Title
Essays on Development Economics

Author
Carpena, Fenella

Publication Date
2017

Essays on Development Economics

by

Fenella C. Carpena

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Edward Miguel, Chair
Associate Professor Benjamin R. Handel
Professor Catherine D. Wolfram

Summer 2017

Essays on Development Economics

Abstract

Essays on Development Economics

by

Fenella C. Carpena

Doctor of Philosophy in Economics

University of California, Berkeley

Professor Edward Miguel, Chair

This dissertation is a collection of empirical essays on three critical microeconomic development issues in India—energy, financial education, and food security—spanning households in both rural and urban settings.

In the first chapter, I examine the adverse consequences rural electrification. Electrification programs have increasingly attracted worldwide interest as a policy tool to boost economic growth and transform the lives of the poor. While some studies have shown that electricity provision brings positive impacts by, among others, increasing employment and improving health outcomes, the potential *negative* effects of electricity provision remain unclear. This chapter argues that electrification unfavorably affects consumers that do not adopt electricity despite its availability. I use a quasi-experimental setting provided by India's national rural electrification program to study the effects of electricity provision on the kerosene markets. I show that when electricity becomes available, the price of kerosene—a substitute good for lighting—increases by 5-10%. This price increase subsequently hurts consumers who do not take up electricity and continue to use kerosene. I present a model that explains why kerosene prices might increase, and I show that the decline in the kerosene market size is a potential channel through which electrification resulted in higher prices of kerosene.

The second chapter considers urban India and household financial well-being. This chapter, co-authored with Shawn Cole, Jeremy Shapiro, and Bilal Zia, employs a large-scale field experiment to study the attitudinal, behavioral, and cognitive constraints that may stymie the link between financial education and financial outcomes. Our research design combines financial education with (1) monetary incentives for correct answers to a financial knowledge test; (2) financial goal setting; and (3) personalized financial counseling. We find no effects of cash incentives on participants' financial knowledge, but significant effects of both goal setting and counseling on real financial outcomes. In particular, combining goal setting with financial education encouraged relatively simple follow-up actions such as writing a budget or starting informal savings. Counseling, in turn, enabled the poor to undertake costlier or more difficult financial activities, including opening a formal bank savings account. Together,

these findings identify important complements to financial education that may successfully bridge the gap between financial knowledge and behavior change.

Finally, in the third chapter, I investigate the impact of dry rainfall shocks on food security among rural Indian households. Although a large literature has established that droughts lead to a significant decline in agricultural yields, much less is known about its effects on household nutrition. On the one hand, low levels of precipitation may reduce household food consumption, for instance due to higher prices or lower income. But on the other hand, such harmful effects may not materialize because trade, food storage, or household savings may offset the impacts of low precipitation. My findings show that a dry rainfall shock negatively impacts households' diet quantity and quality considerably. These negative impacts are evident not only in households' aggregate expenditure, caloric, protein, and fat intake, but likewise when disaggregated across different types of food. Additionally, I find no statistically significant effect of a dry shock on prices. These results therefore suggest that prices may be unresponsive to precipitation in India, so that the negative impacts of drought on food consumption and nutrition more likely come through income rather than prices.

To my papa

Contents

List of Figures

List of Tables

Acknowledgments

Only with the advice, encouragement, and generosity of others have I been able to complete this dissertation. My time in the PhD program has been *the* rollercoaster ride of my life, and a great many people have been there to celebrate the highs and to lift me up through the lows. Words cannot fully express my gratitude to every faculty member, every institution, and each and every individual who has helped me along the way. My mind thanks you. My heart thanks you. And my soul thanks you.

My warmest appreciation goes out to my dissertation chair, Ted Miguel, for his patient supervision throughout graduate school. It was his Econ 172 class at Cal, which I took as an undergraduate student, that got me interested in studying economic development. I am also deeply indebted to my dissertation committee members Ben Handel, for his constant optimism and guidance, and Catherine Wolfram, for her many insightful comments. This work would not have been possible without their support.

Many other professors have contributed enormously to my academic journey. I am grateful to Ben Faber, Fred Finan, Meredith Fowlie, Bryan Graham, Andres Rodriguez-Clare, and Reed Walker for their invaluable feedback and constructive suggestions. I also thank Betty Sadoulet and Jeremy Magruder for sharing their vast knowledge during the development field courses. I cannot begin to explain how Sol Hsiang's mentorship and encouragement have kept me grounded through many difficult times.

Life at the Berkeley Economics Department would not have been the same without our amazing staff. I am particularly thankful to Vicky Lee and Patrick Allen for their tireless efforts for all of us graduate students. I would not have known how to navigate every crossroad without them. Joe Sibol was always at the ready with answers to my questions. My special appreciation also goes out to Domingo and the rest of the custodial team for helping us keep an orderly and bright workspace, despite the dreariness of Evans Hall.

Beyond Berkeley, I would like to acknowledge Shawn Cole, Jeremy Shapiro, and Bilal Zia. Working with them has had a profound impact on my decision to apply to graduate school. During the first three years of my PhD studies, I received generous support from the National Science Foundation Graduate Research Fellowship. The final year of my dissertation was spent in Oslo, and I am grateful to the NUPI staff, especially Arne Melchior, Francesca Jensenius, and Hege Medin for providing an intellectually stimulating research environment as well as the faculty at BI for helpful discussions.

Friends both locally and around the globe have kept me sane during much of my graduate career. At Cal, many thanks to Alisa and Youssef, Dorian and Kristen, Gillian, Marion, Marc, Patrick, and Satoshi for all the lovely memories. I am very fortunate to have found a friend in Stuti in India and Sergiy in DC, both of whom never failed to boost my spirit. Scattered around the world, Anne, Charisma, Star, Jen, Mon, Nikki, Raissa, and my best buddy Kris have all been tremendous friends since our childhood. Though we don't always see each other, we always pick up right where we left off.

My family deserves endless thanks for their unwavering love and confidence in every single one of my endeavors—to my papa, who will always be the light of my life; to my mama, whose extraordinary tenacity I will forever admire; to my kuya Nicco and my pamangkins, Nixxie and Nacho, for reminding me that there are other things in life outside of academia; and to the rest of my family in the US, the Philippines, and everywhere in between, who inspire me to be stronger and to be the best person that I can be.

But family, to me, means more than just genetics. My infinite gratitude goes out to Tito Carlos and Tita Janet, whom I consider as my second set of parents. We may not be related by blood, yet they have treated me as one of their own with an incredible love and kindness that always made me feel at home. I also feel lucky to have an extended family in Belgium, with Ariane, Johan, meter Judith, meter Simonne, the De Lattes, and the Galles. Mijn liefje Simon has been my rock, and I would not have been able to cross the finish line without him.

This dissertation is dedicated to my papa, who passed away while I was in graduate school. He unfortunately did not live long enough to witness the completion of this work, but I know that it would have made him proud.

Chapter 1

The Adverse Effects of Electrification: Evidence from India

1.1 Introduction

In recent years, electrification programs have garnered increasing support from policy makers, international donors, governments, and non-government organizations worldwide, with the aim of bringing modern energy to the approximately 1.2 billion people—one in six of the global population—still living without electricity (International Energy Agency 2016). For example, as one of its Sustainable Development Goals by 2030, the United Nations has targeted ensuring access to "affordable, reliable, sustainable, and modern energy for all" (United Nations 2017). In 2009, global investments in electricity provision amounted to an estimated US$9.1 billion.[1] Today, billions of dollars continue to pour into electrification programs around the world, such as the *Luz Para Todos* ("Light for All") project in Brazil and the *Power Africa* initiative in Sub-Saharan Africa.

Typically designed to boost economic growth and reduce poverty, these large-scale investments in electrification are often expected to deliver substantial improvements in the lives of the poor. Indeed, some research suggests that access to electricity generates a number of positive impacts: it has been demonstrated, among others, to promote female employment, enhance agricultural productivity, improve health outcomes, and increase educational attainment.[2] Nevertheless, the existing literature overlooks the possibility that electrification may

[1]This estimate come from International Energy Agency (IEA) World Energy Outlook for 2011. In addition, the IEA projects that US$48 billion needs to be invested each year to meet the United Nations Sustainable Development Goal of achieving universal access to modern energy by 2030.

[2]Some examples include: Dinkelman (2011), who finds that electrification significantly increases female employment by releasing women from home production; Lipscomb, Mobarak, and Barham (2013), who document large development gains from electrification; Barron and Torero (2017), who show that household electrification leads to substantial reductions in indoor air pollution; and Chakravorty, Emerick, and Ravago (2016), who report substantial short-run welfare gains from electricity expansion, much of which comes from increased agricultural income.

also result in unintended *negative* effects. Understanding these unfavorable consequences is critical to obtain a more comprehensive view of the impact of electricity provision, but it has thus far received only little attention in the ongoing research and policy discourse surrounding electrification programs.

In this paper, I argue that electricity provision adversely affects one particular segment of consumers: those that do not adopt an electricity connection despite its availability. Given that illuminants are one of the main uses of electricity for newly connected households, I provide evidence that electrification impacts the market for alternative lighting sources. Specifically, I show that when electricity becomes available, the price of a substitute good for lighting—namely, kerosene in my empirical setting in rural India—rises by 5 to 10%. This considerable price increase hurts households that forgo electricity and remain kerosene consumers. Importantly, these electricity non-adopters consist primarily of those at the bottom of the income distribution, who also devote more of their scarce earnings on fuel and lighting expenditure than those at the top.

I estimate the causal effects of electrification on kerosene prices by taking advantage of the staggered implementation of India's national rural electrification program, the *Rajiv Gandhi Grameen Vidyutikaran Yojana* (RGGVY). Launched in 2005, the program mandated the construction of basic electricity infrastructure as well as offered households below the poverty line with free electricity connections. Since the program was administered at the district level, my empirical strategy uses differences-in-differences, exploiting variation across districts in both the timing and intensity of the program. Causal identification in this approach requires trends in treatment and control districts to have been identical in the absence of the program. I test this proposition directly using an event study analysis, and the results lend strong support to the validity of the differences-in-differences design.

With a unique combination of market-level information from the Consumer Price Index (CPI) micro-data together with household-level data from the National Sample Surveys (NSS), I investigate the effects of rural electrification on two sets of outcomes. First, I demonstrate that the program led to a statistically significant 1.2% point increase in the propensity to use electricity for the average rural household. This is accompanied by a one-for-one decrease in kerosene use, supporting the idea that electricity and kerosene are substitute forms of lighting. Yet the results also indicate that although electricity is available, a substantial proportion of low-income households still continue to rely on kerosene for lighting. Given that the program provided free connections to households below the poverty line, this result suggests that poor households in this setting face multiple barriers to adopting electricity apart from the costs of household connection.

Second, I show that as a conseqeuence of the program, the market price of kerosene rises by 5 to 10%. This result is robust to different measures of prices as well as regression specifications. Using unit values of kerosene as a proxy for price, I find that the program is associated with a statistically significant increase of Rs. 0.81 in *nominal* and Rs. 0.44 in *real* prices, compared to pre-treatment means of Rs. 15 and Rs. 13.7, respectively. But because unit values may be measured with error and may lead to attenuation bias, I likewise estimate effects on prices by employing CPI micro-data, which is available for about half of all districts. Doing so, I find very similar patterns: the program corresponds with substantially higher prices, at Rs. 1.30 in *nominal* and Rs. 0.92 in *real* terms, both significant at the 1% level. Moreover, these findings hold up whether considering prices in levels or logs. Taken together, the results therefore show that the price increase is unlikely to be due to biases from measurement error, price data availability, or misspecification of the functional form.

That kerosene prices rise as a result of electricity provision is perhaps a somewhat surprising and counterintuitive result, as basic insights from economics suggest that the entry of electricity would lead to *lower* prices of kerosene, a substitute good for lighting. Hence, to investigate theoretically why prices might increase, I turn to the standard model of spatial differentiation typically used in retail contexts due to Salop (1979). In this model, both kerosene sellers and buyers are located on the circumference of circle. Sellers incur a fixed cost of entering the market and compete directly with each of the firms situated to its left and right, while buyers select the retailer that maximize utility given its price and geographic location. Such a model provides an appropriate foundation for my setting because geography represents the main differentiating factor between competitors, and kerosene is otherwise a homogenous good. Furthermore, in practice, customers must travel to the kerosene seller's shop and hence must pay a transportation cost when purchasing the product.

My theoretical framework then augments the Salop (1979) model in two ways. First, whereas agents in Salop (1979) differ only in their location on the circle, I introduce consumer heterogeneity so that individuals either have zero ("type L") or strictly positive ("type H") costs of adopting electricity. Second, I consider how the entry of a firm that is *not* located on the circle (i.e., electricity) impacts retailers that are on the circle (i.e., kerosene sellers). The results from this theoretical framework suggests that the kerosene *market size* is a potential channel linking rural electrification and higher kerosene prices: the model shows that if the price of electricity is relatively low and if both electricity adoption costs and the share of H-types are high, then an equilibrium outcome exists where all L-types choose electricity. Intuitively, this finding indicates that with a smaller pool of kerosene buyers, each seller's average costs of production rises, and as a result, kerosene retailers must hike their prices in equilibrium to continue operating in the market.

This paper builds on a growing literature on the impact of electricity provision in developing countries. In particular, while many studies have considered labor market, health, education, and earnings outcomes, to my knowledge, no existing paper investigates effects on prices of alternative fuels and potential consequences for consumers who do not adopt electricity. This is a critical policy issue especially because many government and non-government organizations around the world continue to implement electrification programs. At the same time, there is an ongoing public debate in countries such as India about whether to phase out kerosene subsidies altogether. If electrification has the unintended consequence of increasing the price of kerosene paid by electricity non-adopters, then it would be important to consider the adverse effects of electrification on such households relative to the potential benefits that electricity may bring.

1.2 Motivating Facts: Energy Use in Rural India

How do rural households in India use fuel and energy in their day-to-day lives? In this section, I use survey data from the National Sample Survey Office, a central government agency responsible for conducting socio-economic surveys, to examine the energy consumption of rural households.[3] Using the survey's sampling weights to obtain nationally representative estimates, I show five separate but related patterns regarding rural energy use. Many of these same patterns have already been documented in a number of previous studies (e.g., World Bank 2003; Barnes, Krutilla, and Hyde 2004; Jain et al. 2015). Nevertheless, in this paper, I highlight the following descriptive statistics as motivating evidence for the potential consequences of electricity provision on rural energy markets.

Motivating Fact 1. Rural households use electricity primarily for lighting and not cooking. Figure 1.1 illustrates the main energy sources used by households for two of the most basic energy-consuming domestic activities: lighting and cooking. In this figure, we see in the left panel that 55% of rural Indian households typically use electricity for their lighting needs. In contrast, in the right panel, we see that almost none use electricity for cooking, with the vast majority of households using firewood instead. These patterns indicate that lighting is likely the first application of electricity among newly grid-connected households in rural India. As a consequence, we would expect electricity provision to directly impact the market for alternative energy sources for lighting, rather than cooking.

Motivating Fact 2. Next to electricity, kerosene is the second most common energy source for lighting among rural households. Figure 1.1, Panel A also shows that kerosene is the lighting fuel of choice among households not using electricity, with 44% of rural households using kerosene as their primary source of lighting. Notably, nearly all households (i.e., 99%) use either electricity or kerosene for lighting; only a very small fringe use other lighting sources such as candles, gas, or other oils. These findings suggest that electricity and kerosene are close substitutes for lighting, and households rely on kerosene when electricity is unavailable or unreliable. Kerosene markets in rural areas are therefore tightly linked with power sector reforms, particularly with rural electrification programs that increase the availability of household electricity connections.

Motivating Fact 3. Electricity (kerosene) use is positively (negatively) correlated with income. Plotting income versus electricity or kerosene use, Figure 1.2 reveals that the overall energy consumption patterns presented earlier in Figure 1.1 belie substantial heterogeneity across households. Here, I follow India's National Sample Survey Organization

[3]Specifically, I use data from NSS Consumer Expenditure Survey Round 61 (2004-2005), which is the most recent round available prior to the implementation of RGGVY. Furthermore, Round 61 is a "thick" round of the NSS surveys, meaning that the sample size is larger than in other rounds. See Section 1.4 for more details about the NSS data.

(NSSO) and use household monthly per capita expenditure (MPCE) as a proxy for income. This figure clearly demonstrates that the likelihood of using electricity as a main source of lighting steadily increases with earnings, while the reverse is true for kerosene. Although these correlations are not causal evidence on the role of income in electricity adoption, they do highlight one possible avenue through which rural electrification may have important distributional consequences. That is, when electricity becomes available, only rich households may be able to connect and use electricity, while poor households may remain unconnected and continue to rely on kerosene.

Motivating Fact 4. Poor households devote a significantly larger percentage of their total expenditure on fuel and lighting than rich households. To further explore energy use across the income distribution, Figure 1.3 depicts fuel and lighting as a share of total household consumption by MPCE deciles. The fuel and lighting budget share monotonically decreases with income, though this relationship is highly non-linear. For households in the two lowest income deciles, fuel and lighting represents a substantial portion of total expenditure at 12.1 %. But for those in the highest-earning decile, the fuel and lighting share in total consumption is much lower at only 8.4 %. This difference is highly statistically significant. Moreover, it implies that higher energy prices will disproportionately impact the poor, who must spend even more of their meager resources on fuel and lighting than their rich counterparts.

Motivating Fact 5. Poor households spend a larger percentage of their total expenditure on kerosene than rich households. Whereas Figure 1.3 considers all types of fuel and lighting expenditure, Figure 1.4 focuses only on kerosene consumption. In India, households may purchase kerosene either through the Public Distribution System (at government-subsidized prices, in rationed quantities) or through the open market (at non-subsidized prices). Whether looking at expenditure for both non-subsidized and subsidized kerosene, subsidized kerosene alone, or non-subsidized kerosene alone, Figure 1.4 shows that the proportion of kerosene spending in the household budget falls as income rises. These patterns echo the results in Figure 1.3, emphasizing that low-income households are more vulnerable to kerosene price increases.

In summary, the patterns described above indicate that electricity and kerosene are close substitutes for lighting, so that increased availability of the former may impact prices of the latter. Importantly, changes in the price of kerosene may have significant consequences for the welfare of the poor, who rely on kerosene to a much greater extent than wealthy households. Given that this study's objective is to understand how electricity provision impacts the market for alternative sources of lighting, for the rest of this paper, I focus mainly on *non-subsidized* kerosene (for which prices are determined by market forces) rather than subsidized kerosene (for which prices are set by government institutions).[4] In particular, I examine non-subsidized kerosene in the context of India's national rural electrification program, which I describe in greater detail in the next section.

1.3 Empirical Context

The Rajiv Gandhi Grameen Vidyutikaran Yojana (RGGVY) or the "Prime Minister's Rural Electrification Scheme" is an ambitious, large-scale, nationwide electrification program which aimed to provide electricity access to all rural villages and households.[5] Launched by the Government of India in 2005, it was implemented primarily during the country's 10th (2002-2007) and 11th (2007-2012) Five-Year Plans.[6] The program's core components are two-fold. First, RGGVY mandated the construction of basic electricity infrastructure—namely, at least one substation in each block, at least one transformer with sufficient capacity in each village, and transmission and distribution lines. Where such infrastructure already exists, the program augmented the extant systems for "intensive electrification," while in remote villages where grid supply is neither feasible nor cost effective, the program installed small off-grid, decentralized generators, and distribution networks. Second, RGGVY offered free electricity connections to all households below the poverty line (BPL). Households above the poverty line (APL) did not receive the same benefit; however, they were able to obtain a connection by submitting an application to the state utility office and paying a one-time connection fee, typically around Rs. 3000 (US$ 45).[7]

Although RGGVY was conceived at the national level, the program was in practice executed through separate, district-level projects. These projects were carried out by local institutions but were heavily subsidized by the central government at 90 percent for infrastructure (with the remaining 10 percent available as loans) and 100 percent for all BPL connections. To initiate an RGGVY electrification project in a given district, the local implementing agency—for example, a distribution company, state electricity board, or central public sector utility—first prepared a project proposal called the "Detailed Project Report" (DPR). Each DPR covered only one district; for that district, the DPR described a comprehensive project plan such as the types of infrastructure to be installed, the number of APL and BPL households to be connected, and estimated costs. In addition, the DPR identified those villages in the district which qualified for RGGVY. Specifically, during the 10th Plan,

[4] To the best of my knowledge, there were no major changes to subsidized kerosene provision during the period I study.

[5] This section draws information from government reports (e.g., Standing Committe on Energy 2009, 2013; Programme Evaluation Organisation 2014), scholarly articles, and in-depth conversations with practitioners and stakeholders. RGGVY excluded Union Territories, and given its focus on rural electrification, it did not include urban areas.

[6] RGGVY was also initially implemented during India's 12th Five-Year Plan (2012-2017). However, beginning in 2014, RGGVY has been subsumed into the Deen Dayal Upadhyaya Gram Jyoti Yojana (DDUGJY).

