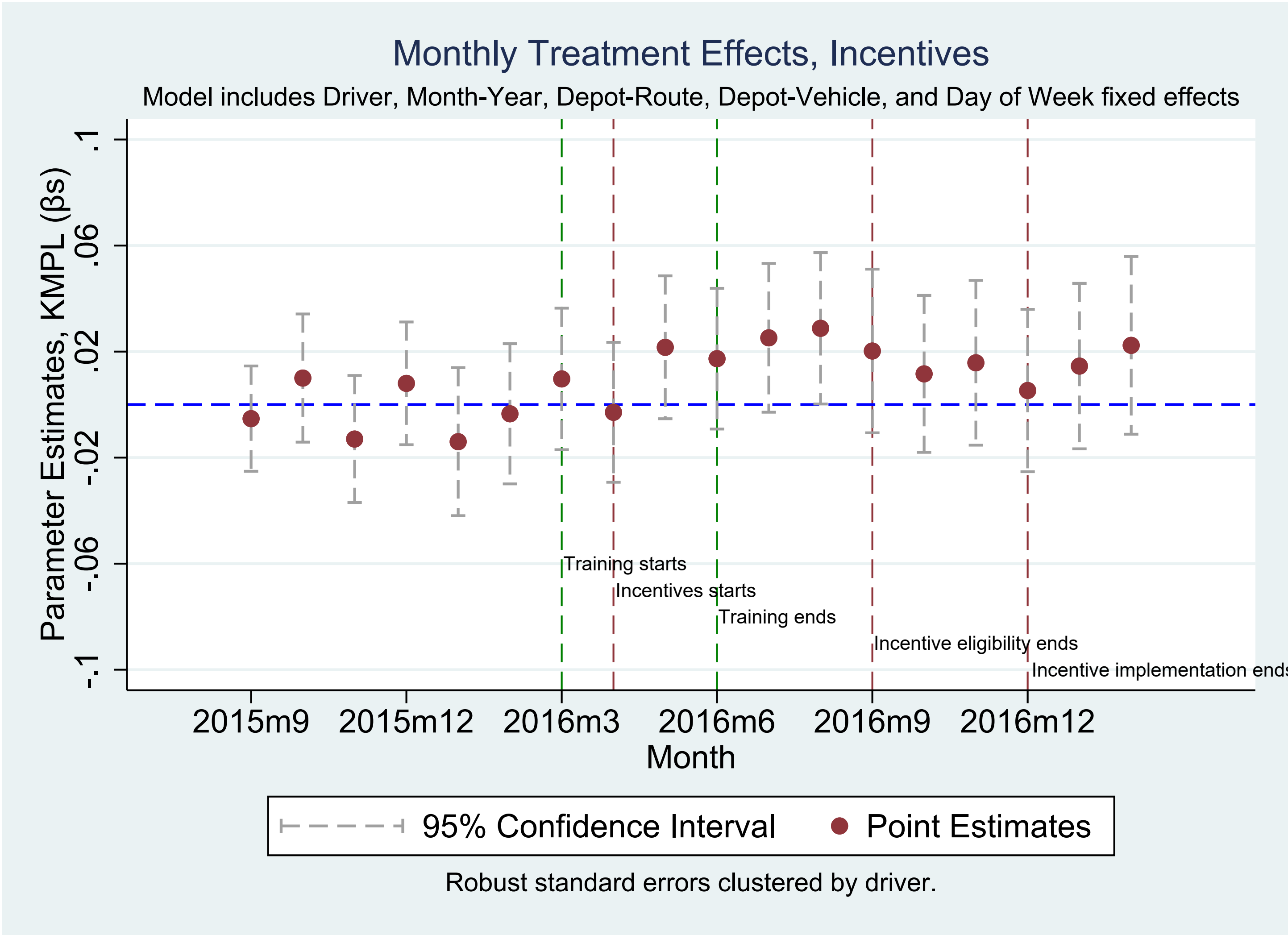


Figure 4: Monthly Treatment Effects, Training $\times$ Incentives ($\beta_{3,m}s$)