[7] Under the program, all households were required to pay for their own electricity consumption; it was not subsidized by RGGVY.

the program stipulated that only villages with population above 300 were eligible for electrification, but in the 11th Plan, this threshold was reduced to 100. Districts may also prepare multiple DPRs over time, and the administrative program data show that about half of all districts have implemented at least 2 DPRs.

Proposed electrification projects outlined in these DPRs required sanctions from the Rural Electrification Corporation (REC) before implementation can begin. A public sector financial institution acting as the "nodal agency" for RGGVY, the REC was responsible for complete oversight of the program including establishing the implementation framework, formulating procurement and bidding conditions, releasing subsidies provided by the national government, and monitoring ongoing RGGVY projects. Moreover, the REC coordinated the approval process for RGGVY electrification proposals: they reviewed DPRs to ensure compliance with program guidelines and thereafter forwarded DPRs to the Ministry of Power (MOP), a central government agency, for approval. Both the REC and the MOP generally reviewed and sanctioned DPRs on a first-come, first-served basis. Once the DPR is approved, the local implementing agency issued a public tender for the project, evaluated bids, and awarded the project contract. At the final approval stage, the REC released the first tranche of program funds amounting to 30% of sanctioned project costs, so that project implementation may commence.

Importantly, the above administrative procedures gave rise to staggered implementation of RGGVY across districts. All districts in India were eventually covered by RGGVY, except those districts in Union Territories and urban areas that are not included in the program. Throughout this paper, I define the "implementation date" of RGGVY in a particular district as the date in which the REC first disbursed RGGVY program funds for that district, as data on actual commencement or completion dates of project fieldwork are unavailable. With this definition at hand, Figure 1.5a illustrates the roll out of RGGVY over time and demonstrates substantial differences in the timing of program introduction across districts. This temporal variation thus allows for a comparison between districts that have already implemented RGGVY and those that have yet to do so. Furthermore, Figure 1.5b plots BPL household connections covered by RGGVY—a measure of the program's intensity—normalized by the total number of households in the district based on the 2001 Census. As before, this figure shows considerable heterogeneity across districts in RGGVY "dosage." My basic empirical approach uses differences-in-differences by taking advantage of these district-level variations in both the timing and intensity of RGVVY, to shed light on how electricity provision influences rural kerosene markets.

Because the roll out of RGGVY over time is central to my empirical design, it is useful to understand the differences between districts that implemented the program earlier versus later. Table 1.1 presents the determinants of RGGVY implementation date using pre-treatment district characteristics from the 2001 Census, four years before the program was launched. In this table, the dependent variable is the month and year in which the district began implementing RGGVY, expressed as an index equal to one beginning in March 2005. The OLS regression results point to two important patterns: (1) districts where villages are relatively more developed (e.g., have facilities such as schools, post offices, telephones, banks) and have a lower proportion of Scheduled Castes in the population are more likely take up RGGVY later; (2) districts where fewer villages have power supply and where more villages are eligible for RGGVY tend to implement RGGVY earlier. Taken together, these findings are consistent with the program eligibility rules as well as the program's overall focus on accelerating rural development. In addition, they suggest that the early adopters of RGGVY are those districts that stood to gain the most from the program.

Despite the significant effects of pre-treatment district characteristics on the timing of program implementation, the identification assumption in this study requires that early and late implementers of RGGVY would have otherwise changed similarly in the absence of the program. I address the validity of this assumption in several ways. First, I provide a direct test of identical counterfactual trends using event study models, explained in more detail in Section 1.5. Second, because initially different districts may also be more likely to change differently, I include interactions of the statistically significant 2001 baseline district characteristics from Table 1.1 with linear time trends in all regressions. I therefore control for observable, parametric differences in trends across districts that may be spuriously correlated with the RGGVY treatment effect. Finally, Figure 1.6 provides scatter plots of baseline district characteristics against RGGVY implementation dates, demonstrating that much of the variation in program timing remains unexplained. Thus, RGGVY adoption dates may potentially be idiosyncratic, conditional on these district characteristics.

1.4 Data Sources

The main datasets I use in this study consist of the following: (1) the National Sample Survey (NSS) from India's Ministry of Statistics and Program Implementation, (2) district-level administrative records from RGGVY, and (3) district-level characteristics from the 2001 Census of India. In what follows, I provide more information on each of these datasets as well as the data construction.

National Sample Survey (NSS). The NSS is an annual socio-economic survey collecting a rich set of household-level information including household characteristics, consumption, living conditions, sanitation, and employment. The survey covers all of India except for inaccessible villages in the union territory of Andaman and Nicobar Islands, and the states of Jammu and Kashmir, and Nagaland. In this paper, I focus on the *Consumer Expenditure* survey module of the NSS, which interviews households regarding their expenses for more than 350 items (including fuel and lighting) over a 30 day period. The survey likewise asks households to report their primary source of energy for lighting and cooking. I employ the ten most recent survey rounds in which this module was fielded by the NSS; these are rounds 57 to 64, 66, and 68, corresponding to years 2001-2007, 2009, and 2010.[8]

[8]Generally speaking, each NSS round lasts for 12 months from July to June. For example, Round 68 was implemented from July 2011 to June 2012. Throughout this paper, I will refer to each NSS round using the year in which the given round started.

Since the NSS is not a panel of households and village identifiers are masked in the public-use files, I use the NSS sampling weights to collapse the household data to the district level, thus building a panel of district-level variables. From these data, the key outcome variables I examine are two-fold: first, the percentage of households in a given district-round that use kerosene or electricity as their main lighting source; and second, the median price per liter of kerosene across all household responses in each district-round. While the NSS does not have price information *per se*, the data indicates both the value and quantity of consumption. I therefore use unit values (i.e., the value in Rupees divided by quantity consumed) as a proxy for price. As this price measure is quite noisy, I take the median unit value within each district and round to avoid the influence of outliers.

Because the above outcome variables are implemented at the district level, it is important to consider the issue of redistricting. District boundaries are quite stable in the period I study: of more than 600 districts covered in the NSS between 2001 to 2010, only 51 districts experienced boundary changes. Nevertheless, in creating a district panel, I address redistricting in the following manner. In the common case where district B is created from a subset of district A, then district B is matched back to district A. It is also possible that district B is created by putting together a subset of district A and a subset of district C. In these rare instances, I combine districts A and C together and track the combined district $A + C$. With this process, the resulting NSS district panel contains a total of 583 districts.

RGGVY Administrative Records. For information on the implementation of RGGVY, I rely on the program's administrative records, specifically from Report D entitled "Physical & Financial Progress of RGGVY Projects Under Implementation (District-wise)."[9] The RGGVY administrative data contains 553 districts, and I merge them to the NSS by using state and district names, yielding a 100% match. For each district covered under RGGVY, I draw two important pieces of information from the administrative records. The first is the date when initial tranche of project funds were awarded. As previously mentioned in Section 3, I use this date as the "implementation date" since data on when distribution infrastructure were completed is not available. The second is the total number of BPL household connections covered under the project. This data provides a continuous "treatment" variable, measuring the intensity of the program within the district.

Census of India, 2001. Aggregating village-level data from the *Village Directory* of the 2001 Census of India to the district level, I obtain district characteristics which are not collected by the NSS Consumer Expenditure Surveys. These include total population of the district, the number of households in the district, and the proportion of villages in the district with various infrastructure (e.g., educational facility, bus services). Since these data come from 2001, four years before the RGGVY program was launched, they provide baseline measures of district characteristics prior to the electrification treatment.

[9]See the RGGVY program website: http://rggvy.gov.in/rggvy/rggvyportal/index.html

1.5 Empirical Methodology

This paper aims to investigate how rural electrification impacts kerosene markets. Consequently, a central requirement for the study is the availability of kerosene market data. Although as described previously in Section 1.3, the RGGVY program involved a population eligibility rule wherein only villages with more than 300 (100) inhabitants qualified for electrification under the 10th Plan (11th Plan), a village-level regression discontinuity approach—and more generally, any village-level empirical design—is unfortunately not feasible due to data constraints.[10] In particular, kerosene market data spanning the RGGVY program do not exist at the village level, and to the best of my knowledge, the finest geographical unit available for such data is at the district level. Hence, my empirical strategy concentrates on districts as a unit of analysis, taking advantage of the staggered implementation of RGGVY across districts over time in a differences-in-differences framework.

I exploit the variation in the timing of RGGVY to estimate three sets of district-level regressions. The first is a standard difference-in-differences model given by

$$y_{dt} = \beta RGGVY_{dt} + \gamma_d + \lambda_t + \delta \mathbf{X}_{d2001} t + \epsilon_{dt}. \tag{1.1}$$

In this equation, the left hand side variable, y_{dt}, denotes kerosene market outcomes such as the price of kerosene in district d at time t. The right hand side variables consist of the following: a dummy variable representing exposure to the rural electrification program, $RGGVY_{dt}$, that is equal to zero for all years prior to RGGVY implementation in the district and is equal to one afterwards; district fixed effects, γ_d, that control for average differences across districts; time fixed effects, λ_t, that control for changes over time common to all districts; and a vector of district characteristics in 2001, $\mathbf{X}_{d2001}$, that includes the statistically significant predictors of RGGVY implementation date from Table 1.1, interacted with a linear time trend.[11]

The coefficient of interest in regression equation (2.1) is β, an estimate of the average effect of RGGVY. In a classic difference-in-difference set-up, this effect is identified under the assumption that if RGGVY did not exist, kerosene market outcomes would have changed similarly across early and late adopters of the program. While Table 1.1 previously showed statistically significant differences between districts prior to the program, these patterns do not necessarily violate the identification assumption if early and late implementing districts have identical counterfactual trends. Nevertheless, to account for the possibility that initially different districts are also more likely to trend differently, I include a vector of pre-RGGVY district-level variables, $\mathbf{X}_{d2001}$, in equation (2.1). I therefore focus on comparing districts that were similar at baseline, with the identification assumption that early implementing

[10] Burlig and Preonas (2016) also examine RGGVY and use regression discontinuity to investigate impacts on economic development. Regression discontinuity is feasible in their setting because the outcomes they study are available at the village level, for instance through the village directories of the Indian census.

[11] The 2001 district characteristics in the vector $\mathbf{X}_{d2001}$ are the natural log of the number of households, proportion of SCs, and individual variables for the proportion of villages with an educational facility, post/telegraph/telephone facility, baking facility, power supply, and population above 300.

districts would have otherwise changed similarly, on average, to late implementing districts *that have similar initial characteristics.*

Whereas equation (2.1) employs a *discrete* treatment variable, my second regression uses a *continuous* treatment to capture variation in the intensity of RGGVY across districts. As explained earlier in Section 1.3, differences in the "dosage" of RGGVY arise from the program component providing free grid connections to all BPL households. I incorporate this dimension in the empirical analysis by estimating the regression

$$y_{dt} = \beta Connections_{dt} + \gamma_d + \lambda_t + \delta \mathbf{X}_{d2001} t + \epsilon_{dt}. \tag{1.2}$$

Here, $Connections_{dt}$ is a continuous variable equal to zero for all years prior to RGGVY implementation in the district. However, in post-implementation periods, $Connections_{dt}$ represents BPL household connections covered by RGGVY in district d at time t, expressed as a proportion of the total number of households in the same district based on the 2001 census.

The regression in equation (2.2) has both important similarities and differences with equation (2.2). For example, equation (2.2) retains the same basic features as before: it includes district fixed effects γ_d, time fixed effects λ_t, and a vector of 2001 district characteristics $\mathbf{X}_{d2001}$ linearly interacted with time. Moreover, identification in equation (2.2) remains analogous to equation (2.1), taking advantage of the variation in the timing of RGGVY across districts and requiring the same identification assumption. Yet in comparison to the binary treatment in the previous regression, the alternative treatment in equation (2.2) makes it possible to understand the effects of different levels of exposure to RGGVY. By formulating the treatment variable to be continuous rather than discrete, equation (2.2) thus allows the changes in kerosene market outcomes to depend not only on the presence of the program itself, but also on the intensity of program implementation.

Finally, my third regression investigates the year-by-year impact of RGGVY using an event study. Specifically, I implement the regression

$$y_{dt} = \sum_{k=-4}^{5} \beta_k D_{dt}^k + \gamma_d + \lambda_t + \delta \mathbf{X}_{d2001} t + \epsilon_{dt} \tag{1.3}$$

where the variables $\beta_k D_{dt}^k$ represent lags (i.e., post-treatment effects) as well as leads (i.e., anticipatory effects). Formally, denoting r_d as the implementation date of RGGVY in district d and $\mathbb{I}[\cdot]$ as the indicator function, I define $D_{dt}^k \equiv \mathbb{I}[t = r_d + k]$. In other words, D_{dt}^k is a dummy variable indicating that in district d and time t, RGGVY was implemented k periods ago for $k > 0$ (or RGGVY will be implemented in k periods for $k < 0$). To facilitate hypothesis tests of program impacts, I exclude the first lead D_{dt}^{-1} in the regression, so the coefficients β_k estimate the effects of RGGVY relative to the year just before implementation. I then present scatter plots of these coefficients in Section 1.6, illustrating the evolution of treatment effects in "event time."

The coefficients β_k from regression equation (1.3) reveal two critical pieces of information. First, they lend themselves to a direct test of the identifying assumption of differences-in-differences. In particular, they allow us to examine whether the coefficients on all leads of the treatment are zero (i.e., $\beta_k = 0 \; \forall \; k < 0$). If the pre-program coefficients β_k are all statistically insignificant and display no trends, these patterns would offer reassurance for the validity of the empirical design. Indeed, as I will discuss in more detail in the results, the event studies in Figure 1.7 are consistent with the identical counterfactual trends assumption since the estimates of leads of β_k for my outcome variables of interest are all very close to zero. Second, the pattern of lagged effects is also of substantive interest: they enable us to examine the dynamic impacts of RGGVY, to understand whether and how the effects of the program change over time.

1.6 Empirical Results

Effects on Household Electricity and Kerosene Use

As a first step to understanding the effects of rural electrification on kerosene markets, I begin by investigating whether RGGVY did in fact lead to meaningful changes in households' energy use. This is a critical exercise particularly in a developing country context like India, where inefficiencies, corruption, and misappropriation of funds are commonplace, resulting in government programs that often fail to reach their intended beneficiaries. Furthermore, the administrative data from RGGVY do not include information on whether the electricity infrastructure and the provision of BPL connections were actually completed. Hence, I seek to establish the "first-stage" effects of RGGVY by estimating equations (1) and (2.2) for two dependent variables both constructed from the NSS data: the percentage of households in the district using electricity as their main source of lighting (henceforth, "electrification rate") and the equivalent measure for kerosene.

The results from these regressions, reported in Table 1.2, point to three important patterns. The first is that implementation of RGGVY is associated with increased use of electricity. As seen in column 1, the RGGVY treatment dummy yields an average of 1.2% point gain in the district electrification rate, representing approximately a 2% increase given that the mean electrification rate across districts was 62% in the year prior to the launch of the program. Moreover, while these findings use self-reported survey data which may be prone to measurement error, they are consistent with the results of Burlig and Preonas (2016) who use nighttime brightness as observed from space. Focusing on the impacts of RGGVY on economic development, the authors compare villages just above and below the program's population eligibility rule, demonstrating that RGGVY led to a substantial increase in nighttime brightness.

The second result we observe from Table 1.2 concerns the impact of RGGVY on electrification rates across different treatment dosages. In particular, districts at the median level of treatment intensity—that is, where 14% of households in the district are provided with free BPL household connections—see an increase in the electrification rate of 1.9% points.[12] Im-

[12]The amount 1.9% points is obtained by calculating $\widehat{\beta}_{BPL} * 0.14 * 100$, where the estimate $\widehat{\beta}_{BPL} = 0.137$ comes from Table 1.2.

portantly, the coefficient in column 2 on the continuous treatment variable, labelled *RGGVY BPL Connections*, is far below one, suggesting that not all BPL households take up the free connections available to them. For example, if 10% of households in the district are provided with free connections, the estimates show that the electrification rate increases by only 1.37% points. Hence, even when provided with free connections, only some BPL households may connect, while the majority remains unconnected and continue to use kerosene instead.

In addition to RGGVY's impacts on electricity use both overall and by varying treatment intensity, Table 1.2 illustrates a third interesting pattern: RGGVY implementation corresponds to a decline in kerosene use. Specifically, districts implementing RGGVY experience an average of 1.4% point decrease in the proportion of households using primarily kerosene for lighting, while at the median treatment intensity, the analogous figure is 2.1% points. Notably, on both binary and continuous treatment variables, the magnitudes of the coefficients across electricity and kerosene outcomes are all very similar. Thus, electricity use appears to displace kerosene use by an equal amount, further supporting the motivating fact that kerosene and electricity are close substitute forms of lighting, as explained earlier in Section 1.2.

Effects on Kerosene Prices

Having established that RGGVY did improve electrification rates among rural households, I now move to the central question of interest: how did rural electrification impact the price of kerosene, a close substitute of electricity for lighting? Since price data are typically rare in developing country settings, I take advantage of the rich household-level data from the NSS Consumer Expenditure Surveys. In particular, the NSS asks households to report their total kerosene expenditure (in Rupees) and the total quantity of kerosene they consumed (in liters) during the last 30 days. By dividing the former by the latter, I obtain unit values—that is, the value per unit consumed—which I subsequently use as a proxy for the kerosene price.

Table 1.3 shows regression results from estimating equations (1) and (2.2) where the outcome variable is the median price per liter of kerosene in the district, measured using unit values from NSS consumption data of rural households. In column (1), I find that RGGVY implementation led to an average increase in kerosene prices of about Rs. 0.81 per liter. This effect is not only statistically significant at the 1% level, but is also economically significant as it corresponds to a 5.4% hike over the mean pre-RGGVY price of Rs. 15. Moreover, in column (2), I observe that for districts at the median level of treatment intensity, the nominal price of kerosene increases by Rs. 0.33 per liter, while for districts at the 90th percentile of treatment dosage (where almost half of all households receive free connections), the price rises by more than Rs. 1 per liter.

Whereas columns (1) and (2) of Table 1.3 investigate effects on *nominal* prices, columns (3) and (4) examine *real* prices. Here, I deflate nominal prices into real terms—expressed in year 2000 Rupees—by using *state-level* consumer price indices (CPI) for rural laborers taken from India's Labour Bureau.[13] Real prices enable me to account for differences in price levels both over time and across space, so that the effects of RGGVY are comparable regardless of when districts were treated. Notably, I find that similar to the results on nominal prices, rural electrification leads to higher real kerosene prices as well. For example, the RGGVY treatment increases real kerosene prices by Rs. 0.44 per liter on average, representing a 3% change compared to the mean real kerosene price in the pre-program years of Rs 13.7.

Identifying these effects of RGGVY on kerosene prices hinges on the assumption that trends in treatment and control districts would have been identical in the absence of the program. In what follows, I evaluate the identification assumption directly using the event study described earlier in regression equation (1.3). Importantly, I restrict the analysis to a *fully balanced sample of districts* both before and after RGGVY. Because districts implementing the program earlier are observed in many more post-treatment years than districts implementing the program later, without a balanced panel, the composition of districts driving the identification of each β_k in equation (1.3) would change for each time period k. A balanced panel is therefore critical to alleviate selection concerns and composition effects when investigating the evolution of treatment impacts.

To achieve a stable composition of districts in event time, I make the following two choices regarding the data. First, I focus on districts treated during the 10th Plan. In so doing, I am able to construct a balanced panel of districts with a time window of four years before and five years after RGGVY implementation. Although this restriction necessarily yields a smaller sample covering only 40% of all districts, the median district treated in the 11th Plan is observed for only two post-treatment periods. Thus, concentrating on 10th Plan-treated districts has the important advantage of a longer event study with which to explore the time path of effects. Second, because the NSS expenditure surveys were not fielded in 2008 and 2010, I linearly interpolate between observation years to obtain missing values. Otherwise, it would not be possible to obtain a balanced panel, as kerosene prices in a district treated in 2006, for example, will be observed at lags 0, 1, and 3, while in a district treated in 2007, prices will be observed at lags 0, 2, 4.

Given the above set-up of the data, Figure 1.7 plots the event study for nominal prices in Panel A and real prices in Panel B. In both panels, each dot denotes the estimate of the coefficient β_k for event time k from equation (1.3), and the vertical lines represent its 95%

[13] The Labor Bureau also provides a price index for agricultural laborers. The results are similar whether the rural laborer CPI or the agricultural labor CPI is used. Furthermore, I exclude fuel and light in the CPI by re-weighting the index as follows. Expenditure items in the CPI are divided in to five groups: (1) food; (2) pan, supari, tobacco, and intoxicants; (3) fuel and light; (4) clothing, bedding, and footwear; (5) miscellaneous items. The Labour Bureau publishes a state-level sub-index for each of the five groups, and their weighted average constitutes the state-level CPI. Denote x_i as the weight for group i, where $i = \{1, \ldots, 5\}$, $i = 3$ corresponds to fuel and light, and $\sum x_i = 100$. I re-calculate the weight for group $i = \{1, 2, 4, 5\}$ as $y_i = x_i/(x_1 + x_2 + x_4 + x_5)$. Then, I use the alternative weights y_i to obtain the weighted average of the sub-indices for groups 1, 2, 4, and 5. Note that the CPI does not cover the following states: Arunachal Pradesh, Chhattisgarh, Jharkhand, Mizoram, Nagaland, Sikkim, and Uttaranchal. For these states, I use the national-level CPI.

confidence interval.[14] The results lend strong support to the validity of the difference-in-difference design. In particular, the pre-treatment trend is flat, and all coefficients on leads are very close to and not statistically different from zero. Hence, there are no systematic differences in kerosene price trends prior to RGGVY implementation. In addition, there is a clear break in the trend once RGGVY is implemented: kerosene prices jump up at event time zero, and prices continue to rise thereafter. Indeed, the estimates suggest that five years after RGGVY, nominal and real kerosene prices increased by Rs. 7 and Rs. 4, respectively, relative to the year just before program introduction.