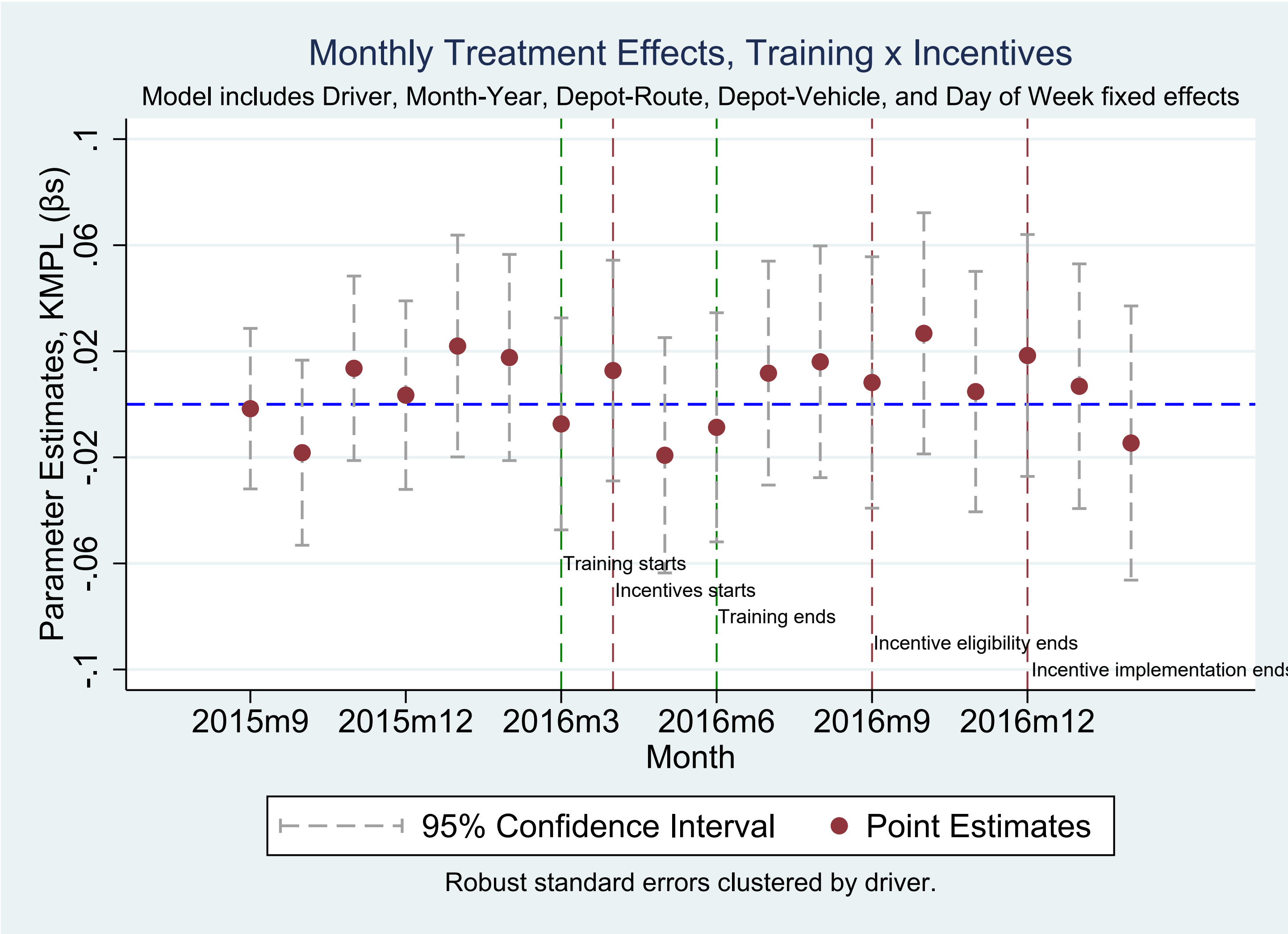


Figure 5: Cost-Effectiveness Analysis

	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
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**

[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 4: Impact on Fuel Efficiency in the entire 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 5: Impact on Fuel Efficiency in Four Sub Periods"

Tables

Table 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	38.81	38.94	38.99	.	0.326	0.253	-0.388
	(8.444)	(8.654)	(8.311)	(8.442)		(0.319)	(0.315)	(0.445)
Education (years)	10.22	10.25	10.31	10.29	.	-0.00170	0.0818	-0.00831
	(1.920)	(1.960)	(1.813)	(1.990)		(0.119)	(0.117)	(0.169)
KSRTC tenure (years)	10.02	10.14	9.903	9.979	.	0.345	-0.0523	-0.369
	(8.059)	(8.550)	(7.896)	(8.060)		(0.369)	(0.360)	(0.508)
Job satisfaction (score)	4.598	4.488	4.614	4.470	.	-0.103*	0.0318	-0.0379
	(0.801)	(0.892)	(0.747)	(0.891)		(0.0619)	(0.0564)	(0.0859)
Risk aversion (score)	3.037	3.021	2.997	3.166	.	0.00168	-0.0234	0.173
	(1.425)	(1.510)	(1.461)	(1.475)		(0.106)	(0.104)	(0.152)
Pro-socialness (Rs.)	25.01	24.58	24.28	24.53	.	-0.470	-0.713	0.787
	(6.765)	(7.763)	(7.826)	(7.716)		(0.536)	(0.528)	(0.773)
Consistently patient	0.582	0.577	0.608	0.552	.	-0.00343	0.0238	-0.0498
	(0.494)	(0.495)	(0.489)	(0.498)		(0.0357)	(0.0358)	(0.0506)
Dice points > 99p	0.0789	0.0923	0.101	0.0802	.	0.0119	0.0156	-0.0302
	(0.270)	(0.290)	(0.301)	(0.272)		(0.0206)	(0.0207)	(0.0295)
Dice points > 95p	0.176	0.161	0.161	0.203	.	-0.0189	-0.0262	0.0682*
	(0.382)	(0.368)	(0.368)	(0.403)		(0.0269)	(0.0270)	(0.0388)
Attended driving school	0.0394	0.0368	0.0341	0.0579	.	-0.000603	-0.00154	0.0249
	(0.195)	(0.189)	(0.182)	(0.234)		(0.0138)	(0.0136)	(0.0207)
Had major accident	0.147	0.156	0.181	0.185	.	0.00591	0.0341	-0.00474
	(0.355)	(0.363)	(0.386)	(0.389)		(0.0259)	(0.0271)	(0.0383)
Had minor accident	0.470	0.507	0.462	0.507	.	0.0431	-0.00393	0.00342
	(0.500)	(0.501)	(0.499)	(0.501)		(0.0362)	(0.0360)	(0.0512)
Attended KSRTC training	0.856	0.858	0.858	0.855	.	0.00225	0.00262	-0.00526
	(0.352)	(0.350)	(0.349)	(0.352)		(0.0254)	(0.0254)	(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. Section 4.4 has additional detail on these behavioral measures.

Table 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 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).

Table 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.

Table 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

Standard errors in parentheses
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
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 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

Standard errors in parentheses

Robust standard errors clustered by driver

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, detailed in Section 4.3

Phase 2 is July-September 2016, detailed in Section 4.3

Phase 3 is October-December 2016, detailed in Section 4.3

Phase 4 is January-February 2017, detailed in Section 4.3

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

Table 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 2.5.

Table 7: 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 Appendix Table A6 for details.)
Dice points > 99p indicates total dice points was above the 99th percentile of the theoretical distribution.

Table 8: 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.
Characteristic mean is from the post-attrition driver subsample. 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).

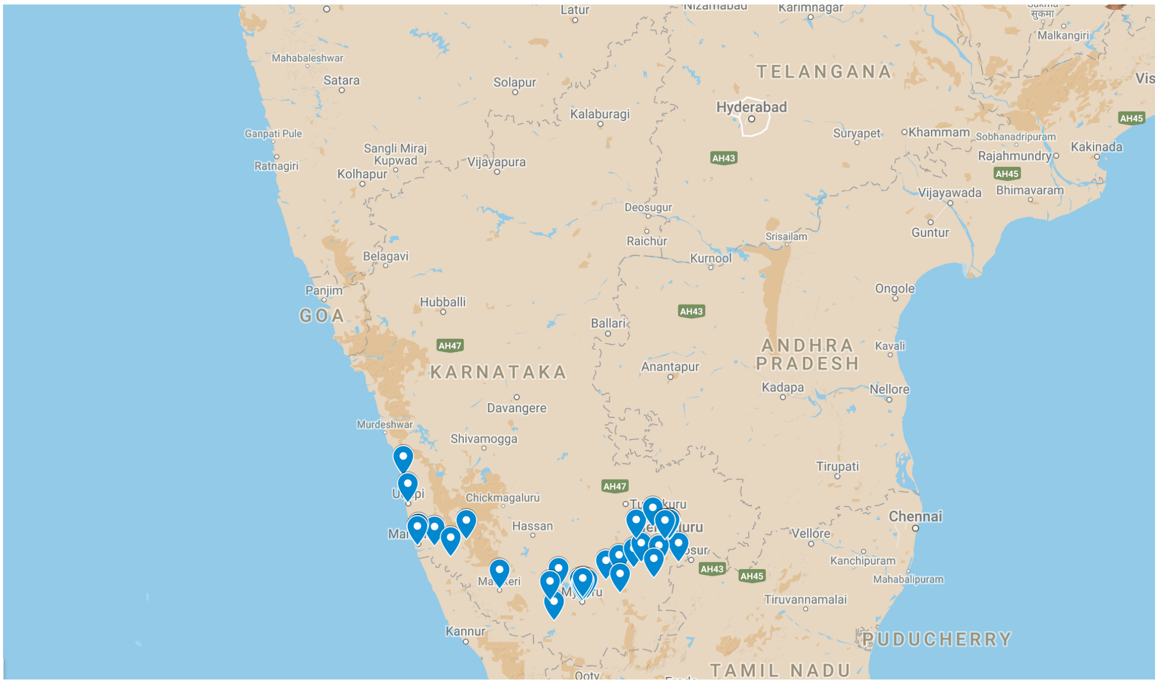
Table 9: Subgroup Analysis by Time Preference

	(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 Appendix Table A6
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.

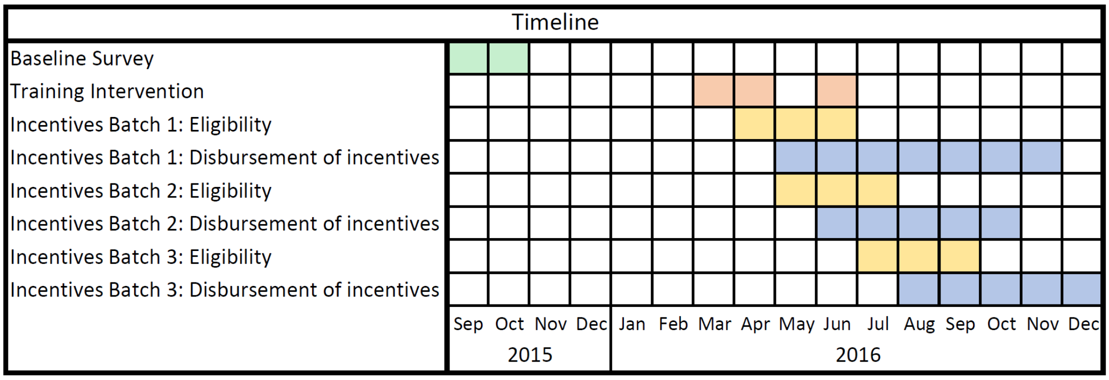
A Appendix Figures and Tables

Figure A1: 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 A2: Timeline of Interventions



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 A3: Monthly Treatment Effects, Training ($\beta_{1,m}s$), With and Without Controls