Altogether, the differences-in-differences and event studies demonstrate that RGGVY led to a significant increase in the price of kerosene whether measured in nominal or real terms. As a consequence, households that continue to rely on kerosene for lighting are potentially made worse off by the rural electrification program. The foregoing analysis is, however, not without caveats because unit values are not prices *per se*. For instance, unit values may be measured with error if respondents inaccurately recall their total expenditure and quantities consumed. Moreover, unit values can be calculated only when households actually consume the product, which may be a challenge if fewer households purchase kerosene as a result of electrification. Section 1.8 discuss these issues in greater detail as well as provides several robustness checks of the preceding results.

Effects on Kerosene Market Size and Aggregate Demand

The *positive* effect of rural electrification on kerosene prices is perhaps a rather unexpected and counterintuitive finding, especially when compared to conventional economics insights. In a traditional supply and demand framework with perfect competition, basic principles from economics would predict a *negative* effect on kerosene prices given that kerosene and electricity are close substitutes for lighting. So why would kerosene prices increase when electricity become available? The rest of the paper delves into this puzzling question. In particular, I propose and examine a potential link between rural electrification and higher kerosene prices: the decline in kerosene market size and aggregate demand.

The intuition behind the market size as a channel through which electricity provision increases kerosene prices is as follows. In a setting with free entry and exit and where kerosene sellers incur fixed costs of operation, prices must equal average costs in equilibrium. At the same time, the availability of electricity causes the pool of kerosene consumers to shrink, given that a subset of households who have previously purchased kerosene for their lighting needs now switch to using electricity. Importantly, this smaller kerosene market size increases equilibrium average cost of production among kerosene retailers, and as a consequence, the equilibrium price of kerosene rises as well.

The earlier findings in Table 1.2, which demonstrated that RGGVY led to a lower proportion of households using kerosene for lighting, already provide preliminary evidence that

[14]Note that by construction, the coefficient β_k for $k = -1$ is zero since the first lead D_{dt}^{-1} is excluded in the event study regression, as described in Section 1.5.

electricity entry is likewise associated with a reduction in the demand for kerosene. Nevertheless, I further investigate effects on demand by estimating regression equations (2.1) and (2.2) for three different measures of kerosene market size and aggregate demand at the district level—that is, the proportion of households in the district purchasing an kerosene, the median quantity of kerosene (in liters) consumed by households in the district, and the total consumption of kerosene per capita (in liters) in the district.

The regression results for the above outcome variables are presented in Table 1.4. Across all measures, the estimates point to a decline in the market size for kerosene as a result of the electrification program; the presence of RGGVY corresponds to a 2% point decrease in the proportion of households who purchase any kerosene (column 1) and a 0.069 liter drop in the median quantity of household kerosene consumption within the district (column 3). With this empirical evidence on the negative effects of electrification on kerosene market size at hand, the next section develops a theoretical framework for understanding how the lower market size and aggregate demand for kerosene translates into higher kerosene prices.

1.7 Theoretical Framework

Using a simple theoretical model of spatial competition, I examine an environment where initially, kerosene retailers are the only purveyors of lighting for rural households, and subsequently, electricity becomes available. I then show that electricity entry may lead to higher prices, and I discuss plausible reasons for this price increase. More generally, I also investigate the conditions under which electricity provision increases or reduces kerosene prices.

Model Set-up

My basic framework builds upon the classic circular spatial model by Salop (1979), which I augment in two ways. First, I consider how the entry of a firm that is not located on the circle (i.e., electricity) affects retailers that are on the circle (i.e., kerosene sellers). Second, whereas consumers in the Salop (1979) model vary only in their location on the circle, I introduce consumer heterogeneity in the cost of adopting electricity. The following subsections describe these different elements of the model in greater detail.

Demand Structure. Consumers of measure M are distributed evenly on a circle with circumference of the same size. For simplicity, I assume that each consumer has unit demand for lighting. All consumers have the same reservation price v, which is assumed to be large enough such that all consumers purchase lighting in equilibrium. Consumers select the lighting supplier—that is, a kerosene retailer or an electricity connection (if available)—that maximizes their utility.

When purchasing kerosene, a consumer incurs transportation cost t for each unit of distance traveled between her location and the kerosene retailer's location on the circle. Therefore, the utility of a consumer located at x who purchases from retailer i is given by

$U = v - t|x - l_i| - p_i$, where $|\cdot|$ is a function representing the arc distance, l_i is the location of retailer i, and p_i is the price of a unit of kerosene from retailer i.

With electricity purchases, on the other hand, I assume consumers are not subject to transportation costs because electricity is provided directly in the consumer's home. However, I assume that the cost of adopting electricity varies across individuals. Each consumer is one of two types: a fraction $\lambda \in (0, 1)$ are of type H ("high") and have adoption cost $\gamma > 0$ when connecting to electricity, while the remaining $1 - \lambda$ are of type L ("low") whose cost of adoption is normalized to 0. Given a price p_e of electricity, H types have utility $U_H = v - p_e - \gamma$ from electricity, while L types have corresponding utility $U_L = v - p_e$.

Supply Structure. Kerosene retailers each have marginal cost of production c and fixed cost F of entering the market. They compete in the following two-stage game. In the first stage, the retailers simultaneously decide whether to enter, and if so, where to locate on the circle. Then, in the second stage, each firm i sets its price p_i and earns profits. To simplify the first stage, I appeal to the *principle of maximal differentiation*: if n firms decide to enter, they will be located equidistant from each other on the unit circle, so that the distance between two sellers is $1/n$.

Equilibrium Concept. I consider the symmetric zero-profit equilibrium (SZPE) as in Salop (1979). The SZPE is defined as a price p and number of kerosene retailers n such that retailers are equally spaced on the circle, p is the Nash price-setting profit maximizing choice, and retailers earn zero profit. I solve for the SZPE under two states of the world: (1) when electricity is *not* available in the market, and (2) when electricity is available at a price p_e.

Results

Benchmark Case: Electricity is unavailable. I first examine the equilibrium when only kerosene sellers are present in the market for lighting. Households in this setting select the kerosene retailer that provides them with the highest utility, after taking into consideration the seller's price and the transportation cost to that seller's location. Meanwhile, each kerosene retailer i competes directly with two other retailers: one located to the left of i and another to the right of i. The market equilibrium is then determined by the demand-side parameters for transportation cost t and market size M, together with the supply-side parameters for marginal cost c and fixed cost F.

Lemma 0 *When electricity is not available, prices $p^* = c + \sqrt{tF/M}$ and number of retailers $n^* = \sqrt{tM/F}$ constitute an equilibrium in the kerosene market.*

Proof Without loss of generality, consider kerosene retailer i located at 0 and retailer j located at $1/n$. Denote retailer i's price as p_i and all other competitors' prices as $\bar{p}$. A consumer located at $\hat{x}$ between i and j is indifferent between the two sellers if both provide her with the same level of utility: $v - t|\hat{x}| - p_i = v - t|\frac{1}{n} - \hat{x}| - \bar{p}$, where $\hat{x} = \frac{\bar{p} - p_i + t/n}{2t}$. Note that all consumers between 0 and $\hat{x}$ purchase from i, and consumers are uniformly distributed on that interval. Moreover, since seller i has two neighbors (i.e., one on each side), its demand function is $D(p_i) = 2\hat{x} = \frac{\bar{p}-p_i+t/n}{t} \cdot M$. Seller i thus solves the profit maximization problem

$$\max_{p_i} (p_i - c) \cdot \left[\frac{\bar{p} - p_i + t/n}{t}\right] \cdot M - F.$$

Taking the FOC with respect to p_i and imposing symmetry $p_i = \bar{p} = p^*$ yields $p^* = c + \frac{t}{n}$. Finally, we obtain the equilibrium number of firms n from the zero profit condition $\Pi(p_i) = (p_i - m) \cdot D(p_i) - F = 0$. From this condition, $n^* = \sqrt{\frac{tM}{F}}$ completes the description of the equilibrium. ∎

Table 1.5, column 1 summarizes the above equilibrium for the case when only kerosene sellers are present in the market. Notably, this simple benchmark case highlights the role of both the market size M and fixed costs F in determining kerosene market outcomes. Comparative statics on the equilibrium price p^* and number of kerosene sellers n^* show that a lower market size and/or higher fixed costs result in fewer firms who charge a higher price. When either the market size falls or fixed costs of entering the market increase, the average costs of production for kerosene sellers rises. As a consequence, sellers charge a higher price in equilibrium to operate in the market.

Now suppose that electricity enters the market and costs p_e, an exogenously set price.[15] Although there are many equilibria in this environment, I do not solve for all possible equilibria. Rather, I focus on the following two cases: (1) type H consumers do not adopt electricity, but *some* type L consumers do; (2) type H consumers do not adopt electricity, but *all* type L consumers do. These two cases demonstrate that electricity provision has ambiguous implications for the kerosene market. In particular, the former shows a *competitive effect* from electricity entry resulting in lower kerosene prices, while the latter illustrates a *market size effect* leading to higher prices of kerosene. The analysis below provides the conditions under which each of these two effects to occur.

Case 1: Some type L and no type H consumers adopt electricity. A necessary condition for this equilibrium to arise is that the price of electricity p_e is *higher* than the price of kerosene p^*. If so, not all L-types would connect to electricity, as those who live close to a kerosene seller receive higher utility from purchasing kerosene instead. At the same time, for H-types, the total cost of buying electricity (i.e., electricity price p_e plus adoption cost γ) must be prohibitive enough such that they all prefer kerosene to electricity, even with the transportation cost. Lemma 1 formalizes these conditions along with the equilibrium outcome in the kerosene market.

[15] I do not explicitly model the interaction between kerosene and electricity providers. This is an appropriate assumption in my empirical context where the government sets the price of electricity regardless of kerosene prices.

Lemma 1 *When electricity is available, the following prices p^* and number of retailers n^* constitute an equilibrium in the kerosene market, given that $p_e + \gamma > \frac{t}{2n^*} + p^* > p_e > p^*$:*

(1) $p^* = c + \sqrt{\frac{tF}{(2-\lambda)M}}$

(2) $n^* = \frac{\lambda t \sqrt{(2-\lambda)M}}{\sqrt{tF}(3-2\lambda) - 2(p_e - c)(1-\lambda)\sqrt{(2-\lambda)M}}$

In this equilibrium, some *type L and* no *type H consumers adopt electricity.*

Proof See Appendix A.1.

Table 1.5, column 2 provides a summary of the above equilibrium in Lemma 1 where electricity entered the market and some L-type consumers adopt electricity. Compared to Lemma 0 where electricity is not available, Lemma 1 shows that availability of electricity leads to a lower price of kerosene. This result emerges even if electricity suppliers are not directly interacting with kerosene retailers in this model.[16] Hence, electricity provision—in and of itself—brings about a *competition effect* that induces kerosene prices to fall: given that some L-type consumers are on the margin between purchasing kerosene and electricity, kerosene retailers would like to set a low price to attract such consumers back into the kerosene market.

In addition to increasing kerosene prices, electricity entry also reduces the equilibrium number of kerosene sellers n^*. To see this, note that n^* is lower than $\sqrt{tF/M}$ (i.e., the equilibrium number of sellers in the pre-electricity case) for the following two reasons: (1) $n^* = 0$ when $\lambda = 1$, and $n^* = \sqrt{tF/M}$ when $\lambda = 1$, and (2) n^* is increasing[17] in $\lambda \in (0, 1)$. That electricity entry results in lower n^* again underscores the importance of market size in determining kerosene market outcomes. Since some consumers switch to electricity and no longer use kerosene, electricity provision causes the pool of kerosene buyers to shrink. As a result, fewer retailers operate in the kerosene market.

Case 2: All type L and no type H consumers adopt electricity. For this equilibrium to occur, the price of electricity p_e must be *lower* than the price of kerosene p^*. If this is the case, then all type L consumers—even those living very close to kerosene retailer—would receive higher utility from connecting to electricity. In addition, the electricity price p_e plus the adoption cost γ for H-types must be large enough so that all type H consumers prefer kerosene, including those who live farthest away from the kerosene store. This latter requirement is identical to Case 1 and is necessary so that no H-type consumers adopt electricity. Given these conditions, Lemma 2 derives the equilibrium prices and number of sellers in the kerosene market.

[16] As will be described in the empirical section, this assumption is appropriate in the context of India where the government sets electricity prices regardless of kerosene prices.

[17] Note that n^* is increasing in $\lambda \in (0, 1)$ since the numerator of n^* is increasing with λ over the interval $(0, 1)$, while first derivative of the denominator of n^* is negative, so that it decreasing with λ. Therefore, n^* overall is increasing with λ.

Lemma 2 *When electricity is available, the following prices p^* and number of retailers n^* constitute an equilibrium in the kerosene market, given that $p_e + \gamma > \frac{t}{n^*} + p^*$ and $p_e < c + \sqrt{\frac{tF}{(2-\lambda)M}}$:*

(1) $p^* = c + \sqrt{\frac{tF}{\lambda M}}$

(2) $n^* = \sqrt{\frac{t\lambda M}{F}}$

In this equilibrium, all *type L and* no *type H consumers adopt electricity.*

Proof See Appendix A.2.

Table 1.5, column 3 provides an overview of the equilibrium price and number of kerosene sellers in Lemma 2. This lemma demonstrates that compared to the pre-electricity equilibrium indicated in Lemma 0, the provision of electricity results in higher kerosene prices. In particular, kerosene prices increase because of a *market size effect*: the negative demand shock from electricity entry causes the market size for kerosene to fall, and when electricity is priced at a very low cost, kerosene retailers find it difficult to compete with this already low price. As a result, kerosene sellers are unable to expand the pool of potential kerosene buyers by enticing electricity users back into the kerosene market.

This *market size effect* increases kerosene prices and decreases the number of kerosene retailers due to the higher equilibrium average cost of production retailers incur when selling kerosene. In the equilibrium in Lemma 2, all L-type consumer switch to electricity because of its relatively low price, and only H-types—for whom electricity adoption costs are relatively high—are the only consumers left buying kerosene. With a smaller set of kerosene buyers, each kerosene retailer sells a lower quantity of products. Consequently, in equilibrium, the market can support fewer kerosene retailers, the average cost of selling kerosene rises, and hence, the price of kerosene also increases.

Summary and Discussion. The preceding analysis reveals that Case 1 and Case 2 result in vastly different outcomes for kerosene consumers. Although the equilibrium number of kerosene *sellers* decreases in both cases, the equilibrium kerosene *price* falls in Case 1 while the opposite is true in Case 2. Households who do not adopt electricity and remain kerosene users are therefore better off in terms of household fuel and lighting expenditure in the first case. In contrast, in the second case, the price hike hurts those who continue to rely on kerosene, particularly type H households who, given the high costs of adopting electricity, forgo an electricity connection despite its availability.

To improve the welfare of such households, the theoretical results likewise point to a number of potential policy tools: first, decreasing the costs of electricity adoption γ, for example through subsidies not only for connection fees but also for light bulbs other equipment that allow households to take up electricity for lighting; second, easing kerosene prices by reducing kerosene sellers' fixed costs of operation F or marginal cost m; third, lowering the transportation cost t of traveling to a kerosene retailer; and finally, providing cash transfers and/or kerosene subsidies to compensate kerosene users for the higher prices. These programs, when implemented in parallel with rural electrification, may aid in combating the possible adverse effects of electricity provision among electricity non-adopters.

1.8 Robustness Checks

In this section, I provide several robustness checks on the effects of RGGVY on kerosene prices as well as on the overall validity of the empirical research design.

Alternative measure of prices. Using unit values as a proxy for price raises two important issues (e.g., Deaton 1998, 1990, 1997). First, unit values can be prone to measurement error. In particular, the measurement error from respondents' recall bias may be compounded when dividing household expenditures by quantity of consumption. Second, unit values can only be calculated when the respondent actually consumes the product. If households in treated districts no longer consume kerosene—and indeed, the estimates in Table 1.4 show that fewer households purchase kerosene as a result of electrification—then the unit value for kerosene will no longer be observed. Hence, unit values from households in electrified districts are more likely to be missing in post-treatment periods.

I address the above issues by taking advantage of an alternative measure of prices from the *Rural Price Collection* (RPC) Survey implemented by the NSS. This survey collects rural retail prices every month for a fixed basket of goods, including kerosene, from 603 markets spread over 26 states. Although the RPC covers only half of all districts participating in the RGGVY program, prices from these data form the basis for the CPI for Agricultural and Rural Laborers,[18] and therefore, are potentially more accurate than NSS unit values. Figure 1.8 shows a scatter plot of kerosene prices as measured by the RPC data and by NSS unit values. The correlation between these two measures is 0.68, but as can be seen in the figure, they do not always correspond with each other.

In Table 1.6, I present regression estimates of the impact of RGGVY on kerosene prices using RPC data. The results in this table confirm my previous findings in Table 1.3—which used NSS unit values as a proxy for price—that electricity provision is indeed associated with higher kerosene prices. For example, Table 1.6 indicates that the presence of electricity corresponds to an increase in the nominal and real kerosene prices of Rs. 1.30 and Rs. 0.92, respectively. Interestingly, the magnitudes of the coefficients when using the RPC data are between 40 to 100% larger than those using NSS unit values. These results thus suggest that the measurement error from NSS unit values may be causing attenuation bias.

[18] The CPI for Agricultural and Rural Laborers is compiled and published by the Indian government's Ministry of Labor.

Placebo Test. As an additional test for the credibility of the differences-in-differences estimates, I explore the effects of the electrification treatment among *urban* households. Because RGGVY covered only rural villages, the program should have had no impact on the electricity use of households in urban areas. Otherwise, if the program had non-zero effects for urban households, then the estimated impacts on *rural* households may be biased. Table 1.7 reports regression results where the outcome variable is the proportion of urban households in the district who use electricity. The findings from this placebo test show that the coefficients on both the discrete and continuous RGGVY treatments are very close to zero. Additionally, they are not statistically significant at conventional levels, lending further support to the validity of the empirical approach.

1.9 Conclusion

In contrast to much of the existing literature which finds *positive* effects of electrification on socio-economic outcomes, this paper demonstrates that electricity provision may have *negative* effects particularly among household who do not take up electricity. Using India's national rural electrification program as a quasi-experiment, I show that electrification increases the price of kerosene, an alternative source of lighting, by 5 to 10%. In addition, I provide evidence that the decline in the kerosene market size may be a potential channel through which electricity provision leads to higher kerosene prices. Overall, the results indicate that households who do not avail of an electricity connection and remain kerosene consumers become worse off, as they end up paying a higher price to fulfill their lighting needs.

Since the non-adopters of electricity also tend to be poor households who also devote more of their meager income on fuel and lighting than their rich counterparts, it is critical to understand how electricity provision impacts those who continue to rely on kerosene or other sources of lighting. With the widespread growth of electrification programs around the world, the results of this study suggests that it may be necessary to implement parallel interventions to mitigate the potential adverse effects of electricity provision on those who do not take up an electricity connection. These insights may aid policy makers, governments, and non-government organizations to design electrification programs that deliver positive benefits for all households, including electricity adopters and non-adopters alike.

Table 1.1: Determinants of RGGVY Implementation Date

	(1)	(2)
Ln(number of households in the district)	–2.103*	–0.427
	(1.088)	(1.366)
Proportion of the district population that is:		
Scheduled Caste (SC)	–48.783***	–29.532*
	(12.157)	(15.189)
Scheduled Tribe (ST)	3.236	–3.875
	(4.536)	(4.528)
Proportion of villages in the district with:		
Educational facility	35.158***	21.343**
	(9.426)	(10.629)
Post, telegraph, and telephone facility	14.307*	18.262**
	(7.417)	(7.577)
Bus services	–9.290	2.606
	(6.181)	(9.296)
Banking facility	46.052***	52.183***
	(7.689)	(13.046)
Credit society	4.538	0.603
	(5.771)	(5.967)
Approach road/path	–0.510	–2.352
	(4.573)	(6.522)
Access to newspapers and magazines	–6.748	–0.489
	(5.117)	(6.300)
Power supply	17.200***	10.501
	(4.818)	(7.477)
Population above 300	–30.529***	–31.803***
	(7.277)	(11.998)
State FEs	No	Yes
R-squared	0.279	0.515
N	554	554

Notes: The dependent variable is the month and year in which the district began implementing the RGGVY program, normalized to 1 in March 2005. Hence, April 2005 takes on the value 2, May 2005 takes on the value 3, March 2006 takes on the value 13, and so on. District characteristics come from the 2001 Census village directory (excluding uninhabited villages) and have been aggregated to the district level. Data on RGGVY implementation dates come from program administrative records. *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table 1.2: Effects of RGGVY on Rural Households' Propensity to use Electricity or Kerosene as a Main Source of Lighting

	Electricity		Kerosene	
	(1)	(2)	(3)	(4)
RGGVY Dummy	0.012*		–0.014**	
	(0.007)		(0.007)	
RGGVY BPL Connections		0.137***		–0.153***
		(0.029)		(0.029)
District FEs	Yes	Yes	Yes	Yes
NSS Round FEs	Yes	Yes	Yes	Yes
2001 District Vars × Linear Time	Yes	Yes	Yes	Yes
Adj. R-squared	0.828	0.829	0.831	0.832
N	5399	5399	5399	5399

Notes: The dependent variable is the proportion of households in the district using electricity (columns 1-2) or kerosene (columns 3-4) as a main source of lighting. RGGVY Dummy is a binary variable equal to zero for all years prior to RGGVY implementation in the district and is one afterwards. RGGVY BPL Connections is a variable equal to zero for all years prior to RGGVY implementation and in post-implementation periods represents BPL household connections covered by RGGVY in the district, expressed as a proportion of the total number of households in the district based on the 2001 Census. District characteristics for 2001 include ln(number of households in the district), proportion of SCs in the district population, and individual variables for the proportion of villages in the district with an educational facility, post/telegraph/telephone facility, banking facility, power supply, and population above 300. Data come from NSS Consumer Expenditure Survey, RGGVY administrative records, and the 2001 Census. Standard errors are clustered at the district level. *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table 1.3: Effects of RGGVY on Kerosene Prices Levels

	Nominal Kerosene Price		Real Kerosene Price	
	(1)	(2)	(3)	(4)
RGGVY Dummy	0.809***		0.436**	
	(0.272)		(0.207)	
RGGVY BPL Connections		2.358*		1.282*
		(1.258)		(0.711)
District FEs	Yes	Yes	Yes	Yes
NSS Round FEs	Yes	Yes	Yes	Yes
2001 District Vars × Linear Time	Yes	Yes	Yes	Yes
Adj. R-squared	0.672	0.671	0.491	0.491
N	5122	5122	5122	5122

Notes: The dependent variable is the price of non-subsidized kerosene per liter both in nominal (columns 1-2) and real (columns 3-4) terms. Data on kerosene prices were constructed from the NSS Consumer Expenditure Survey by taking the median unit value for each district and NSS round. RGGVY Dummy is a binary variable equal to zero for all years prior to RGGVY implementation in the district and is one afterwards. RGGVY BPL Connections is a variable equal to zero for all years prior to RGGVY implementation and in post-implementation periods represents BPL household connections covered by RGGVY in the district, expressed as a proportion of the total number of households in the district based on the 2001 Census. District characteristics for 2001 include ln(number of households in the district), proportion of SCs in the district population, and individual variables for the proportion of villages in the district with an educational facility, post/telegraph/telephone facility, banking facility, power supply, and population above 300. Standard errors are clustered at the district level. *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table 1.4: Effects of RGGVY on Market Size and Aggregate Demand for Kerosene

	Proportion of HHs Purchasing Any		Median Quantity Consumed by HHs		Total Consumption per Capita	
	(1)	(2)	(3)	(4)	(5)	(6)
RGGVY Dummy	–0.020*		–0.069**		–0.002	
	(0.010)		(0.033)		(0.009)	
RGGVY BPL Connections		–0.120***		–0.156		–0.071*
		(0.046)		(0.141)		(0.038)
District FEs	Yes	Yes	Yes	Yes	Yes	Yes
NSS Round FEs	Yes	Yes	Yes	Yes	Yes	Yes
2001 District Vars × Linear Time	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R-squared	0.562	0.563	0.401	0.401	0.205	0.206
N	5399	5399	5399	5399	5399	5399

Notes: The dependent variables are the proportion of househlds in the district purchasing any non-subsidized kerosene (columns 1-2), the median quantity of non-subsidized kerosene consumed by households in the district (columns 3-4), and total consumption of kerosene per capita in the district (columns 5-6). All dependent variables were constructed from NSS Consumer Expenditure Surveys. RGGVY Dummy is a binary variable equal to zero for all years prior to RGGVY implementation in the district and is one afterwards. RGGVY BPL Connections is a variable equal to zero for all years prior to RGGVY implementation and in post-implementation periods represents BPL household connections covered by RGGVY in the district, expressed as a proportion of the total number of households in the district based on the 2001 Census. District characteristics for 2001 include ln(number of households in the district), proportion of SCs in the district population, and individual variables for the proportion of villages in the district with an educational facility, post/telegraph/telephone facility, banking facility, power supply, and population above 300. Standard errors are clustered at the district level. *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table 1.5: Model Results

Kerosene Market Equilibrium	Lemma 0	Lemma 1	Lemma 2
Electricity Available	No	Yes	Yes
Electricity Adopters	None	Some type L, No type H	All type L, No type H
Price, p^*	$c+\sqrt{\frac{tF}{M}}$	$c+\sqrt{\frac{tF}{(2-\lambda)M}}$	$c+\sqrt{\frac{tF}{\lambda M}}$
Number of sellers, n^*	$\sqrt{\frac{tM}{F}}$	$\frac{\lambda t\sqrt{(2-\lambda)M}}{\sqrt{tF}(3-2\lambda)-2(p_e-c)(1-\lambda)\sqrt{(2-\lambda)M}}$	$\sqrt{\frac{t\lambda M}{F}}$
Market share	M	$M\left[(1-\lambda)+2n^*\cdot\frac{p_e-p^*}{t}\lambda\right]$	$(1-\lambda)M$

Notes: This table summarizes the results from the theoretical model. Each of the three columns considers the kerosene market equilibrium for the following scenarios: in column 1, electricity is not available; in column 2, electricity is available, some L types adopt electricity, and no H types adopt electricity; in column 3, electricity is available, all L types adopt electricity, and no H types adopt electricity.

Table 1.6: Effects of RGGVY on Kerosene Prices using RPC Data

	Nominal Kerosene Price		Real Kerosene Price	
	(1)	(2)	(3)	(4)
RGGVY Dummy	1.295***		0.919***	
	(0.363)		(0.336)	
RGGVY BPL Connections		4.962***		1.802**
		(1.664)		(0.872)
Market FEs	Yes	Yes	Yes	Yes
Month FEs	Yes	Yes	Yes	Yes
2001 District Vars × Linear Time	Yes	Yes	Yes	Yes
Adj. R-squared	0.834	0.835	0.799	0.798
N	27361	27361	27361	27361

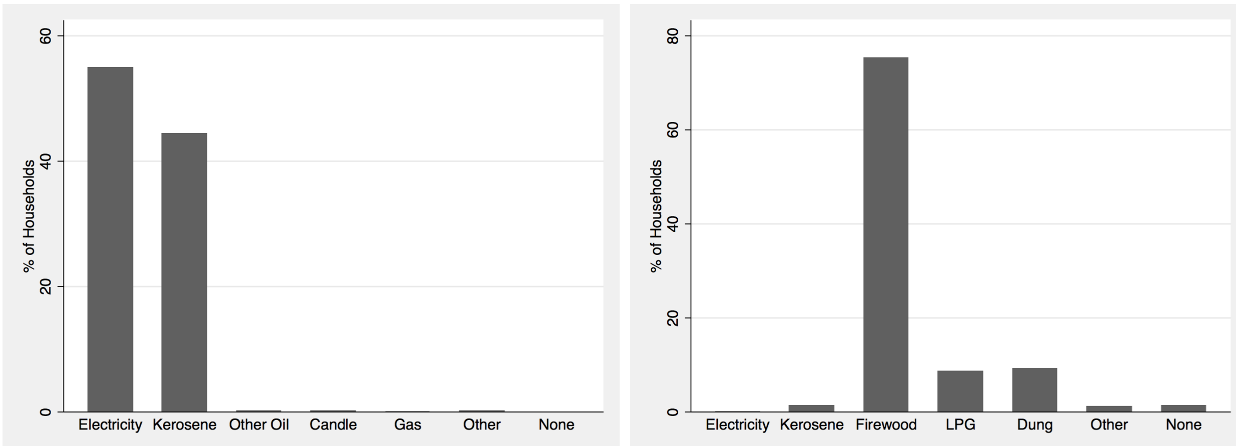
Notes: The dependent variable is the price of non-subsidized kerosene per liter both in nominal (columns 1-2) and real (columns 3-4) terms. Data on kerosene prices are from the NSS Rural Price Collection Data. RGGVY Dummy is a binary variable equal to zero for all years prior to RGGVY implementation in the district and is one afterwards. RGGVY BPL Connections is a variable equal to zero for all years prior to RGGVY implementation and in post-implementation periods represents BPL household connections covered by RGGVY in the district, expressed as a proportion of the total number of households in the district based on the 2001 Census. District characteristics for 2001 include ln(number of households in the district), proportion of SCs in the district population, and individual variables for the proportion of villages in the district with an educational facility, post/telegraph/telephone facility, banking facility, power supply, and population above 300. Standard errors are clustered at the district level. *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table 1.7: Effects of RGGVY on Urban Households Electricity Use

	(1)	(2)
RGGVY Dummy	0.007	
	(0.005)	
RGGVY BPL Connections		0.032
		(0.022)
District FEs	Yes	Yes
NSS Round FEs	Yes	Yes
2001 District Vars × Linear Time	Yes	Yes
Adj. R-squared	0.458	0.458
N	5067	5067

Notes: The dependent variable is the proportion of urban households in the district using electricity. RGGVY Dummy is a binary variable equal to zero for all years prior to RGGVY implementation in the district and is one afterwards. RGGVY BPL Connections is a variable equal to zero for all years prior to RGGVY implementation and in post-implementation periods represents BPL household connections covered by RGGVY in the district, expressed as a proportion of the total number of households in the district based on the 2001 Census. District characteristics for 2001 include ln(number of households in the district), proportion of SCs in the district population, and individual variables for the proportion of villages in the district with an educational facility, post/telegraph/telephone facility, banking facility, power supply, and population above 300. Data come from NSS Consumer Expenditure Survey, RGGVY administrative records, and the 2001 Census. Standard errors are clustered at the district level. *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

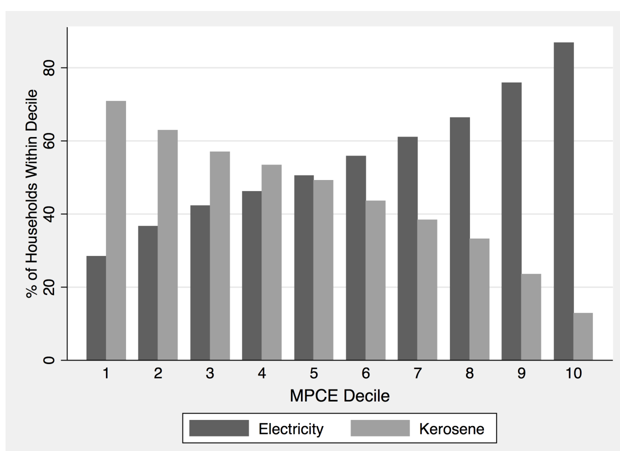
Figure 1.1: Main Lighting and Cooking Energy Sources for Rural Households



(a) Lighting (b) Cooking

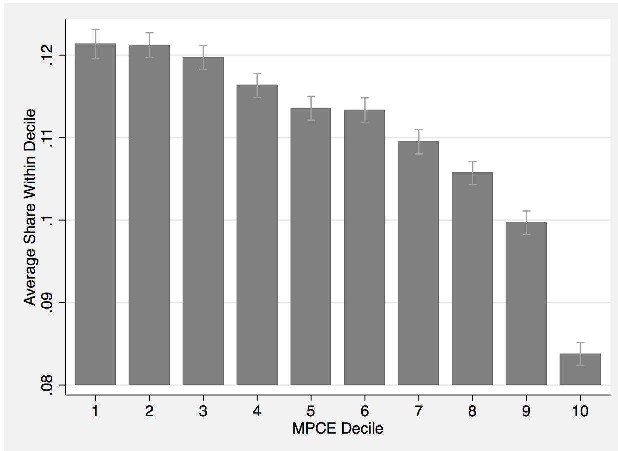
Notes: Each bar represents the percentage of households that use the given item as a primary source of energy for lighting (Panel A) and cooking (Panel B). The data come from Round 61 (2004-2005) of the NSS Consumer Expenditure Survey, and the sample consists of rural households. NSS sampling weights are used.

Figure 1.2: Kerosene and Electricity Use of Rural Households by Household Expenditure Deciles



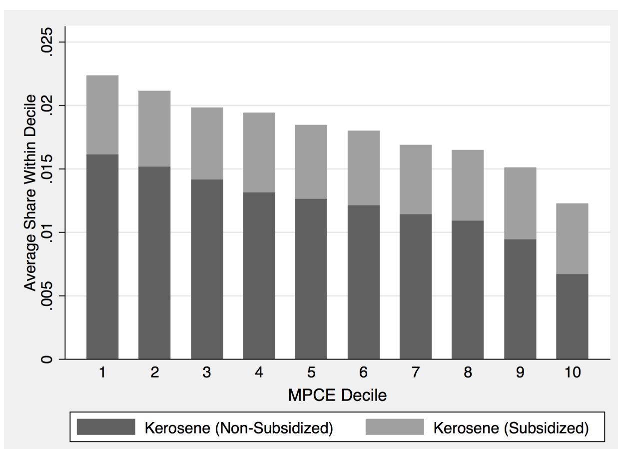
Notes: This bar chart plots the percentage of households that use kerosene (light gray bars) or electricity (dark gray bars) as a primary source of lighting against deciles of household Monthly Per Capita Expenditure (MPCE). The data come from Round 61 (2004-2005) of the NSS Consumer Expenditure Survey, and the sample consists of rural households. NSS sampling weights are used.

Figure 1.3: Fuel and Lighting Share of Total Household Expenditure



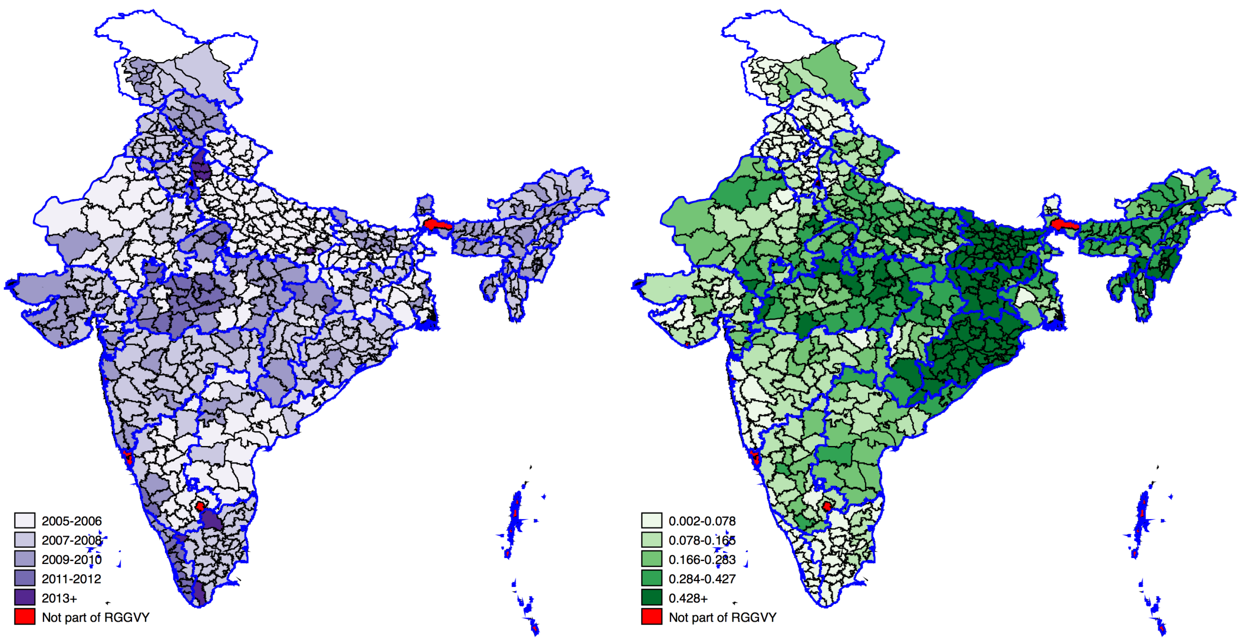
Notes: This bar chart shows the average share of fuel and lighting in total household expenditure for each decile of household Monthly Per Capita Expenditure (MPCE). The vertical lines at the top of each bar represent the 95% confidence interval. The data come from Round 61 (2004-2005) of the NSS Consumer Expenditure Survey, and the sample consists of rural households. NSS sampling weights are used. Fuel and lighting expenditure in the survey consists of the following items: coke, firewood and chips, electricity, dung cake, subsidized kerosene, non-subsidized kerosene, matches, coal, LPG, charcoal, candle, gobar gas, and other fuels.

Figure 1.4: Kerosene Share of Total Household Expenditure



Notes: This bar chart shows the average share of kerosene in total household expenditure for each decile of household Monthly Per Capita Expenditure (MPCE). Each bar is broken down into two components: subsidized kerosene (light gray bars) and non-subsidized kerosene (dark gray bars). The data come from Round 61 (2004-2005) of the NSS Consumer Expenditure Survey, and the sample consists of rural households. NSS sampling weights are used.

Figure 1.5: RGGVY Implementation Across Districts

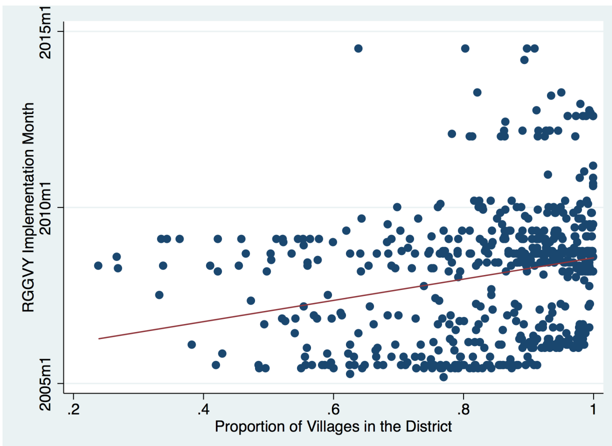


(a) Timing of Implementation

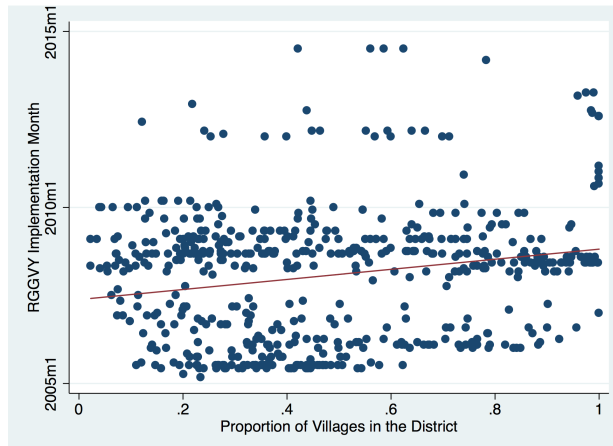
(b) BPL Connections Covered

Notes: In both panels, blue and black lines correspond to 2001 state and district borders, respectively. Red-shaded districts are those which are not part of RGGVY (e.g., union territories, urban areas). Panel A shows implementation of RGGVY across districts over time, with a darker blue shade denoting a later implementation date. The timing of implementation is defined by the date in which RGGVY program funds were first disbursed for the district. Panel B shows BPL connections covered by RGGVY, expressed as a proportion of the the total number of households in the district based on the 2001 Census. The darker green shades indicate a higher proportion of BPL connections covered.