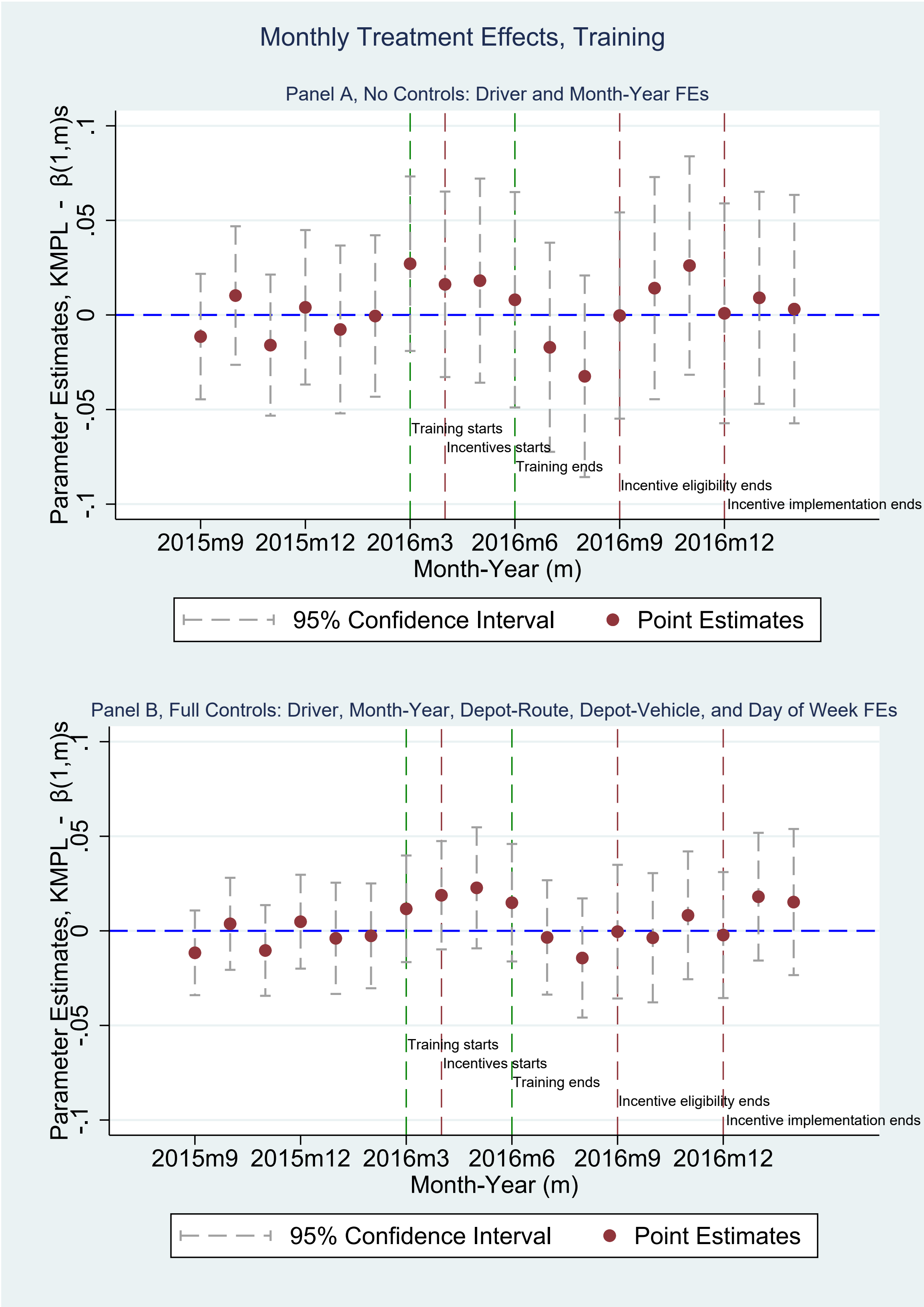


Figure A4: Monthly Treatment Effects, Incentives ($\beta_{2,m}s$), With and Without Controls

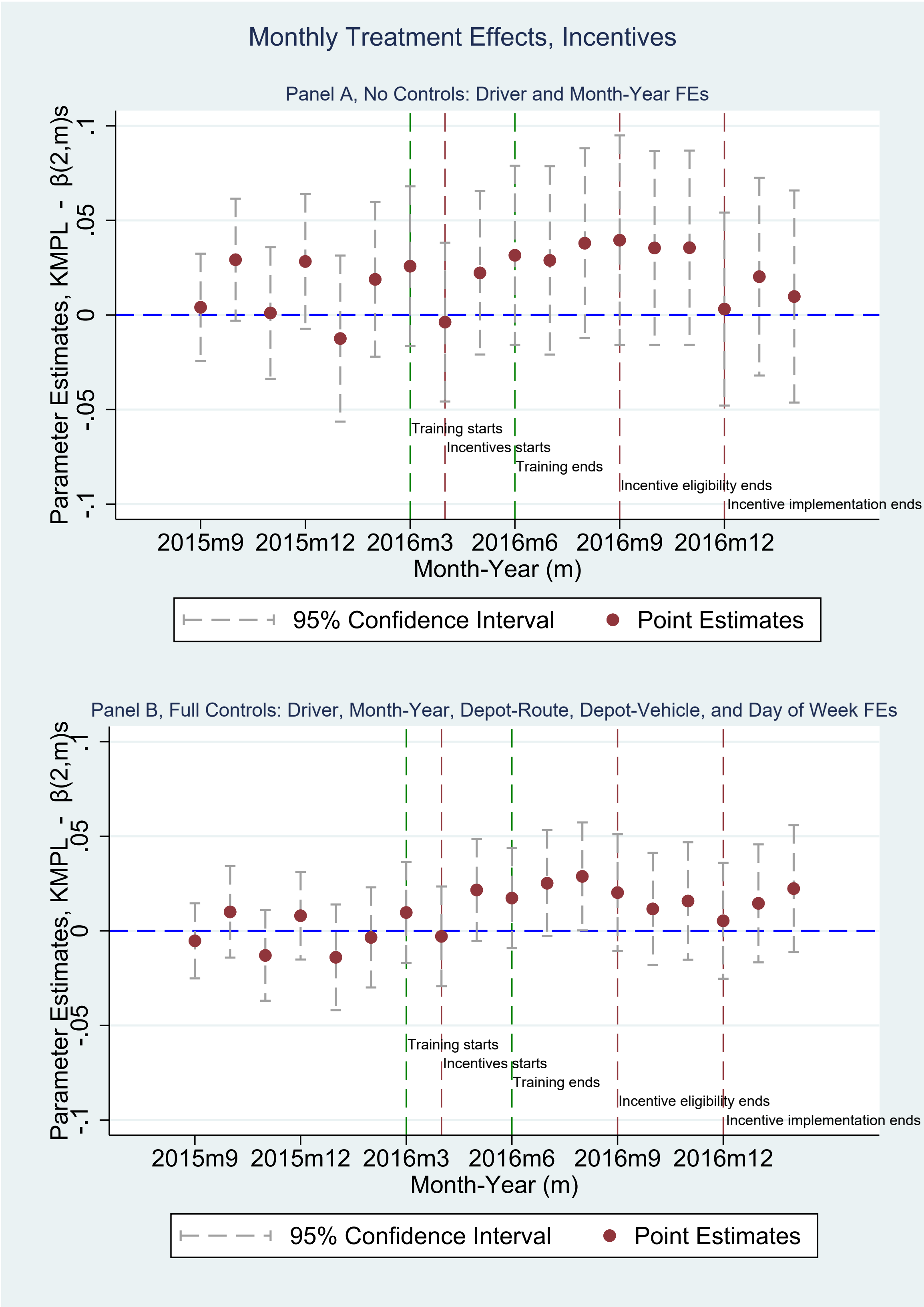


Figure A5: Monthly Treatment Effects, Training × Incentives ($\beta_{3,m}s$), With and Without Controls