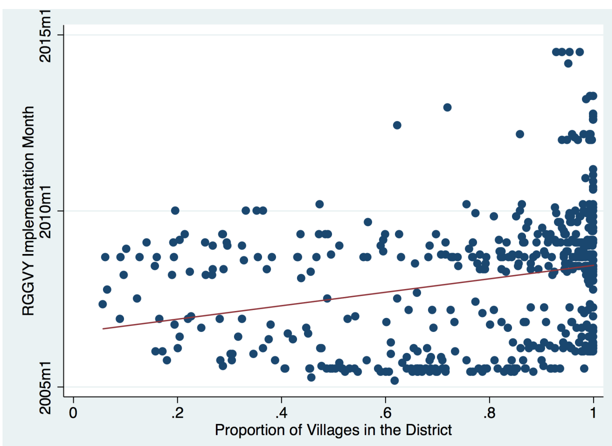
Figure 1.6: Baseline District Characteristics and RGGVY Implementation Date



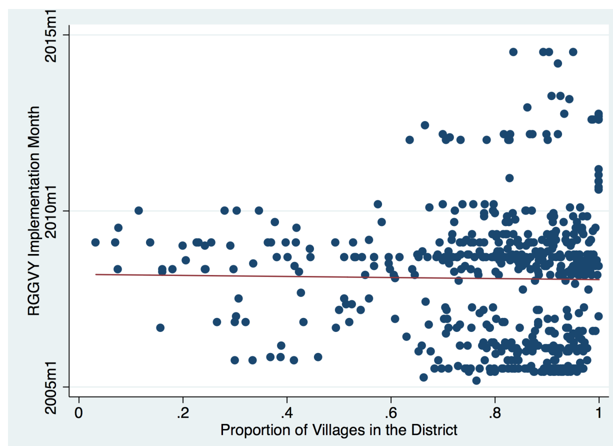
(a) Educational Facility



(b) Post, Telegraph, and Telephone Facility



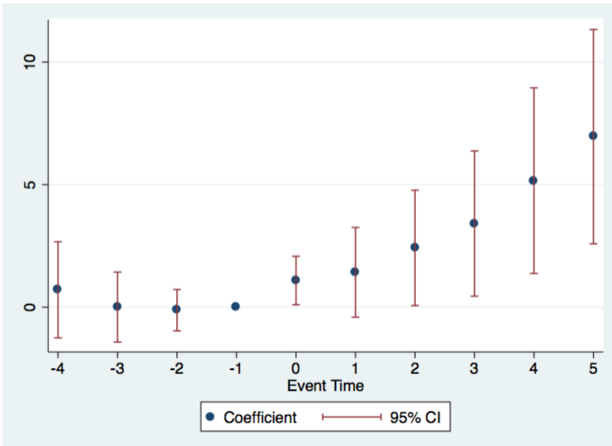
(c) Power Supply



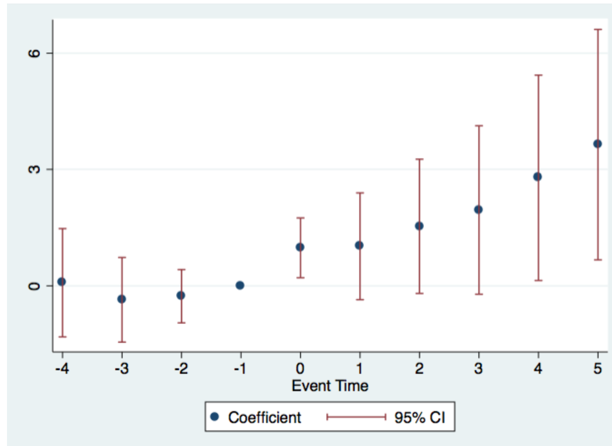
(d) Population above 300

Notes: Each subfigure provides a scatter plot of baseline district characteristics (x-axis) against the month and year of RGGVY implementation in the district (y-axis). The figures also contain the best-fit line. Baseline district characteristics come from the 2001 Census village directory (excluding uninhabited villages) and have been aggregated to the district level. Data on RGGVY implementation dates come from program administrative records.

Figure 1.7: Effects of RGGVY on Kerosene Price Levels



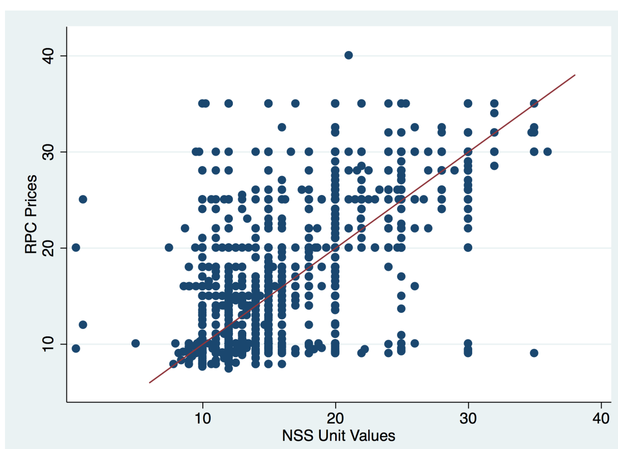
(a) Nominal Prices



(b) Real Prices

Notes: This figure shows event study estimates of RGGVY over the period 2001-2012. In Panel A, the outcome variable is the nominal price of non-subsidized kerosene while in Panel B, it is the real price in year 2000 Rupees. Nominal prices are constructed from the NSS Consumer Expenditure Survey by taking the median unit value reported by rural households in each district-year. Real prices are obtained by deflating nominal prices using state-level CPI for rural laborers, compiled by India's Labor Bureau.

Figure 1.8: Comparison of Kerosene Prices using NSS Unit Values and RPC Data



Notes: This scatter plot compares kerosene prices as measured using NSS unit values (x-axis) and RPC data (y-axis). Prices are at the district-year-level. The NSS unit values are constructed by taking the median unit value reported by rural households in each district-year. The RPC data were collapsed from monthly values to district-year values by taking the median price for each year. The red line on the figure represents the 45-degree line, so that a point is on the line if both NSS and RPC indicate the same price. The correlation between the two measures is 0.68.

Chapter 2

The ABCs of Financial Education: Experimental Evidence on Attitudes, Behavior, and Cognitive Biases

with Shawn Cole, Jeremy Shapiro, and Bilal Zia

2.1 Introduction

In the modern economic landscape, financial literacy—the ability to make informed decisions regarding money—plays a critical role in ensuring both the well-being of households and the stability of the financial system (Lusardi and Mitchell 2014; Lusardi 2015). As financial services such as microfinance and mobile money expand in many parts of the world, so too do concerns, fueled by the recent financial crisis, that many consumers lack knowledge to judiciously utilize the new financial products at their disposal. Consequently, numerous private institutions, non-profit organizations, and governments have responded with implementing financial education programs.[1] Yet the empirical evidence on the efficacy of such programs provide only mixed results, and little is known about which aspects of financial education initiatives successfully (or unsuccessfully) enhance financial behavior and financial outcomes.[2]

[1] The United States adopted a President's Advisory Council on Financial Literacy in 2008 to help promote financial education at all levels of the economy; the UK government mandated compulsory financial education in schools from 2012 onwards; the Indonesian government declared 2008 as the year of financial education; the Reserve Bank of India launched a series of financial literacy and counseling centers across the country in 2007; Brazil and many other developing countries have incorporated national strategies for improving financial education; and private and multilateral agencies such as Citibank and the World Bank have multi-million dollar programs on financial education throughout the developed and developing world.

[2] For example, Duflo and Saez (2003) find that a benefits information session improved retirement savings contributions at a US university, Cole, Sampson, and Zia (2011) show that a financial education program in Indonesia had no impact on bank savings, and Chong, Karlan, and Valdivia (2010) discontinued their study of a video and radio- based financial education course due to logistical challenges and low take up. See

This paper investigates two related questions that have so far been overlooked in the growing literature on financial education. First, what barriers prevent individuals participating in financial education programs from translating financial knowledge into action? And second, what mechanisms are most effective for delivering financial education interventions that meaningfully improve financial outcomes? We shed light on these questions using a randomized evaluation in India, with a large study sample of over 1,300 low-income individuals from a major metropolitan area. Two-thirds of this sample was randomly selected to receive a five-week, high-quality, video-based financial education program which covered budgeting, savings, credit, and insurance. The remaining one-third of the sample received health training with a similar video-based and logistical format, to control for Hawthorne effects.

We employ a rich research design by combining financial education with three additional treatments, all randomly assigned at the individual level, to examine different frictions that can stymie the link between financial education and financial outcomes. In our first treatment, we offer cash incentives to half of the sample for correct answers to a financial knowledge test (henceforth, "pay for performance"). While a number of studies in the education literature have demonstrated that financial incentives improve student achievement (e.g., Angrist and Lavy 2009), our paper is the first to test such incentives in the context of financial education. Theoretically, cash rewards may be necessary to foster the success of financial literacy initiatives if participants have insufficient motivation, have high discount rates, or have deeply ingrained and "sticky" financial habits, all of which may result in sub-optimal effort on the part of participants to learn from financial training. Indeed, existing studies have found little interest among adults in joining a financial education workshop, as well as poor attendance during the program itself (Bruhn, Ibarra, and McKenzie 2014).

For our second treatment, we encourage half of participants receiving financial education to set short-term, achievable, but non-compulsory financial goals, with target dates made visible on a calendar provided by the study (hereinafter, "goal setting"). This intervention allows us to assess the role of self-imposed non-binding goals in attenuating self-control problems, in contrast to standard economic models wherein only binding goals such as pre-commitment or externally enforced contracts can affect motivation and behavior. A large body of literature in experimental psychology dating back to the 1960s consistently confirms the effectiveness of goals for behavior change (Locke and Latham 2002), encompassing a wide variety of fields including worker productivity (e.g., Goerg and Kube 2012), household energy conservation (e.g., Harding and Hsiaw 2014), and health management (e.g., Shilts, Horowitz, and Townsend 2004).[3] Nevertheless, non-binding goal setting remains relatively unexplored

Fernandes, Lynch, and Netemeyer (2014), Miller et al. (2015), Hastings, Madrian, and Skimmyhorn (2013), and Xu and Zia (2012) for literature reviews.

[3]As discussed in Locke and Latham (2002), goals affect performance through the following four mechanisms: (1) providing direction, (2) invoking energy, (3) affecting persistence, and (4) leading to the discovery, and/or use of relevant knowledge and strategies. Goal setting has also been shown to have positive effects on financial decisions such as savings, spending, and debt repayments (e.g. Agarwal et al. 2014; Bartels and Sussman 2015; Salisbury 2014; Soman and Zhao 2011; Ülkümen and Cheema 2011).

in financial education, particularly for addressing psychological constraints that are often more pronounced for the poor and the unbanked (Mullainathan and Shafir 2009, 2013).

Whereas the previous two treatments consider behavioral barriers that financial education participants may face—namely, sub-optimal effort and self-control—our third intervention focuses on structural factors that may hinder the efficacy of financial education. Specifically, we deliver additional financial counseling services involving intensive one-on-one instruction and individualized advice, to half of participants receiving financial education. In the current debate on financial education initiatives, opponents such as Willis (2011, p.431) have argued that because of the heterogeneity of households' circumstances and needs, effective financial education needs to be structured "in a one-on-one setting, with content personalized for each consumer." Even so, rigorous empirical evidence on the impact of financial counseling especially in developing countries remains scarce.[4] Our study therefore contributes to the ongoing debate and informs policies on designing financial education initiatives, as we are able to experimentally evaluate the merits of augmenting a one-size-fits-all financial education program with an individually-tailored counseling approach.

With the explosive growth of financial education initiatives around the world over the last decade, our results are both revealing and optimistic. In a financial knowledge test administered shortly after the five-week program, we find that those who received financial education but not pay for performance achieved 10 percent higher test scores, for questions measuring awareness of and attitudes towards financial products, relative to participants who received neither financial education nor pay for performance. However, the marginal impact of pay for performance is economically and statistically insignificant, and we precisely estimate this null effect. Even more strikingly, the same results still hold 6 to 12 months after the conclusion of the program. Hence, pay for performance led to no improvement in participants' financial knowledge, either in the short- or long-run.

Although we do not find evidence in support of pay for performance, we do identify substantial effects of goal setting and counseling on participants' financial outcomes. In particular, our results show that while financial education alone did not bring about changes in financial behavior, combining goal setting with financial education encouraged relatively simple follow-up actions, such as attempting to write a budget, starting savings, and avoiding borrowing for unforeseen expenses. For example, in comparison to those who received only financial education, participants who additionally received the goal setting treatment were 6 percentage points more likely to join an informal community savings group, corresponding to a 78 percent increase over the control group. This large, statistically significant effect is quite surprising given that the goals were non-binding. Furthermore, it provides a testament to the potential of self- chosen, non-binding goals in mitigating self-control problems in the context of financial planning.

Our results likewise indicate that financial counseling services enable the poor to undertake costlier or more difficult activities to better manage their finances, including regularly writing a budget and opening a formal bank savings account. For instance, participants

[4]Collins and O'Rourke (2010) provide a literature review on financial counseling services.

who also received financial counseling were 13 percentage points (or 45 percent) more likely to open a formal bank savings account relative to the control group. This impact was also significantly higher than for participants who received goal setting. Importantly, this result suggests that an intensive, one-on- one medium for financial education is critical for complex economic decisions and financial products. Given that today's marketplace confronts consumers with evermore sophisticated financial instruments, our findings underscore counseling as a potent mechanism to empower individuals amidst a complicated financial environment.

Taken together, our study demonstrates that participants of financial education programs face two broad sets of constraints that prevent them from bridging the gap between financial knowledge and taking action. One set of barriers is internal to the individual, as it relates to their own behavior, such as lack of self-control, in carrying out their financial goals. At the same time, they encounter external impediments as well, particularly those imposed by a sub-optimally structured financial education program that fails to address their unique needs. Our study reveals that overcoming both sets of constraints simultaneously may be necessary to enhance the effectiveness of financial education initiatives. Indeed, we find that the subsample receiving a very high intensity treatment involving all three interventions of financial education, goal setting, and financial counseling exhibited larger positive changes on all outcomes we examined, compared to those who received only financial education. These findings also offer an explanation for why the financial education literature thus far finds only mixed evidence of impact: the programs studied are fairly heterogeneous with wide variation in topics covered, training emphasis, and medium of instruction, with many failing to concurrently address both internal and external constraints.

More generally, our results show that on its own, financial education is not a panacea for improving the financial well-being of low-income households in developing countries. Nonetheless, we do not view this as evidence to warrant broad pessimism about financial education programs, but rather, we highlight financial education as one strategy in the policy toolbox. We find that financial education fosters participants' knowledge and attitudes towards financial products, yet in and of itself, falls short in promoting improved financial behavior and financial outcomes. In contrast, complementing financial education with inexpensive but personalized add-ons, such as goal setting and counseling, allows consumers to successfully apply their knowledge to financial decision making. We believe these insights can aid financial education policy makers, stakeholders and NGOs to allocate resources more efficiently, and to design financial education initiatives that deliver meaningful impact.

The rest of this paper proceeds as follows. Section 2 describes the sample and the study design. Section 3 presents the empirical strategy, summary statistics, and randomization checks, while Section 4 discusses the results. Finally, Section 5 concludes with implications of the study. Appendices 1 and 2 provide information on the content of our financial education and health videos, and present the financial knowledge survey questions, respectively.

2.2 Sample and Study Design

Our study sample consists of over 1,300 urban poor households in Ahmedabad, a metropolitan city in the state of Gujarat, India. To manage the large sample size, we conducted the study in four waves. The sample size in each wave is reported in Table 2.1, Panel A. Respondents came from different *chalis* (neighborhoods), which were mutually exclusive across waves. Furthermore, all respondents were associated with Saath, our non-government partner organization. About half were clients of Saath's microfinance services, while the other half were participants of Saath's other urban development programs such as livelihood training.

The recruitment of study subjects proceeded in the following manner. Our field staff first obtained a geographic listing of all households in a given neighborhood. Using this list, a field officer visited every fourth household[5] in the neighborhood and selected the decision maker, earning member of the family, or his/her spouse as the respondent for that household. The field officer then invited this person to participate in a life skills training program that was marketed as a program to help them better their lives, without mentioning the specific program content. The field officer also informed subjects about the time commitment required for the training (i.e., five consecutive sessions over five weeks, once a week for two hours) as well as the program location. If the respondent chose to participate, the field officer recorded the necessary contact information together with training days and times that were convenient for the respondent. All field staff were trained extensively prior to these household visits to ensure adherence to these project protocols.

The research design consisted of two main components. First, two-thirds of the study sample was randomly assigned to a comprehensive classroom-based financial education program. The remaining one-third of the sample was assigned to a similar classroom-based health education program. These assignments were only revealed to study participants when they attended their first training session. Second, the design included three additional treatments: pay for performance, financial counseling, and goal setting, which are described in detail below. The pay for performance treatment was orthogonal to all other treatments in the study. Specifically, half of all study subjects—selected individually at random, independent of whether they received financial or health education—were paid Rs. 10 (US$ 0.20) for correct answers on test questions *related* to their program: financial education participants received payments for financial questions, and similarly, health training participants received payments for health questions. The other half also received the same cash reward, but for correct answers to questions *unrelated* to their video training: financial education participants were paid for health questions, and vice versa.

The goal setting and financial counseling treatments were offered only among the set of individuals who were offered financial education classes. Within this subset of the full sample, we administered the goal setting exercise to a randomly selected half of participants, independent of their pay for performance treatment and financial counseling treatment status.

[5]During our field operations, it was also possible that every third or fifth household was selected, depending on the size of the neighborhood.

Likewise, we offered financial counseling to a randomly selected half of participants in the financial education group, independent to their pay for performance and goal setting status.

Table 2.1, Panels B and C indicate the percentage of our sample that received these treatments. We also note that all treatments were stratified within each wave based on the respondent's gender, whether the respondent was currently a client of Saath microfinance, and the respondent's neighborhood.

Data collection included a comprehensive baseline survey prior to program implementation, a post-intervention knowledge survey administered in the respondents' households three weeks after the final training session, and an endline survey implemented ten months later.

Classroom-Based Financial Education

Both the financial education and health training programs consisted of five consecutive weekly sessions, each lasting two to three hours. To control for Hawthorne-type effects, the control group was provided with health training instead of no training at all to ensure that both treatment and control groups experienced similar levels of "disruption" in their daily activities due to the weekly sessions. All respondents were assigned to attend a particular treatment or control class of about 20 participants. For each wave of the study there were about 15 classes (10 treatment and 5 control), which met at the same time every week for the duration of the program. Classes were held at a nearby training center equipped with computers, where the respondents watched their respective training videos.

All respondents received a Rs. 50 (US$ 1) show-up fee for each session they attended, and were provided free transportation to and from their homes to the training center for each visit. This fee amount was chosen deliberately to serve only as a token of gratitude for participation. Each video screening lasted for two to three hours and took place at a time of day when typically, men have to go to work and women have to do household chores. Our baseline data shows the median household income in our sample was Rs. 5,900, which translates to Rs. 195-235 per day in wages given reasonable assumptions about the number of days worked in a month. Thus, the opportunity cost of attending the program was likely much higher in comparison to the show-up fee and if so, individuals who chose to attend did so because they valued other benefits (e.g., learning from the course).

Earlier work on classroom-based financial education for adults has found limited effects of such trainings across a number of different settings.[6] One criticism of these previous studies is that perhaps the education programs were not comprehensive enough, not long enough, or not engaging enough. Our financial education program was designed specifically to address these concerns and in this regard, differs in structure from those studied previously. For example, in comparison to several existing studies which examine a short one-off financial training session,[7] the program we evaluate was more intensive—carried out over five weekly

[6]See Fernandes, Lynch, and Netemeyer (2014) for a review.

[7]Miller et al. (2015) find in their meta-analysis that more than one-third of financial education programs are delivered within one week or less.

meetings, each lasting 2 to 3 hours—to account for the possibility that a longer engagement might be needed to influence the financial habits of adults.

Additionally, whereas several prior financial education initiatives have been unsuccessful because of lack of interest among participants,[8] our financial education program included several interactive features to ensure a high level of participation. Each video was broken up into shorter clips of less than ten minutes each. To maintain participants' interest, each clip covered only two or three concepts, and there was a short break in between any two clips. Moreover, a skilled moderator led a group discussion at the end of each video session. The moderator engaged with the participants by answering their questions, eliciting their opinions, and touching on their real-life situations. Additionally, the moderator used worksheets, write-boards, and picture cards to make the discussions more interactive as well as games to foster participants' enthusiasm.

Although the implementation structure of the financial education program we study makes it more intensive and engaging than others, the course content itself was similar to those used in previous research. In particular, our curriculum was developed through an iterative process starting from standard materials developed by Freedom from Hunger, Microfinance Opportunities, and Citi Foundation that have been used in other studies. We adapted these materials to our local context of urban India together with our local research partner, our local implementation partner, and a local media company. These adapted materials were then used to professionally produce original videos using real life examples in familiar neighborhoods and with locally-known actors. The financial education videos included the following five topics: budgeting, savings, loans, insurance, and a final summary video. The health training videos covered topics unrelated to financial knowledge, specifically: cleanliness and hygiene, midwifery, maternal and child health, condoms, AIDS and syphilis, and night-blindness.[9]

While financial concepts can be taught using a broad variety of approaches, our research team deliberately chose videos as a medium with the intention that they can be used to facilitate replication, reach a wider audience, and achieve scale—all at potentially lower operation costs—and that they could serve as a foundation for governments and NGOs in other countries to adapt to their respective settings. The portability of videos thus makes it possible to scale up and generalize our financial education program to other contexts, both within and outside of India.

[8]For example, Chong, Karlan, and Valdivia (2010) report that a video- and radio-based financial literacy program in Peru had to be discontinued due to low implementation levels and meager attendance. Bruhn, Ibarra, and McKenzie (2014) echo these results, describing low take-up of financial education among adults in Mexico as well as poor participation rates during the workshop itself.

[9]The study team did not produce the health videos; rather, we utilized videos previously used in Gujarat by the United Nations for health education promotion. Appendix B.1 explains the content of both the financial education and health videos further.

Additional Treatments

Pay for Performance

The objective of the pay for performance treatment was to test whether financial learning is constrained by motivational factors in addition to the knowledge barriers examined with classroom-based financial education training. By offering payments for correct responses on a knowledge-based test, the treatment assessed whether concrete monetary incentives can more effectively induce individuals to lean, retain, and apply financial knowledge.

The impact of monetary incentives on academic achievement is fairly inconclusive in the education literature. Some studies find positive effects: for example, Angrist and Lavy (2009) show that in Israel, cash rewards led to a significant increase in high school certification rates for girls, and Kremer, Miguel, and Thornton (2009)report that in Kenya, a merit scholarship including school fees and a cash grant substantially raised test scores for primary school pupils. But other papers estimate minimal impact, such as Fryer (2011) who find zero effects of financial incentives on student achievement in Dallas, New York, and Chicago; or Bettinger (2012) who finds that cash payments improved elementary student scores in math but not reading, science, nor social science. Still others find positive but small impacts of monetary rewards on sub-groups of college students, as in Angrist, Lang, and Oreopoulos (2009) and Leuven, Oosterbeek, and Klaauw (2010). Together, these studies suggest that the effect of cash payments varies widely across different scholastic levels and environments; our study aims to determine whether such incentives are productive in the financial education context.

In addition, although much of this literature considers academic outcomes among students, our paper is, to our knowledge, the first to test incentives specifically for increasing financial literacy among adults.[10] Ex ante, it is unclear whether financial incentives in schools will have similar impacts in financial education because both environments vastly differ. For instance, in comparison to students, adults participating in financial education programs have distinct family commitments, financial experiences, and financial situations, among others. Adult behavior may also be more difficult to change than that of the youth, for example due to stickier preferences and more binding day-to-day constraints (Bruhn et al. 2016). The effects of monetary rewards among students may therefore not directly translate to adult financial education, and this is precisely the knowledge gap that our study seeks to fill.

The pay for performance analysis also contribute to the ongoing debate surrounding the value of financial education. Proponents and opponents alike have argued that weak enthusiasm for gaining financial skills presents a significant barrier to the success of financial education initiatives. For example, the Financial Literacy Foundation (2007) cautions that providing education resources alone is inadequate because the key challenge lies in promoting engagement among "those who, for reasons of disinterest in the issue, lack of perceived relevance, stress or other obstacles [...] are not currently seeking to build their money skills."

[10] A related paper is Bruhn, Ibarra, and McKenzie (2014), which focuses on monetary rewards for attending a financial education program. Our study differs in that we focus on cash incentives to learn financial concepts (conditional on attendance), and not attendance itself.

Similarly, in her article "Against Financial Literacy Education," Willis (2008) points out that although voluntary personal finance courses are widely available, participation is low unless some "perk" is awarded, making lack of interest a very costly obstacle for financial education to overcome. The pay for performance treatment allows us to evaluate these foregoing claims: we investigate whether small monetary incentives can help shift effort towards learning financial concepts and thus surmount behavioral barriers arising from disinterest or insufficient enthusiasm among financial training participants.

The logistics of the pay for performance treatment were as follows. Payments were made based on performance on a post-intervention knowledge test administered three weeks after classes ended and comprising three dimensions of financial knowledge (see Appendix B.2 for the exact wording of these questions). The first set of questions tested financial numeracy. Respondents were asked questions that involved numerical calculations, such as comparing monthly versus weekly interest rates and adding household income/expenses. The second set of questions measured respondents' financial awareness. These focused not on computation but instead on general concepts related to financial products and financial planning. For example, respondents were asked about the purpose of a household budget, minimum bank account opening requirements, and whether bank savings accounts had deposit insurance. Finally, the third set of questions assessed financial attitudes and perceptions, measured by asking respondents what financial advice they would give to their friends. For instance, respondents were asked whether they would suggest buying insurance or increasing savings to a friend who had a risky job. The purpose of these questions was to assess whether individuals understood the financial situation described and were capable of identifying the correct type of product or advice for each setting.[11]

Concrete Goal Setting

An important research question we address in this study is whether behavioral factors such as lack of self-control influence financial decisions and outcomes. We theorize that low self-control may be an impediment to the conversion of financial knowledge into positive financial outcomes. To examine this behavioral constraint, we implemented a treatment that encouraged a randomly selected subsample participants to set concrete yet non-binding financial goals with designated completion dates.

The role of non-binding goals in alleviating procrastination and self-control problem remains understudied in economics thus far, but is a growing field of inquiry at the research

[11] For example, one hypothetical question asked what type of financial advice would be appropriate for a family where the main income earner had an inherently risky job working on the exterior of tall buildings. Respondents were asked to choose between the following three recommendations (a) He should quit his job; (b) The family should start saving; or (c) He should buy accident insurance. The responses were then graded on a scale with (c) obtaining the highest score followed by (b). Hence a higher score on the financial attitudes measure represents a better understanding of financial situations and attitudes towards appropriate solutions. The full set of questions is available in Appendix B.2. We acknowledge that responses to these types of questions may be influenced by the respondents' level of risk aversion.

frontier. One prominent hypothesis at the forefront of behavioral economics is that non-binding goals act as reference points against which individuals measure losses and gains. Indeed, Heath, Larrick, and Wu (1999) explicitly argue that "mere" goals serve as reference points and systematically alter outcomes in the prospect theory value function. Moreover, Hsiaw (2013) asserts that a non-binding goal provides internal motivation, since the future "self" of a present-biased individual with reference-dependent preferences inherits the goal as a reference point in his utility function. In addition, Koch and Nafziger (2011) assume loss aversion and illustrate that goals make future "selves" strive harder because of fear of failing to reach the goal.

That goals act as reference points also finds strong support in the psychology literature, where research consistently confirms the positive effects of goals on task motivation. Reviewing empirical goal research in psychology, Locke and Latham (2002) write that "[g]oals serve as the inflection point or reference standard for satisfaction versus dissatisfaction [...] exceeding the goal provides increasing satisfaction as the positive discrepancy grows, and not reaching the goal creates increasing dissatisfaction as the negative discrepancy grows" (p. 709). Locke and Latham (2002) also summarize thirty-five years of empirical psychology research on goal-setting theory and identify four main mechanisms through which non-binding goals affect performance. First, goals provide direction both behaviorally and cognitively by focusing attention and effort towards goal-related activities.[12] Second, goals serve an energizing function as evidenced by high goals leading to greater effort than low goals.[13] Third, goals affect persistence by prolonging effort and increasing work intensity.[14] Finally, goals impact action indirectly by leading to the discovery and/or use of relevant knowledge and strategies.[15]

Goal setting has also been shown to be important in financial decision-making. Existing research has studied the effects of goal setting among consumers on repayment, spending, and saving behavior. For instance, Karlan et al. (2016) find that text messaged or mailed reminders highlighting a client's particular savings goal were twice as effective as those that did not. Soman and Zhao (2011), using a field study among Indian households, also find that setting specific goals had a significant positive effect on savings rates. Moreover, setting a single goal (in this case, financing their children's education) resulted in higher savings than setting multiple goals (savings for education, health care, and retirement). Ülkümen and Cheema (2011) observe that for more ambitious savings targets, having concrete goals as opposed to general goals can increase the perceived importance of and commitment to the target.

[12]As an example, Rothkopf and Billington (1979) had high school students study a passage with goal-relevant and non-relevant text. Recording students' eye movements revealed that students fixated on goal-relevant sentences over twice as long as non-relevant sentences due to the direction provided by the goal.

[13]Empirical evidence include Bryan and Locke (1967) and Bandura and Cervone (1983), among others.

[14]For instance, LaPorte and Nath (1976) found that subjects who were presented with a difficult goal for answering questions correctly about a reading passage studied longer, more persistently, and produced more correct answers when tested.

[15]For example, a study by Earley and Perry (1987) showed that when individuals are trained with the proper strategies, those who had high-performance goals experienced improvements because they were more likely to use the given strategies.

On the credit and repayment side, several studies find that when confronted with different credit card payment options, individual financial goals often determine the payment amount selected by consumers (Bartels and Sussman 2015; Salisbury 2014; Agarwal et al. 2014). (Thaler 1999) and Soman and Cheema (2011) further explore goal setting as a form of "mental accounting" and find that people are more disposed to honor spending targets that are earmarked for certain product categories.

In our study, participants who received the goal setting treatment were encouraged to set short- term achievable but non-compulsory financial goals. This treatment involved a household survey, implemented within four weeks of the financial knowledge exam, wherein respondents were interviewed about their use of financial services. Notably, respondents were also asked to voluntarily choose a target date for completing one or more financial goals: opening a savings account, increasing savings, reducing expenditure, and/or purchasing insurance. Surveyors recorded these target dates on a calendar provided by the study at no cost and posted in the respondent's home, so that subjects may be reminded of their self-chosen, non-binding financial goals.

To measure marginal effects beyond financial education alone, we administered this goal setting exercise by design to a randomly selected half of participants assigned to financial training. The remaining half served to isolate the effect of goal setting from that of the household visit, as they received a similar household survey on financial services during the same period, but they were neither asked to set financial goals and target dates nor given any calendars. To summarize, the treatment group received a household survey, a calendar, and were asked to set a target date for a financial goal on this calendar, while the control group received only the household survey. Hence, goal setting measures the combined effect of both the calendar and the target dates, which we consider together as one treatment.

We acknowledge that the goal setting treatment we study does not allow us to identify a single underlying mechanism through which non-binding goals help attenuate self-control problems. Instead, our approach tests the value of "a foot in the door," whereby prompting individuals to develop a plan of action (such as setting target dates for goals) increases the likelihood of attaining the goal. For example, in a field experiment on influenza vaccination, Milkman et al. (2011) show that prompting individuals to write down date and time they plan to be vaccinated increased vaccination rates. Similarly, in a field experiment during the 2008 US presidential elections, Nickerson and Rogers (2010) show that facilitating the formation of a voting plan increased voter turnout. In both of these studies, neither the target vaccination time nor the voting plan were binding, and yet these "implementation intentions" resulted in meaningful positive effects.

Individualized Counseling

The final treatment in our study was designed to test whether the intensity of financial education and the medium in which it is delivered affects knowledge acquisition and application. Our hypothesis is that traditional classroom-based financial education trainings may be insufficiently suited to individuals' specific learning needs. We test the role of the education medium by supplementing the financial education trainings with individualized counseling. This treatment consisted of one-on-one, in-person counseling at home, where the counselors aided in tasks such as preparing a budget, opening a bank account, paying a loan, or buying insurance. Such counseling may be more effective in changing behavior as it provides guidance specific to the needs of the participant.

Medical and public health studies have found individualized or segmented counseling to be effective in promoting better health behaviors. For example, individualized risk counseling for women with a family history of breast cancer has been shown to improve understanding of their personal risk (Lerman et al. 1995). Similarly, Proper et al. (2003) find positive and significant effects of individual counseling on physical fitness. In the financial context, Dalal and Morduch (2010) find that having an insurance representative present after trainings significantly improves take-up rates. Similarly, Bertrand et al. (2006) find that allowing banking workshop participants the opportunity to complete account opening paperwork as part of the learning workshop and having a bank representative present on-site significantly improves take-up and adoption of complementary banking products such as ATM cards, direct deposit, and electronic fund transfers. Finally, psychologists have long advocated the benefits of human interaction in individualized counseling over inanimate information sources such as pamphlets, text messages, or computer messages (King et al. 2007).

The counseling treatment in our study was randomly assigned among financial education participants. Half were randomly selected to receive an offer of financial counseling, independent of their goal setting treatment status. Specifically, within one month of the classroom sessions, financial counselors visited the counseling treatment group in their homes to provide individualized financial counseling services. The financial counselors assisted participants on several issues—including, but not limited to, preparing a budget, opening a bank account, paying off or re-financing loans, and purchasing an insurance policy—depending on their individual needs. Financial counselors were trained rigorously by our partner research organization in India, the Center for Microfinance, prior to visiting respondents. The treatment involved monthly household visits by the counselors for the duration of the study.

As a final point, we note that all of the treatments in our study were agnostic about specific financial products or providers: the financial education program focused on explaining concepts related to savings, loans, and insurance (e.g., interest rates, premiums) as well as the importance of scouting the market for financial products appropriate for participant's individual needs; similarly, in the goal setting and counseling treatments, respondents selected their own goals and counseling agendas, respectively. Participants were never pressured to adopt, purchase, or join any specific financial product, service, or provider.

2.3 Empirical Methodology and Summary Statistics

Empirical Methodology

The main analysis of this paper estimates causal intent-to-treat (ITT) impacts on financial knowledge and behavior. First, we analyze impacts on three distinct components of financial knowledge, namely financial numeracy, awareness, and attitudes. We study pay for performance impacts using data from both the short-term and endline surveys.

Since financial education and pay for performance were orthogonal treatments both randomized at the individual level, we estimate causal effects on financial knowledge with the following OLS model:

$$\begin{aligned} Y_i = \alpha + \beta_1 FinEd_i + \beta_2 PayforPerf_i + \beta_3 FinEd\ and\ PayforPerf_i \\ + \beta_4 DiscountRate_i + \sum_k StrataDummy_{ik} + \epsilon_i \quad (2.1) \end{aligned}$$

where outcomes Y represent financial knowledge measures from the survey; $FinEd$ is a dummy equal to 1 for an individual i who was assigned the financial education treatment; $PayforPerf$ is a dummy equal to 1 for an individual i who was offered pay for performance on financial knowledge questions; and $FinEd\ and\ PayforPerf$ is the interaction term.

Next, we estimate treatment impacts on financial behavior using data from our endline survey. Since we have four treatment combinations (i.e., financial education alone; financial education with counseling; financial education with goal setting; and financial education with both goal setting and financial counseling), we analyze results with a saturated model to simplify interpretation:

$$\begin{aligned} Y_i = \alpha + \beta_1 FinEd_i + \beta_2 FinEd\ and\ Goal_i + \beta_3 FinEd\ and\ Couns_i \\ + \beta_4 FinEd\ and\ Couns\ and\ Goal_i + \beta_5 DiscountRate_i + \sum_k StrataDummy_{ik} + \epsilon_i \quad (2.2) \end{aligned}$$

Here, the outcomes Y represent responses to financial behavior questions from the endline survey. $FinEd$ is a dummy equal to 1 for an individual who received the financial education treatment, but not the financial counseling or the goal setting treatments. $FinEd\ and\ Goal$ is a dummy equal to 1 for an individual who received both the financial education and goal setting treatments, but not the financial counseling treatment. Similarly, $FinEd\ and\ Couns$ is a dummy equal to 1 for an individual who received both the financial education and counseling treatments, but not goal setting. And finally, $FinEd\ and\ Couns\ and\ Goal$ is a dummy equal to 1 for an individual who received all three treatments. The omitted category is group that did not receive any financial education, the control group.

For both equations (2.1) and (2.2), we include a control for the baseline discount rate, which shows an imbalance in Table 2.2. We also control for strata dummies for precision, since in each wave of the study we stratified the randomization. Strata are defined by gender, whether the respondent is currently a client of Saath microfinance, and neighborhood. Note that since neighborhoods were mutually exclusive across waves, we do not add wave fixed effects. Furthermore, in each study wave, participants were assigned to attend a particular class that met at the same time every week for the duration of the training program. Classes consisted solely of either financial education training participants or health training participants. In estimating equations (2.1) and (2.2), we cluster standard errors at the wave-class level.

Summary Statistics and Randomization Checks

Baseline characteristics for our sample are presented in Table 2.2. Households in our sample comprised 6 members on average, with a mean monthly income of Rs. 7017 (US$ 120). A little more than half (58 percent) of our respondents were female, and a vast majority was married. Respondents in our sample also had limited schooling, with 47 percent having completed elementary school, but only 4 percent having completed secondary school.

In addition to standard data on household demographics and respondent characteristics, our baseline survey measured financial knowledge, attitudes, and preferences. First, we note that almost everyone in our sample (94 percent) reported having difficulty saving. Next, we measured discount rates in the standard manner, by asking respondents to provide the minimum amount they would be willing to hypothetically accept in one month in lieu of a hypothetical payment of Rs. 350 today. Respondents in our sample reported relatively high monthly discount rates: the median was 0.14, while the average was 1.52. We also measured risk aversion by allowing respondents to choose between a payment of Rs. 10 with certainty, or playing a lottery that pays out Rs. 25 or Rs. 0 with equal probability. 18 percent of our sample chose the safe payment, and these respondents were coded as risk averse.

We also measured basic computational skills through a series of eight mathematics questions. The mean score for these mathematics questions was 4.73 out of 8. We find similar computational skill levels as in Cole, Sampson, and Zia (2011) in Indonesia. Specifically, almost all respondents could answer a simple addition question ("How much is 4+3?"), but only about 50 percent were able to answer a multiplication question correctly ("What is 3 multiplied by 6?"). Even fewer respondents were able to make percentage calculations correctly ("What is 8 percent of 100?"), with close to half responding "do not know" to this question.

Finally, we measure baseline levels of financial knowledge based on the following three questions, which are a standard set provided by Lusardi and Mitchell (2009): 1) "If you borrowed Rs. 5,500 and were charged 12 percent interest per month, how much interest would you pay in the first month?"; 2) "Suppose you had Rs. 100 in a savings account and the same amount saved at home, which of the two will yield returns at the end of the year?"; and 3) "Suppose your friend inherits Rs. 10,000 today and his brother inherits Rs. 10,000 three years from now. Who is richer because of the inheritance?" Measured financial literacy was low in our sample, with an average score of 1.6. Similar to the mathematics questions, few respondents (less than 10 percent) were able to calculate interest rates correctly in question 1, and over 60 percent responded "do not know" to this question. In contrast, almost all respondents were aware that a savings account yields positive returns (question 2), but only

58 percent of our sample was able to correctly identify the time value of money (question 3), lower than what Lusardi and Mitchell (2009) find among respondents in the US.

Table 2.2 provides a test of the randomization. The p-values in column 4 report the statistical significance of a joint test for the difference between the baseline means across all treatments including the control group. As the table shows, the p-values are fairly large, suggesting no significant difference across the treatments in baseline measures. The only baseline variable that shows imbalance across treatments, the monthly discount rate, is controlled for in all regression specifications.

Finally, attrition in our sample was very low, at less than 6 percent of the entire sample over the four waves from baseline to final follow-up, and uncorrelated with treatment status.

2.4 Results and Discussion

In this section, we present and discuss results on the short-term impacts on financial knowledge as well as the longer-term impacts on both financial knowledge and behavior.[16]

As described in Section 3.1, the regression analysis presented in this paper estimates intent-to- treat (ITT) effects. In particular, the sample in the regressions includes all study participants regardless of whether they actually attended the screenings, which is an endogenous choice. Note that the take-up of our various interventions was quite high: 93% of those assigned to financial education attended at least one video screening, 88% of those assigned to goal setting chose to set at least one financial goal, and finally, 64% of those assigned to financial counseling accepted the counselor's services.

Given the high attendance rate, we focus on the ITT effects rather than the treatment-on-treated (TOT) estimates. Moreover, the TOT results only provide local average treatment effects (LATE)—that is, the average impact among those induced to change their choice by the instrument—which would be difficult to extrapolate to the whole population. The focus of our analysis therefore remains on estimating impacts using ITT.

Financial Knowledge and Pay for Performance

We find varied short-term effects of traditional financial education, with no impact on participants' financial numeracy scores but strong positive effects on aggregate measures of financial awareness and attitudes. Table 2.3 presents results on aggregate measures of financial knowledge, while Appendix Tables B.1 to B.3 present regression results on individual questions for each category of numeracy, awareness, and attitudes. The longer term effects are likewise reported in Table 2.4 (aggregate measures) and Appendix Table B.4 (individual questions).

We consider a variety of different outcome variables as proxies for financial numeracy, including questions on selecting financial products and budgeting capabilities. The short-term

[16] A related paper discusses measurement issues on financial literacy and how our measures of financial knowledge allow for disaggregated impacts on numeracy, awareness, and attitudes. See Carpena et al. (2015).

results presented in Table 2.3 and Appendix Table B.1 indicate no impact on financial numeracy. Moreover, even the addition of pay per performance did not yield a positive effect in the short-run on financial numeracy skills. Table 2.4 and Appendix Table B.4 validate these findings for the long- term as well. These results show that financial education failed to help individuals choose the loan option that minimizes expenses, to select the most appropriate savings or insurance product, or to write a budget effectively. Incentivizing individuals with payments on correct answers led to no significant improvement in financial numeracy scores.

These results corroborate the existing literature which finds that financial education, no matter what form it takes, has little effect on financial numeracy skills. For instance, Jamison, Karlan, and Zinman (2011) find no effect of a ten-week financial education course on financial numeracy among youth clubs in Uganda. Similarly, Doi, McKenzie, and Zia (2014) find no effect on numeracy skills of migrant workers in Indonesia who attended financial education classes prior to being assigned to work overseas. Carpena et al. (2015) discuss these limitations of financial education and proposes measuring financial knowledge in terms of awareness and attitudes rather than strictly in terms of numeracy.