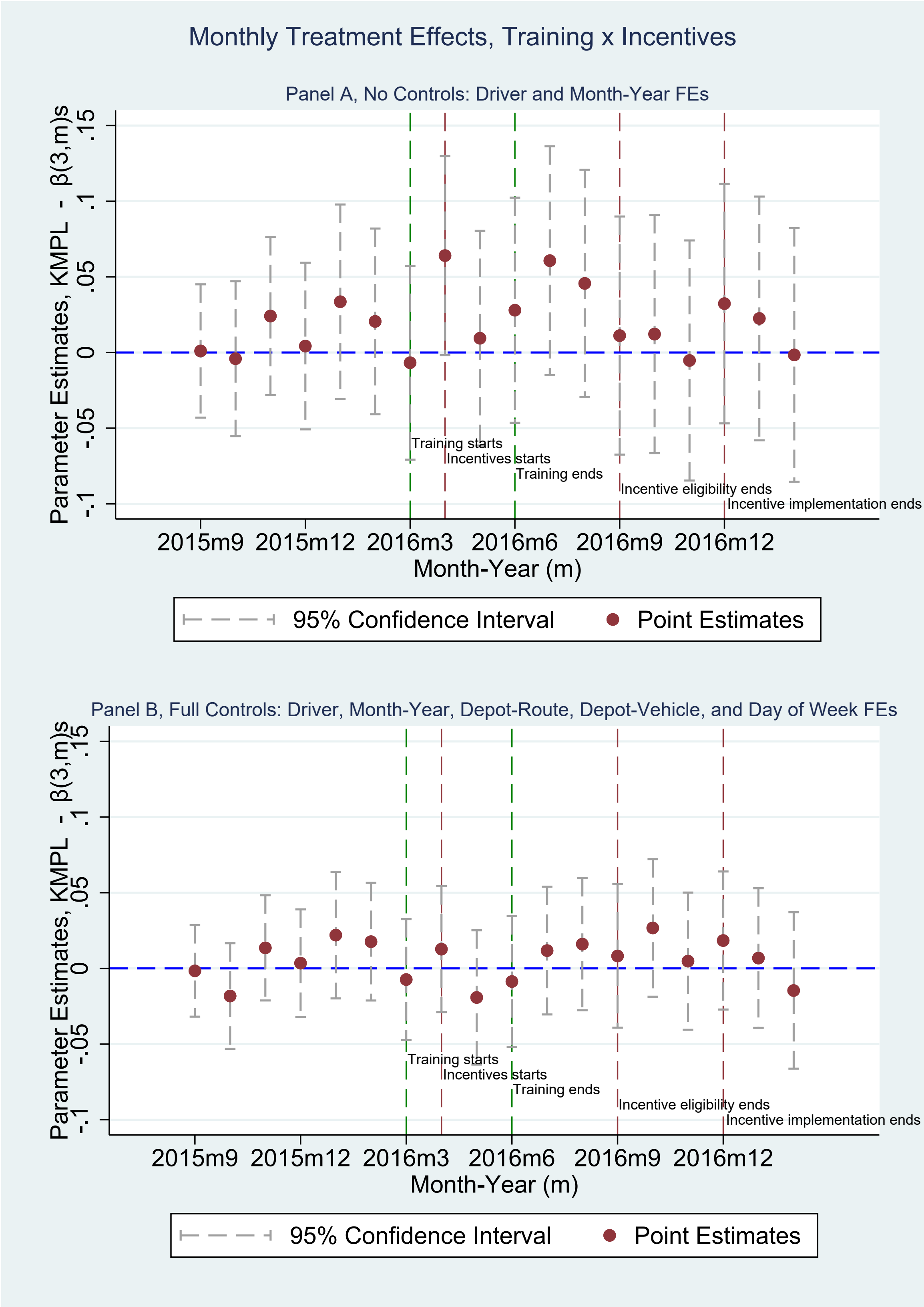
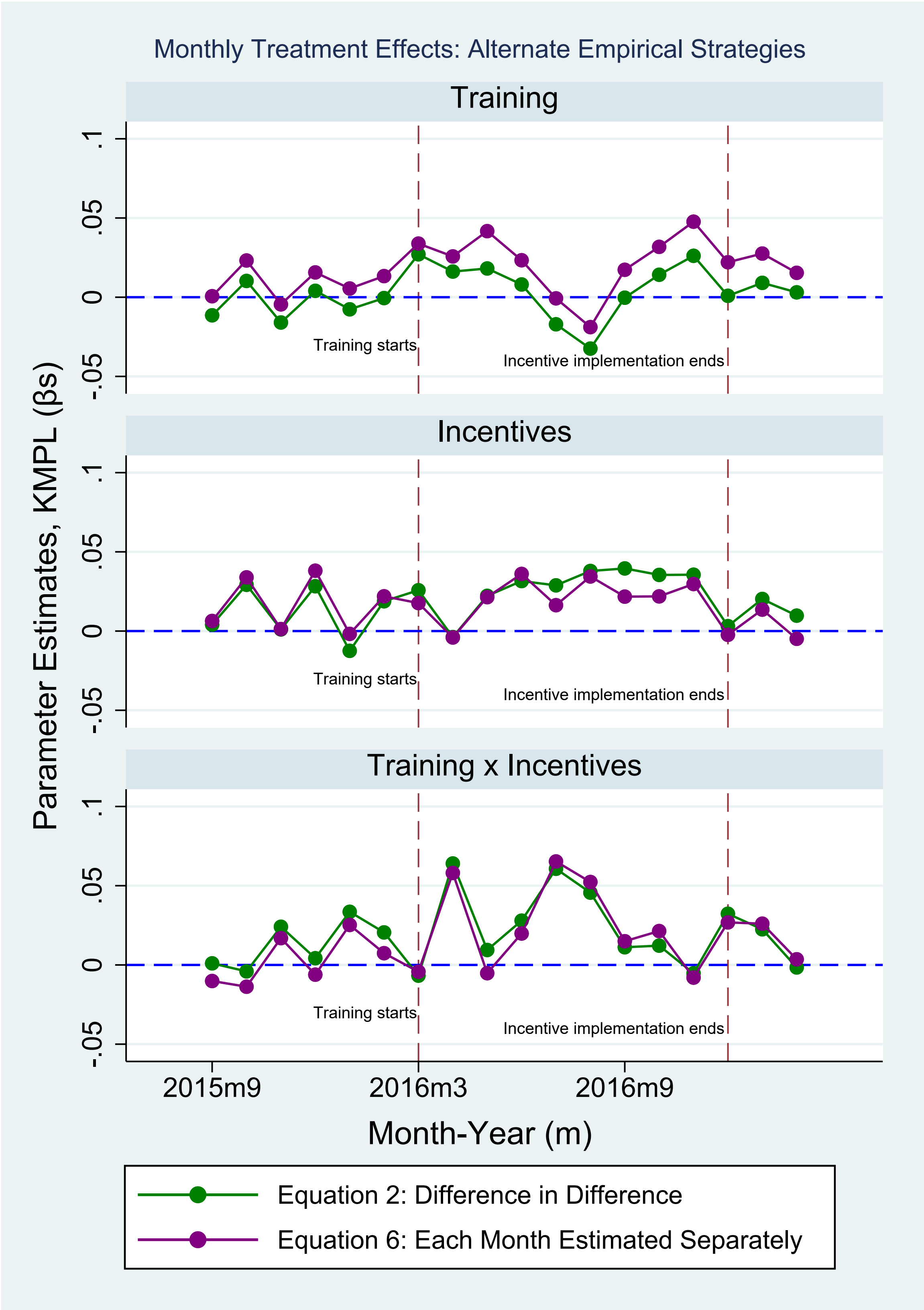


Figure A6: Monthly Treatment Effects: Robustness Check



The difference in difference model does not include controls, i.e. Column 1 of Table A5. Robust standard errors clustered by driver.

Table A1: 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 A2: 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 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 A3: Post-Attrition Subsample: Take-up and Compliance

	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	0.00557	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 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 A4: 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 A5: 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)
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)
Training $\times$ Incentives $\times 1(m = Nov2015)$	0.0241	0.0174	0.0141	0.0136
	(0.0266)	(0.0210)	(0.0177)	(0.0177)
Training $\times$ Incentives $\times 1(m = Dec2015)$	0.00425	0.00526	0.00334	0.00345
	(0.0281)	(0.0218)	(0.0181)	(0.0181)
Training $\times$ Incentives $\times 1(m = Jan2016)$	0.0336	0.0194	0.0219	0.0220
	(0.0327)	(0.0268)	(0.0213)	(0.0213)
Training $\times$ Incentives $\times 1(m = Feb2016)$	0.0205	0.0104	0.0179	0.0176
	(0.0313)	(0.0235)	(0.0198)	(0.0198)
Training $\times$ Incentives $\times 1(m = Mar2016)$	-0.00674	-0.0175	-0.00732	-0.00738
	(0.0326)	(0.0248)	(0.0204)	(0.0204)
Training $\times$ Incentives $\times 1(m = Apr2016)$	0.0640*	0.0189	0.0122	0.0127
	(0.0335)	(0.0261)	(0.0212)	(0.0212)
Training $\times$ Incentives $\times 1(m = May2016)$	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

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

Table A6: Summary 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.

Table A7: 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

Standard errors in parentheses

Robust standard errors clustered by driver

KMPL is winsorized at 1% and 99%

Only post-intervention shifts ($m >= March2016$) are used

Baseline KMPL is the average of KMPL in all pre-intervention shifts ($m <= February2016$)

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

Table A8: 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

Standard errors in parentheses

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.

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.

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

B 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 A2 and Section 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

$$\begin{aligned} Y_{imd\tau} = {} & \beta_1 Training_d \times Post_{i\tau} + \beta_2 Incentives_d \times Post_{i\tau} \\ & + \beta_3 Training_d \times Incentives_d \times Post_{i\tau} + \alpha_d + \delta_m + Controls_{imd} + \epsilon_{imd\tau} \end{aligned} \tag{7}$$

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. Results are presented in Appendix Table B1.

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

$$\begin{aligned} Y_{imd\tau} = {} & \sum_{\tau=-6}^{\tau=11} \beta_{1,\tau}(Training_d \times I_\tau) + \sum_{\tau=-6}^{\tau=11} \beta_{2,\tau}(Incentives_d \times I_\tau) \\ & + \sum_{\tau=-6}^{\tau=11} \beta_{3,\tau}(Training_d \times Incentives_d \times I_\tau) + \alpha_d + \delta_m + Controls_{imd} + \epsilon_{imd\tau} \end{aligned} \tag{8}$$

Where I_τ is a set of indicator variables for event-month τ, and the rest is as in Equation 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 B1, B2, and B3.

As described in Section 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 7 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 B1, B2, and B3 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.