In contrast to the null effects on financial numeracy, our results show that the financial education program significantly improved financial awareness and attitudes towards financial products. The results presented in Table 2.3 show that individuals who received financial education improved financial awareness and financial attitudes by 7 percentage points and 8 percentage points, respectively compared to the control group. Analyzing the individual questions in Appendix Table B.2, those who received financial education were 16 percentage points more likely to know minimum bank account opening requirements, 13 percentage points more likely to distinguish bank processing fees, and 20 percentage points more likely to understand unproductive loans relative to the control group. Appendix Table B.3 shows similar positive impacts on short-term financial attitudes—when hypothetically asked to give financial advice, treated individuals were 10 percentage points more likely to suggest insurance cover for a dangerous work environment and 20 percentage points more likely to suggest making a budget to track household income and expenditure relative to the control group. Table 2.4 and Appendix Table 2.4 show that these results hold in the long run as well.

These results on financial awareness and attitudes corroborate findings from several previous studies that show similar effects. For example, while Jamison, Karlan, and Zinman (2011) find no effect on numeracy scores among Ugandan youth clubs, they do find a significant positive effect on aggregate financial knowledge scores among those who were offered financial education. Likewise, Doi, McKenzie, and Zia (2014) find significant improvements in measures of awareness and attitudes similar to ours among migrant workers in Indonesia.

Next, we analyze the impact of pay for performance and find that it did not lead to any significant marginal improvements over the standard curriculum on either of the aggregate measures of awareness or attitudes, just as it did not induce variation in treatment effects on financial numeracy. We can rule out the concern that the financial incentives offered were not large enough to be salient to participants since, with correct answers to all 18 questions in the financial knowledge test, respondents could have earned up to Rs. 180, an amount close to a full day's wage.

The results in Table 2.3 and 2.4, therefore, suggest that participant motivation was not a critical barrier in improving financial knowledge in our sample, and we estimate these null effects relatively precisely. Specifically, in Table 2.3, we see that the marginal effect of pay for performance in the short-run—obtained by summing the coefficients for "Pay for Performance" and "Interaction of Financial Education and Pay for Performance"—are all very close to zero, with an estimate of -0.007 for numeracy, 0.015 for awareness, and -0.016 for attitudes. We consider these null effects to be reasonably precise since their respective 95% confidence intervals, i.e., [-0.023, +0.036], [-0.013, +0.042], and [-0.046, + 0.014], suggest quite small effects in comparison to the control group means and standard deviations. Even more strikingly, we see in Table 2.4 that the same results still hold almost one year after the program ended. Hence, integrating pay for performance into financial education led to no additional improvements in financial knowledge either in the short- or long-run, and these relatively precise null results can aid financial education policy makers, stakeholders, and NGOs in optimally designing financial education programs.

Financial Behavior

Our analysis on financial behavior comes from the endline survey. Data from this survey also helps distinguish impacts of additional treatments of goal setting and individualized counseling over traditional financial education. The specific behaviors we study are the ones targeted by the financial education program: budgeting, savings, borrowing, and insurance adoption.

Budgeting

We first consider changes in household budgeting, the theme of one of the five financial education video sessions. Existing research has shown important benefits of writing down income and expenses for planning finances, starting savings, and managing spending (Miller et al. 2015). Record keeping and tracking expenditures are often cited as critical elements of gaining control of one's finances, much the way that many fitness and diet programs focus on recording eating and exercise habits to control weight and improve health. This is a behavior that is fully under the control of the individual as compared to decisions to default or even to save money which may be influenced by factors outside one's control such as unexpected illness (and medical fees), loss of a job, or other problems that lead to financial distress. The meta-analysis of prior literature in Miller et al. (2014) indicates that financial education may positively encourage record keeping behaviors. From a policy perspective, budgeting and record keeping are relatively simple to target since advocating for regular record keeping does not require institutional change or the creation of new financial products as would be the case for some other financial behaviors such as formal savings, loans, and insurance.

In columns (1), (3), and (5) of Table 2.5, we report impacts of being invited to any financial education treatment on beliefs that budgeting is helpful (column 1), attempts to make a budget in the last six months (column 3), and making a regular monthly budget (column 5). Without distinguishing between treatments, we see a strong positive treatment effect on all these dimensions, though the effects weaken as we move from beliefs to actions and outcomes. Specifically, while those individuals invited to any financial education treatment were 22 percentage points more likely than the control group to understand the benefits of making a budget and 28 percentage points more likely to have attempted to make a budget, they are only 3 percentage points more likely to actually make a regular budget every month.

We delve into mechanisms by analyzing the treatments separately in columns (2), (4), and (6). Our results show that the medium of delivery makes a substantial difference in longer-term budgeting behavior. Providing classroom-based financial education alone generally yields weaker results than when it is complemented with higher-intensity, personalized treatments. We find that those who received the single financial education treatment were 17 percentage points more likely than the control group to think that budgeting is helpful, while combining financial education with the other two treatments yielded a 26.5 percentage point improvement. Importantly, the p-value on the F-test comparing all three treatments against financial education alone is 0.026, suggesting significant marginal improvements over financial education alone due to the add-on treatments. Similar results are reported when either goal setting or counseling alone are part of the financial education package.

The regression results also find important distinctions across treatments when moving from beliefs about budgeting to action. Notably, the effect of financial education classes and goal setting is limited to raising awareness about budgeting, but stops short of regular behavior change. For instance, in column (4), those who were invited to financial education alone are 13 percentage points more likely than the control group to have attempted to make a budget in the last six months, but this effect disappears when it comes to making a regular monthly budget (column 6). Adding goal setting improves the attempt to make a budget slightly to 16 percentage points (not statistically distinguishable from financial education alone) but again the effect does not persist for regular monthly budgets.

The significant effect on sustained behavior change comes from adding counseling to the mix. Individuals invited to financial education with personal counseling are 39 percentage points more likely to have attempted a budget and 4 percentage points more likely than the control group to make a regular monthly budget. These effect sizes are even larger for the highest intensity treatment (financial education with goal setting and counseling) at 43 percentage points and 5 percentage points, respectively.

The fact that financial counseling is key to sustained budgeting behavior is important. While both financial education classes and goal setting can highlight the importance of budgeting, individuals may still lack the necessary skills to actually maintain a regular budget given their unique individual circumstances. Our results suggest that personalized counseling acts as a critical bridge that enables individuals to apply their acquired financial knowledge to improve behavior.

Savings

We next turn to long-term impacts on household savings behavior. A long line of research in development economics shows that the incomes of poor households in the developing world are not only low, but are also extremely irregular and unpredictable (e.g. Morduch 1995). This is particularly true in our context, urban India, where many of those employed are casual laborers (such as helpers and cooks) who may or may not have work on any given day, or own-account workers (such as auto-rickshaw drivers and street vendors) whose earnings largely depend on sales. In such an environment with highly variable earnings, storing past income through savings becomes an essential financial tool for the poor, enabling them to put food on the table every day, and fundamentally, to manage their uneven cash flows (Collins et al. 2009).

During our study period, households in our sample had access to "no-frills" savings accounts, a type of bank savings account designed specifically for low-income individuals and mandated by the Reserve Bank of India (RBI), the country's central bank, to increase financial inclusion. In particular, these "no-frills" savings accounts have initial deposits, minimum balances, and other charges that are either zero or very low. For example, during our study, the State Bank of India offered such accounts with an initial deposit of Rs. 50 and zero maintaining balance thereafter, while UCO Bank required Rs. 250 for the initial deposit and Rs. 5 maintaining balance. Importantly, these "no-frills" accounts earn a strictly positive interest rate that is similar to other regular savings accounts (between 2.5% to 4% per year during our study), and they are also reliable because in the event that the bank shuts down, all deposits are insured by the Deposit Insurance and Credit Guarantee Corporation, an institution similar to the FDIC in the US. Taking these factors together—zero or very low maintaining balances and fees, positive interest rates,and deposit insurance—holding formal savings accounts thus posed little to no costs among the households in our setting.

Apart from availability of suitable formal products, previous research has shown that savings products may offer additional advantages beyond the interest earnings. By keeping money inaccessible, savings accounts may protect against financial demands from family members or neighbors (e.g., Ashraf 2009) and they may discourage temptations to spend especially for those with present-biased preferences (e.g., Laibson 1997). In addition, they allow households to create large sums for big-ticket purchases such as furniture and education, while building financial relationships which can be leveraged for accessing loans (e.g., Collins et al. 2009). Moreover, research has shown that access to savings facilities yield real welfare benefits for the poor beyond positive interest rates: they increase savings, productive investments, and food expenditures, and importantly, reduce overall poverty (Dupas and Robinson 2013; Burgess and Pande 2005, e.g.,).

The regression results on savings in our study are presented in Table 2.6. In columns (1), (3), and (5) we report impacts of being invited to any financial education treatment on holding of informal savings (column 1), holding of formal savings (column 2), and investments in fixed or recurring deposits (column 5). Without distinguishing between treatments, we find that participants who received any form of financial education intervention were 4 percentage points more likely to hold informal savings and 8 percentage points more likely to hold formal savings in a bank account.

As with budgeting, the medium of instruction is critical for motivating sustained behavior change. Financial education alone produced no effect on any of the savings outcomes we measured—participants who received only the financial education treatment were no more likely to hold savings, formally or informally, than the control group. These findings suggest that it may be more difficult to influence households' savings compared to altering budgeting behavior by using a traditional program of financial education. Moreover, classroom-based models may not be adequate to address cognitive barriers or resource constraints that are likely to inhibit households from changing their current savings practices.

The results on add-on treatments indicate that the type and intensity of the intervention has a significant influence on savings. Incorporating goal setting and/or counseling did produce changes in savings behavior and our results offer insights into the mechanisms of impact. Participants who received goal setting in addition to financial education were 6 percentage points more likely than the control group to save informally (in a neighborhood fund or at home) and 8 percentage points more likely to save formally at a bank. Both these results are significant at the 5 percent level. In contrast, the results for counseling are different: we find no significant effect of adding counseling on informal savings but a 13 percentage point improvement in the likelihood of opening a formal bank account over the control group, a result that is statistically significant at the 1 percent level. The p-value on the F-test comparing the combined financial education and counseling treatments against financial education alone is 0.021, suggesting significant marginal improvements over financial education alone due to the add-on counseling treatment.

These results suggest that while financial education classes and goal setting can inform and encourage people to save, respectively, they still may lack the skills needed to open and maintain a bank account. While goal setting did appear to increase in the likelihood of saving, the effects of counseling are honed in on formal savings, with an effect size nearly double that of goal setting. Counseling thus appears to enable participation in the formal financial sector and, as with budgeting, serves as a bridge that enables individuals to convert their acquired financial knowledge into financial actions.

Finally, sustained behavior change in household investments in fixed or recurring deposits (column 6) appears more difficult to achieve. Combining financial education with both goal setting and counseling led to a modest 4 percentage points increase in the likelihood of repeated deposits; however, this result is only significant at the 10 percent level. Neither financial education alone nor financial education combined with personalized counseling yielded significant effects on fixed or recurring deposits.

Borrowing

Households in our study had the ability to borrow money from many different sources. These included private banks, cooperative societies, microfinance institutions (MFIs), credit and savings groups, moneylenders, employers, shopkeepers, pawnbrokers, and family and friends, among others. But almost all of these options share two important features. First, the loans charged a very high interest rate, driven not only by the risk in lending to the poor, but also by the short-term nature of the loans, the relatively small size of the principal, and other transaction costs (e.g., Collins et al. 2009). In our sample period, banks and MFIs in

India typically priced loans at 17-24% interest per annum, while moneylenders were at 50% per annum or above. Second, many lenders often charged additional costs when taking out a loan. These could be loan processing fees which were usually 1-2% of the loan amount, documentation fees or stamp duties which are fixed amounts regardless of the loan size typically around Rs. 110 (US$ 2), or upfront interest charges that ultimately increase the effective interest rate of the loan.

Since the cost of credit in our research setting was quite high, the financial education course taught the importance of borrowing wisely, the responsible use of loans, and healthy borrowing behavior. Specifically, the video session on borrowing instructed participants on understanding the different components of loan costs: comparing interest rates across different options, accounting for additional or potentially hidden fees, and recognizing loan terms that are likely to impact the overall price of the loan. Notably, the video also explained in simple, accessible language the distinction between productive loans (e.g., borrowing to buy an asset) versus unproductive loans (e.g., borrowing for consumption). The outcomes that we consider are therefore those behaviors directly related to the financial education course such as whether respondents are aware of the terms of their loan. We also examine outcomes on whether respondents borrowed for the purpose of business, education, or purchasing durable goods, which are examples of productive loans discussed in the program, rather than borrowing for unforeseen expenses or repaying other debt, which may be unproductive loans as well as potential warning signs of unhealthy financial habits.

The results for borrowing are presented in Table 2.7. In columns (1), (3), (5), (7), (9), and (11), we report impacts of being invited to any financial education treatment on outstanding loans (column 1); planned borrowing in the next year (column 3); and among the sample who took out loans since the conclusion of financial education classes: knowledge of loan terms (column 5); positive borrowing, e.g. for business, education, or durable goods (column 7); negative borrowing for unforeseen circumstances (column 9); and borrowing to repay other debt (column 11). Findings suggest that financial education yields only modest effects on household borrowing. While the treatment effect on knowledge of interest rates is positive and significant, we observe no significant difference in outstanding loans, planned borrowing, or use of loans for productive purposes.

As with budgeting and savings, the medium of instruction is important for influencing household borrowing outcomes. First, the positive impact on knowledge of interest rate terms of loans is not statistically significant among participants who received only financial education or financial education with goal setting. In contrast, those who received financial education with financial counseling or all three treatments show significant improvements in knowledge. Consistent with the results on budgeting and savings, this result highlights the value of individualized counseling in improving the readiness of households for financial products.

On financial behavior, the results show a similar pattern and financial education alone had no impact on participants' propensity to borrow or their reasons for borrowing. In contrast, adding goal setting and financial counseling did affect these outcomes. Goal setting had a suppressive effect on borrowing for the future (column 4) with participants 6 percentage points less likely to take out a loan in the next two years, an effect that is statistically significant at the 10 percent level. The coefficients on other treatments are negative as well but not statistically significant.

Goal setting also significantly reduced the likelihood of taking out loans for unforeseen expenses (column 10), and this effect is similar in the financial counseling and the combined groups. Financial counseling additionally had a positive impact on borrowing for productive purposes with borrowers 12 percentage points more likely than the control group to have borrowed for business, education, or purchase of durable goods.

Insurance

Like much of the developing world, urban poor households in India encounter substantial risks in their everyday lives including non-chronic and chronic illnesses, loss of life, loss of work, theft, and fire (e.g., Kantor and Nair 2003). Although these risks have important implications for both rich and poor households, the consequences are likely much direr for the poor because of their low and unstable income. Our study households represent a microcosm of this larger picture. At baseline, for example, 71% of subjects reported that at least one member of their household was ill in the last three months. Of those that did visit a medical facility for their sickness, the mean cost of one visit at a health facility was Rs. 1349 with a standard deviation of Rs. 3820—quite substantial in comparison to the baseline per capita monthly income of Rs. 1272.

Under such settings, insurance can be an important product in a household's financial portfolio. However, take-up remains very low—for example, baseline health insurance ownership in our sample was very low at a meager 8 percent. The financial education program therefore aimed to raise awareness about the value and suitability of insurance. In our sample period, several actuarially fair insurance products were available for study participants. For instance, in the case of health insurance, subjects were eligible for India's national insurance scheme for the poor, the Rashtriya Swasthya Bima Yojna (RSBY), which is still operational. The RSBY provides insurance of Rs. 30,000 for a family of up to five members, covering pre-existing conditions, hospitalization expenses, surgeries, and child deliveries, among others. The premiums for this insurance scheme are fully subsidized by both the central and state governments: households are required to pay only Rs. 30 (less than US $1) per year in registration fees.[17] This very low cost, together with the relatively high incidence of illness that our subjects reported at baseline, suggests that failure to adopt insurance may not have been an optimal choice on the part of respondents.

In the case of life insurance, no similar public option was available, but the market had a variety of options with 23 life insurers operating across India during our study period.[18] Another pertinent outcome we study is adoption of debt insurance. In India, debt insurance is a financial instrument that insures a loan so that in the event of the borrower's death, the outstanding loan amount is settled by the insurer. These insurance policies are typically available through banking institutions, MFIs, and insurance providers. Some lenders, such

[17] More details on RSBY can be found in `http://www.rsby.gov.in/about_rsby.aspx`

[18] There were 24 non-life insurers present in India during our study.

as banks and MFIs, often require borrowers to adopt debt insurance before the loan can be disbursed. And because such insurance is part of the package of obtaining credit, it poses additional costs to the loan that borrowers may not have been aware of. Considering debt insurance take-up as an outcome therefore allows us to examine whether respondents understood this aspect of the credit market, especially given the financial education program's emphasis on fostering participants' understanding of loan costs.

Finally, we note that all insurance providers in India operate under the umbrella of the Insurance Regulatory and Development Authority (IRDA), a government agency whose main mission is "to protect the interest of and secure fair treatment of policy holders."[19] The IRDA has a Consumer Affairs Department that is devoted exclusively to issues faced by policy holders. In addition, the Policyholder Protection Act of 2002 requires insurance companies to have effective and timely grievance redress mechanisms in place. The IRDA monitors the redress systems of insurers and operates a Grievance Call Center which provides an additional avenue through which consumers may file complaints. To prevent any fraud, the IRDA also conducts on-site inspections of insurance companies and reviews the qualifications of insurance agents. Together, these regulations on grievance redress, monitoring, and scrutiny help to ensure the viability of insurance products and providers as well as to safeguard policyholders against any unfair practices.

The results for insurance as the outcome of interest are presented in Table 2.8. In columns (1), (3), and (5) we report impacts of being invited to any financial education treatment on purchases of health insurance (column 1), life insurance (column 3), and debt insurance (column 5) within the last six months. Despite the foregoing discussion of the value and availability of suitable products on the market, our financial education program had very limited impact on insurance outcomes, with minimal adoption of insurance products in the six-to-ten months following the program.

While traditional financial education alone caused no significant changes in the take-up of insurance products compared to the control group, there is some modest evidence that a combination of high-intensity interventions yielded impacts on adoption of certain types of insurance. In particular, those who received all three treatments were 5 percentage points more likely to purchase life insurance. However, these effects do not hold for other types of insurance, such as debt or health insurance. No combination of financial education, goal setting, or counseling enticed participants to purchase these products.

These results indicate that participants likely faced additional cognitive and behavioral constraints to the take-up of insurance. One reason is that participants may view insurance products as a luxury that will not add value in the short run. Financial education will thus have a limited impact on insurance adoption relative to budgeting and savings, which are cognitively and financially easier for participants to implement. Insurance products are also relatively new in India, and the absence of peer effects and knowledge of long-term returns may partly explain participants' reluctance to purchase insurance. Due to these constraints,

[19]See https://www.irdai.gov.in/

decisions regarding insurance may be more difficult to influence through financial education compared to decisions regarding savings, borrowing, and budgeting.

Discussion of Findings on Financial Behavior

Our finding that traditional adult financial education alone did not lead to substantial changes in financial behavior is not surprising and draws parallels with the existing literature. For example, Collins (2013) finds that a mandatory financial education course for low-income families enrolled in the Federal Housing Choice Voucher program had no significant effect on savings. Cole, Sampson, and Zia (2011) similarly find that financial literacy training in Indonesia had no significant effect on the likelihood of a household opening a bank account except among those with low initial levels of education and financial literacy.

In general, the literature finds that financial behavior changes are much harder to elicit using traditional financial education programs. A meta-analysis conducted by Fernandes, Lynch, and Netemeyer (2014) finds that interventions to improve financial literacy explain only 0.1% of the variance in financial behaviors studied, with weaker effects in low-income samples. Gartner and Todd (2005), for example, evaluate a randomized credit education plan for first-year college students in the U.S. and find no statistically significant differences between the control and treatment groups in their credit balances or timeliness of payments.

The literature does find better impacts when we move away from traditional delivery channels for financial education. For instance, Drexler, Fischer, and Schoar (2014) examine the impact of two different financial education programs targeted at micro-entrepreneurs in the Dominican Republic. Members of the first treatment group participated in several sessions of traditional, principles-based financial education, while members of the second treatment group participated in several sessions of financial education oriented around simple financial management rules of thumb. Relative to the control group, the authors find no difference in the financial behaviors of the treatment group who received traditional financial education. They do, however, find statistically significant and economically meaningful improvements in the behavior of the rule-of-thumb treatment group. The results of this study suggest that the structure of financial education matters in determining its effects on behaviors, and might help explain why many other studies have found much weaker links between financial education and economic outcomes.

Results from other non-traditional financial education interventions have also shown significant effects on outcomes. Bruhn et al. (2016) find that a comprehensive financial education program targeting Brazilian high school students improved financial knowledge, attitudes toward financial products, and financial behaviors. Similarly, Berg and Zia (2017) use entertainment media to deliver financial education messages on debt management to the public in South Africa and find statistically significant improvements in content-specific financial knowledge and borrowing behavior.