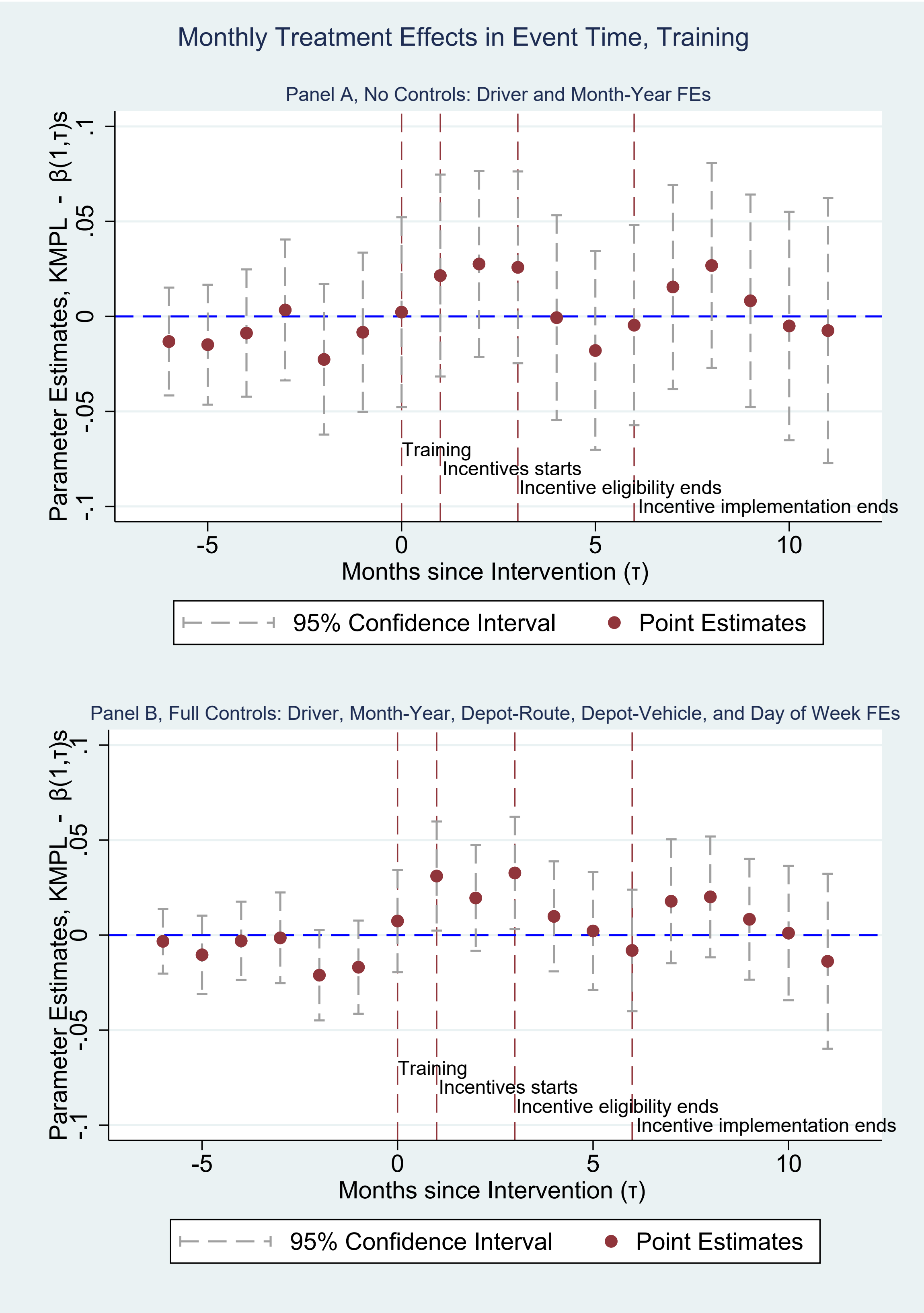
Overall, results are qualitatively similar with both the single treatment date and multiple treatment date approaches. The time patterns seen in the main paper for training and incentives are more clear with the event-time approach. The training program has a small positive impact in the short term for 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 B1: Impact on Fuel Efficiency using a Difference in Difference in Event-Time

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

Standard errors in parentheses
Robust standard errors clustered by driver
KMPL is winsorized at 1% and 99%
Post is an indicator for $Y\_imd\tau \in [\tau >= 0]$
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

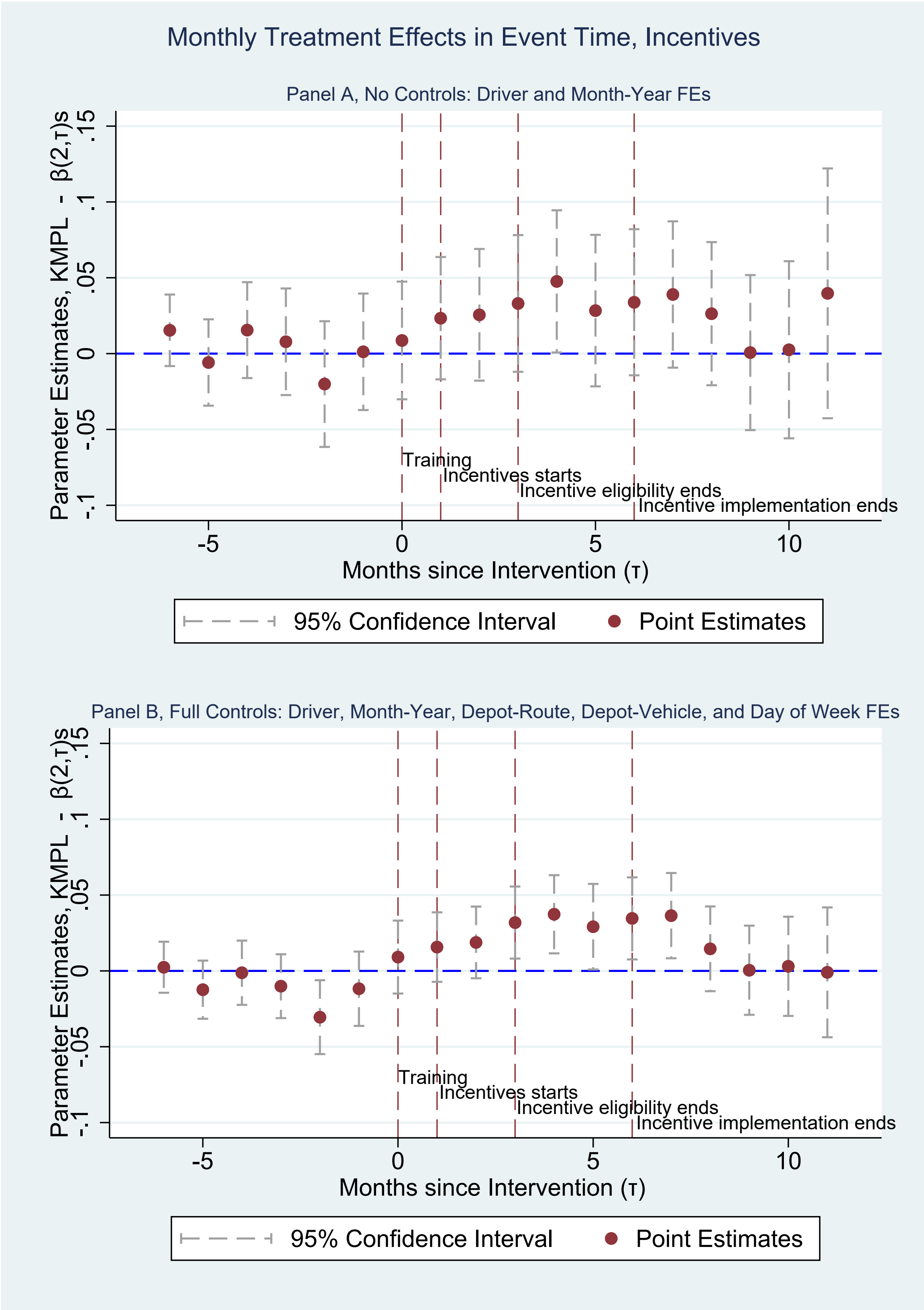
Figure B1: Monthly Treatment Effects in Event Time, Training ($\beta_{1,\tau}s$)



s

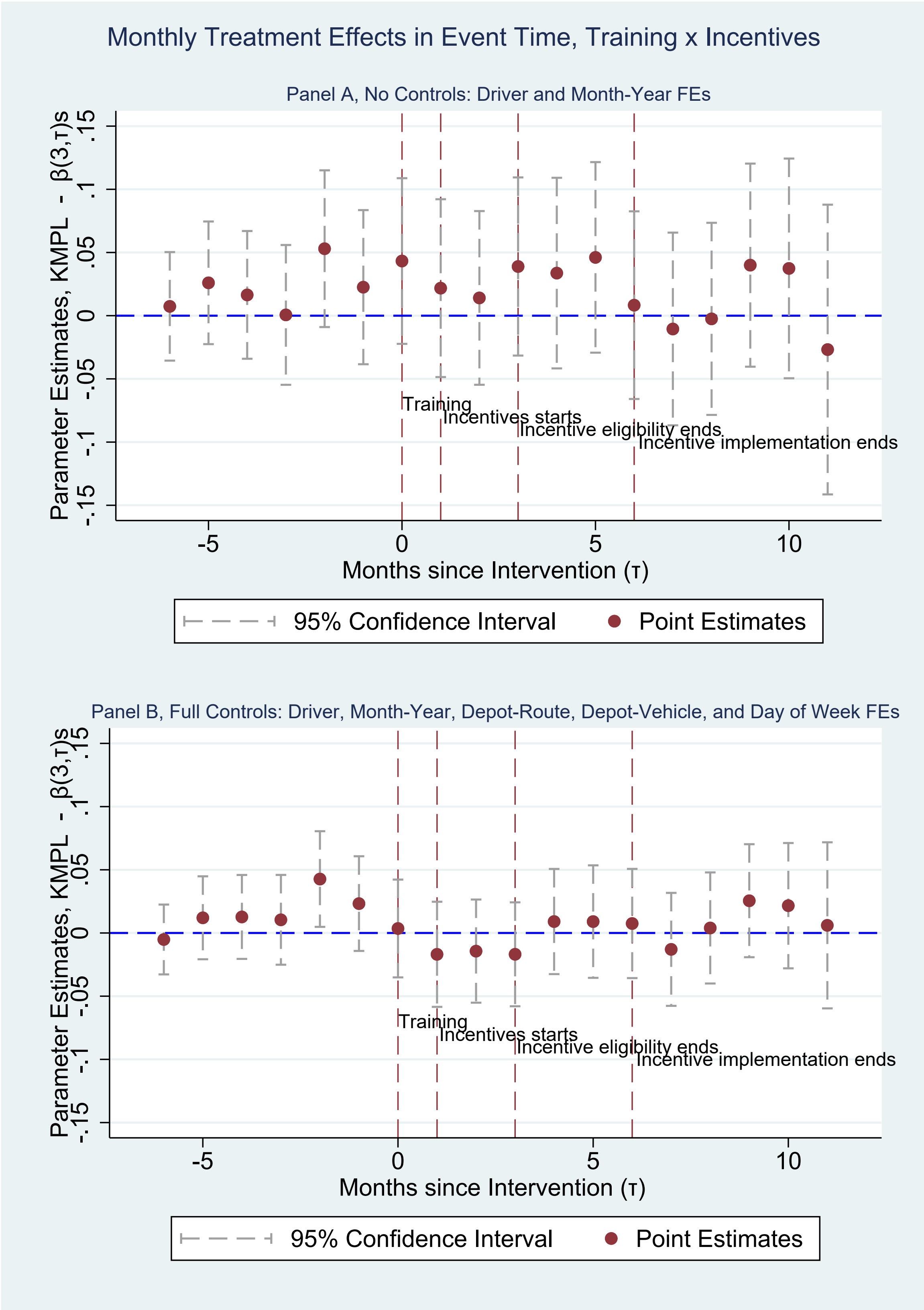
Robust standard errors clustered by driver.

Figure B2: Monthly Treatment Effects in Event Time, Incentives ($\beta_{2,\tau}s$)



Robust standard errors clustered by driver.

Figure B3: Monthly Treatment Effects in Event Time, Training × Incentives ($\beta_{3,\tau}s$)



Robust standard errors clustered by driver.