The important distinction to note is that these programs are quite different from traditional financial education interventions. The program analyzed by Bruhn et al. (2016) targeted high school students and included study materials, teacher training, monitoring, and participation awards. The program was delivered by regular teachers and integrated into classroom curricula, and schools with high levels of participation received awards and public recognition; treatment intensity, then, was much higher than in most financial education initiatives. The soap opera intervention in South Africa analyzed by Berg and Zia (2017) is also unique. Instead of relying on a traditional classroom approach, the program targeted people in their home environments without placing an emphasis on financial education, and the storyline lasted for approximately two months.

In light of the knowledge that innovative methods for delivering financial education can improve its effects on outcomes, our paper investigates the ways that classroom-based financial education can be improved. By addressing three specific factors—participant motivation, goal setting, and program intensity—that may prevent financial education from benefitting recipients unless exclusively addressed, our results provide new insight into effective strategies to promote financial education.

We find that traditional financial education is largely ineffective in terms of changing financial behaviors, despite its positive effects on financial attitudes and awareness. Yet, we also find that certain small changes to financial education can strengthen the "link" between education and outcomes. While interventions that include behavioral components like goal setting have been studied before, our analysis is the first to study these interventions in the same experiment and in combination with other approaches to encourage behavioral change. Previous literature has not fully explored the disconnect between financial education and financial outcomes, and has therefore missed a crucial element of any program attempting to help improve financial behaviors.

First, we find that simple non-binding goals can address some of the hard-to-change financial behaviors, including setting of monthly budgets and savings. Second, financial counseling facilitates further sustained action, such as making a household budget regularly, saving in formal bank accounts, knowing details of loan terms, and borrowing for productive purposes. We note four components of our findings and study setting that shed light on the mechanisms behind the strong effects of financial counseling. First, respondents in our sample had limited schooling, and as seen in Table 2.2, 47 percent report completing elementary school, but only 4 percent for secondary school. Second, even though participants in the financial counseling treatment could request the counselor for assistance on any aspect of money management (e.g., preparing a budget, contacting an insurance provider), our data shows that a majority of these respondents sought the counselor's help for opening a formal bank savings account. Third, the financial education program emphasized the benefits of both informal and formal savings, yet adding counseling yielded much larger effects on the take-up of formal savings accounts than financial education on its own. And fourth, financial counseling had significant effects on formal savings, which require specific documents, but not informal savings, which often rely only on personal relationships.

These findings indicate several underlying reasons which might explain the positive impacts of counseling on financial behaviors that are typically difficult to change. For formal savings, they could suggest that respondents' inability to fill out forms could have been a critical constraint to adopting a bank account. Respondents may have found application forms overwhelming given the low levels of education in our sample. We note, however, that the bank account opening requirements in our setting were already quite minimal; respondents could open a "no-frills" savings account by submitting only a photograph, signature, and a nominal amount for the bank account opening balance. On the other hand, the significant effects of counseling on formal savings that we observe could also be due to respondents' apprehension about interacting with formal banks. Existing studies have shown that trust may be an important barrier to adoption of formal financial services (e.g., Cole et al. 2013). And because the counselors accompanied respondents to the bank, doing so may have increased the respondents' level of ease in dealing with a formal institution.

As most subjects in our study were illiterate, the financial counselor provided households with assistance throughout all steps in the process of opening a bank savings account, from gathering the required documents, to filling out applications, and to accompaniment to the bank branch. As a consequence, we are ultimately unable to isolate the specific mechanism at play, and in particular, to assess the merits of simplifying the bank account opening requirements. We believe these are excellent empirical questions for future research.

2.5 Conclusion

This paper studies a large-scale field experiment among urban households in India to highlight the limitations of financial education and identify important complements that can enable financial education to successfully improve financial behavior. Specifically, we find that financial education alone improves financial awareness and attitudes but falls short of improving longer term behavioral outcomes on budgeting, savings, and borrowing. In comparison, the addition of individually tailored interventions in the form of financial goal setting and particularly financial counseling are more successful in helping individuals circumvent behavioral and cognitive constraints.

Taken together, our findings suggest that financial education can yield significant improvements in financial knowledge and behavior when sufficient attention is paid to the delivery model. Moreover, our results suggest that traditional classroom-based financial education alone has limited ability to affect long-term financial behavior, but adding more personalized and motivational complements can enable such outcomes. An important avenue for future research is to carefully examine the trajectory of effects over a longer time horizon as the impacts of financial education, goal setting, and financial counseling may sustain differently over time.

Table 2.1: Sample Size and Experimental Design

Panel A. Sample Size per Wave

Wave	Sample Size
1	279
2	422
3	312
4	315
Total	1328

Panel B. Experimental Design: Financial Education and Pay for Performance

Financial Education Videos	Pay for Performance	N	% of Sample
No	No	218	16
No	Yes	224	17
Yes	No	445	34
Yes	Yes	441	33

Panel C. Experimental Design: Financial Education and Additional Treatments

Financial Education Videos	Counseling	Goal Setting	N	% of Sample
No	No	No	442	33
Yes	No	No	232	17
Yes	No	Yes	209	16
Yes	Yes	No	215	16
Yes	Yes	Yes	230	17

Notes: This table describes the sample size and experimental design. The study was conducted in four waves and Panel A describes the number of respondents in each wave of the study. Panel B and Panel C describe the experiment design and randomization across the various treatments.

Table 2.2: Baseline Summary Statistics

	Median	Mean	Standard Deviation	Test of Joint Equality of Means Across All Treatments (F-test p-value)
Household characteristics				
Household size	6.00	5.85	2.47	0.711
Household monthly income (Rs.)	5900.00	7017.48	5635.51	0.164
Household monthly income per capita (Rs.)	1050.00	1272.96	922.26	0.121
Household has phone		0.84		0.361
Household has non-farm enterprise		0.26		0.517
Household has water connection		0.77		0.813
Respondent characteristics				
Female		0.58		
Age	38.00	38.56	9.07	0.367
Married		0.98		0.503
Hindu		0.82		0.866
Completed elementary school		0.47		0.339
Completed secondary school		0.04		0.830
Saath MFI client		0.48		
Math score (out of 8)	5.00	4.73	2.03	0.788
Financial knowledge score (out of 3)	2.00	1.61	0.62	0.215
Has hard time saving (self-report)		0.94		0.551
Interested in financial matters (self-report)		0.87		0.460
Monthly discount rate	0.14	1.52	4.72	0.087 *
Inconsistent time preferences		0.48		0.809
Risk averse		0.18		0.934

Notes: This table provides baseline summary statistics for our sample which consists of urban poor households in Ahmedabad, India. Column (4) reports the p-value of the F-test of joint significance across all treatment coefficients in regressions of the baseline characteristics on treatment dummies. The four treatments are i) financial education video only, ii) financial education video and goal setting, iii) financial education video and counseling, and iv) financial education video, goal setting and counseling. Column (4) regressions control for strata dummies where a strata is defined by gender, location and whether the household was an MFI client. Standard errors are clustered at the wave-class level. * indicates statistical significance at the 10% level.

Table 2.3: Short-term Impact on Financial Knowledge

	Aggregate Measure of Financial Numeracy (1)	Aggregate Measure of Financial Awareness (2)	Aggregate Measure of Financial Attitudes (3)
Financial Education	-0.008	0.072***	0.082**
	(0.018)	(0.017)	(0.034)
Pay for Performance	0.001	0.004	-0.017
	(0.019)	(0.021)	(0.049)
Interaction of Financial Education and Pay for Performance	0.006	0.011	0.001
	(0.024)	(0.024)	(0.051)
R-squared	0.186	0.177	0.208
Number of Observations	1256	993	591
Mean of Dependent Variable in Control Group	0.646	0.691	0.800
F-test p-value: Financial Education + Interaction = 0	0.896	0.000	0.013

Notes: This table presents regression results on short-term impacts from a survey conducted three weeks after the conclusion of the financial education program. The table shows intention-to-treat effects. The dependent variables are aggregate measures of financial knowledge in three dimensions: numeracy, awareness, and attitudes. Regression results for individual questions are presented in Appendix Tables 1-3. *Financial Education* is a dummy equal to 1 for an individual who was invited to the financial education treatment. *Pay for Performance* is an orthogonal treatment and is a dummy equal to 1 for an individual who was offered a monetary incentive for correct answers to financial knowledge questions. Results are reported with robust standard errors clustered at the wave-class level. All regressions include monthly discount rate at baseline as well as strata dummies, where strata are defined by gender, chali (neighborhood), and microfinance borrower status. *** indicates statistical significance at the 1% level, ** at the 5% level, * at the 10% level.

Table 2.4: Longer-term Impact on Financial Knowledge

	Aggregate Measure of Financial Numeracy (1)	Aggregate Measure of Financial Awareness (2)	Aggregate Measure of Financial Attitudes (3)
Financial Education	-0.021	0.104***	0.095***
	(0.028)	(0.020)	(0.024)
Pay for Performance	-0.029	-0.025*	-0.025
	(0.043)	(0.014)	(0.022)
Interaction of Financial Education and Pay for Performance	0.031	0.051**	0.024
	(0.050)	(0.020)	(0.028)
R-squared	0.152	0.216	0.203
Number of Observations	972	972	972
Mean of Dependent Variable in Control Group	0.720	0.682	0.734
F-test p-value: Financial Education + Interaction = 0	0.779	0.000	0.000

Notes: This table presents regression results on longer-term impacts from an endline survey conducted ten months after the conclusion of the financial education program. The table shows intention-to-treat effects. The dependent variables are aggregate measures of financial knowledge in three dimensions:numeracy, awareness, and attitudes. Regression results for individual questions are presented in Appendix Tables 1-3. *Financial Education* is a dummy equal to 1 for an individual who was invited to the financial education treatment. *Pay for Performance* is an orthogonal treatment and is a dummy equal to 1 for an individual who was offered a monetary incentive for correct answers to financial knowledge questions. Results are reported with robust standard errors clustered at the wave-class level. All regressions include monthly discount rate at baseline as well as strata dummies, where strata are defined by gender, chali (neighborhood), and microfinance borrower status. *** indicates statistical significance at the 1% level, ** at the 5% level, * at the 10% level.

Table 2.5: Household Budgeting

	Believes budgeting is helpful		Has tried making a budget in last 6 months		Makes a regular monthly budget	
	(1)	(2)	(3)	(4)	(5)	(6)
Any Treatment	0.221***		0.275***		0.028**	
	(0.029)		(0.027)		(0.012)	
Financial Education Only		0.168***		0.126***		0.026
		(0.040)		(0.039)		(0.018)
Financial Education & Goal Setting		0.241***		0.158***		0.003
		(0.038)		(0.041)		(0.023)
Financial Education & Financial Counseling		0.214***		0.385***		0.036*
		(0.036)		(0.040)		(0.020)
All Three Treatments		0.262***		0.433***		0.048*
		(0.038)		(0.039)		(0.025)
R-squared	0.247	0.252	0.244	0.294	0.265	0.267
Number of Observations	1235	1235	1235	1235	1235	1235
Mean of Dependent Variable in Control Group	0.602	0.602	0.194	0.194	0.065	0.065
F-test p-value: Financial Education & Goal Setting = Financial Education		0.102		0.499		0.363
F-test p-value: Financial Education & Financial Counseling = Financial Education		0.260		0.000		0.689
F-test p-value: All Three Treatments = Financial Education		0.026		0.000		0.421

Notes: This table presents regression results on household budgeting from an endline survey conducted ten months after the conclusion of the financial education program. The sample consists of respondents from all four waves of the study and the table shows intention-to-treat effects. *Any Treatment* is a dummy equal to 1 for an indivdual who received any financial education treatment. *Financial Education Only* is a dummy equal to 1 for an individual who was invited to the financial education classes, but did not receive either financial counseling or goal setting. *Financial Education and Goal Setting* is a dummy equal to 1 for an individual who received the financial education and goal setting treatments, but not the financial counseling treatment. *Financial Education and Financial Counseling* is a dummy equal to 1 for an individual who received the financial education and counseling treatments, but not the goal setting treatments. *All Three Treatments* is a dummy equal to 1 for an individual who received all three financial education, financial counseling, and goal setting treatments. Results are reported with robust standard errors clustered at the wave-class level. All regressions include monthly discount rate at baseline as well as strata dummies, where strata are defined by gender, chali (neighborhood), and microfinance borrower status. *** indicates statistical significance at the 1% level, ** at the 5% level, * at the 10% level.

Table 2.6: Household Savings

	Informal savings		Formal bank savings account		Investments in fixed or recurring deposits	
	(1)	(2)	(3)	(4)	(5)	(6)
Any Treatment	0.036***		0.082***		0.022	
	(0.012)		(0.022)		(0.014)	
Financial Education Only		-0.000		0.023		0.006
		(0.020)		(0.030)		(0.021)
Financial Education & Goal Setting		0.062**		0.082**		0.020
		(0.025)		(0.039)		(0.020)
Financial Education & Financial Counseling		0.015		0.132***		0.022
		(0.023)		(0.040)		(0.022)
All Three Treatments		0.070***		0.095**		0.041*
		(0.020)		(0.038)		(0.022)
R-squared	0.184	0.190	0.207	0.211	0.133	0.135
Number of Observations	1235	1235	1235	1235	1235	1235
Mean of Dependent Variable in Control Group	0.079	0.079	0.293	0.293	0.043	0.043
F-test p-value: Financial Education & Goal Setting = Financial Education		0.060		0.194		0.588
F-test p-value: Financial Education & Financial Counseling = Financial Education		0.591		0.021		0.587
F-test p-value: All Three Treatments = Financial Education		0.014		0.105		0.228

Notes: This table presents regression results on household savings from an endline survey conducted ten months after the conclusion of the financial education program. The sample consists of respondents from all four waves of the study and the table shows intention-to-treat effects. *Any Treatment* is a dummy equal to 1 for an indivdual who received any financial education treatment. *Financial Education Only* is a dummy equal to 1 for an individual who was invited to the financial education classes, but did not receive either financial counseling or goal setting. *Financial Education and Goal Setting* is a dummy equal to 1 for an individual who received the financial education and goal setting treatments, but not the financial counseling treatment. *Financial Education and Financial Counseling* is a dummy equal to 1 for an individual who received the financial education and counseling treatments, but not the goal setting treatments. *All Three Treatments* is a dummy equal to 1 for an individual who received all three financial education, financial counseling, and goal setting treatments. Results are reported with robust standard errors clustered at the wave-class level. All regressions include monthly discount rate at baseline as well as strata dummies, where strata are defined by gender, chali (neighborhood), and microfinance borrower status. *** indicates statistical significance at the 1% level, ** at the 5% level, * at the 10% level.

Table 2.7: Household Borrowing

	Has outstanding loan		Plans on taking a loan in next two years		Knows details of loan terms		Loan purpose: Business, education or purchase of durable goods		Loan purpose: Unforeseen expenses		Loan purpose: Repay other debt	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Any Treatment	0.025		-0.035		0.107**		0.056		-0.016		-0.007	
	(0.030)		(0.028)		(0.046)		(0.046)		(0.010)		(0.031)	
Financial Education Only		0.024		-0.026		0.099		0.094		0.001		-0.051
		(0.038)		(0.041)		(0.066)		(0.070)		(0.020)		(0.046)
Financial Education & Goal Setting		0.005		-0.064*		0.057		-0.018		-0.025**		-0.001
		(0.043)		(0.033)		(0.066)		(0.073)		(0.012)		(0.043)
Financial Education & Financial Counseling		0.041		-0.027		0.103*		0.118*		-0.020*		-0.014
		(0.047)		(0.038)		(0.057)		(0.063)		(0.011)		(0.041)
All Three Treatments		0.029		-0.023		0.169**		0.023		-0.018*		0.037
		(0.043)		(0.045)		(0.065)		(0.062)		(0.011)		(0.044)
R-squared	0.212	0.212	0.136	0.136	0.341	0.346	0.282	0.290	0.273	0.277	0.281	0.287
Number of Observations	1235	1235	1235	1235	404	404	536	536	536	536	536	536
Mean of Dependent Variable in Control Group	0.619	0.619	0.293	0.293	0.698	0.698	0.320	0.320	0.023	0.023	0.110	0.110
F-test p-value: Financial Education & Goal Setting = Financial Education		0.662		0.388		0.549		0.231		0.190		0.303
F-test p-value: Financial Education & Financial Counseling = Financial Education		0.747		0.966		0.957		0.777		0.280		0.412
F-test p-value: All Three Treatments = Financial Education		0.915		0.961		0.297		0.381		0.323		0.160

Notes: This table presents regression results on household borrowing from an endline survey conducted ten months after the conclusion of the financial education program. The sample in columns 1 to 4 consists of respondents from all four waves of the study and shows intention-to-treat effects. The sample in columns 5 to 12 consists of respondents from all four waves of the study who took out a loan since the conclusion of the financial education program. *Any Treatment* is a dummy equal to 1 for an indivdual who received any financial education treatment. *Financial Education Only* is a dummy equal to 1 for an individual who was invited to the financial education classes, but did not receive either financial counseling or goal setting. *Financial Education and Goal Setting* is a dummy equal to 1 for an individual who received the financial education a nd goal setting treatments, but not the financial counseling treatment. *Financial Education and Financial Counseling* is a dummy equal to 1 for an individual who received the financial education and counseling treatments, but not the goal setting treatments. *All Three Treatments* is a dummy equal to 1 for an individual who received all three financial education, financial counseling, and goal setting treatments. Results are reported with robust standard errors clustered at the wave-class level. All regressions include monthly discount rate at baseline as well as strata dummies, where strata are defined by gender, chali (neighborhood), and microfinance borrower status. *** indicates statistical significance at the 1% level, ** at the 5% level, * at the 10% level.

Table 2.8: Household Insurance

	Bought health insurance in last 6 months		Bought life insurance in last 6 months		Bought debt insurance in last 6 months	
	(1)	(2)	(3)	(4)	(5)	(6)
Any Treatment	-0.000		0.018		0.002	
	(0.009)		(0.011)		(0.003)	
Financial Education Only		-0.003		-0.004		0.009
		(0.011)		(0.015)		(0.007)
Financial Education & Goal Setting		0.000		-0.011		0.002
		(0.011)		(0.015)		(0.005)
Financial Education & Financial Counseling		0.006		0.033		-0.003
		(0.010)		(0.022)		(0.003)
All Three Treatments		-0.004		0.054**		-0.002
		(0.014)		(0.022)		(0.003)
R-squared	0.138	0.138	0.141	0.151	0.065	0.070
Number of Observations	1235	1235	1235	1235	1235	1235
Mean of Dependent Variable in Control Group	0.014	0.014	0.034	0.034	0.002	0.002
F-test p-value: Financial Education & Goal Setting = Financial Education		0.804		0.667		0.462
F-test p-value: Financial Education & Financial Counseling = Financial Education		0.507		0.112		0.122
F-test p-value: All Three Treatments = Financial Education		0.937		0.011		0.144

Notes: This table presents regression results on household insurance from an endline survey conducted ten months after the conclusion of the financial education program. The sample consists of respondents from all four waves of the study and the table shows intention-to-treat effects. *Any Treatment* is a dummy equal to 1 for an indivdual who received any financial education treatment. *Financial Education Only* is a dummy equal to 1 for an individual who was invited to the financial education classes, but did not receive either financial counseling or goal setting. *Financial Education and Goal Setting* is a dummy equal to 1 for an individual who received the financial education and goal setting treatments, but not the financial counseling treatment. *Financial Education and Financial Counseling* is a dummy equal to 1 for an individual who received the financial education and counseling treatments, but not the goal setting treatments. *All Three Treatments* is a dummy equal to 1 for an individual who received all three financial education, financial counseling, and goal setting treatments. Results are reported with robust standard errors clustered at the wave-class level. All regressions include monthly discount rate at baseline as well as strata dummies, where strata are defined by gender, chali (neighborhood), and microfinance borrower status. *** indicates statistical significance at the 1% level, ** at the 5% level, * at the 10% level.

Chapter 3

Rainfall Shocks and Food Security in Rural India

3.1 Introduction

Food security remains one of the most pressing economic development challenges in the world today. Recent estimates suggest that in 2013, almost 850 million people across the globe experienced chronic hunger, with the most vulnerable individuals lacking the necessary nourishment for normal growth and physical development (FAO, IFAD and WFP 2013). The causes of hunger are multidimensional, including shortages in food supply, insufficient purchasing power, and the inadequate use of food at the household level. Yet all of these causes contribute to one critical consequence: a situation wherein food insecure individuals cannot live an active and healthy life, thus preventing them from contributing fully to the socio-economic development of their households, their local communities, their nations, and the global stage.

At the same time, changing weather and climate patterns loom large over the existing challenges to achieving food security; the frequency, intensity, and duration of droughts are expected to increase around the world in the coming years, posing enormous threats to both global and local food systems.[1] To develop effective and well-targeted policies for mitigating the effects of climate change in the future, it will be critical to understand the micro-dynamics of food security along its multiple dimensions, from food production, to households' ability to acquire food, and to households' diet. Nevertheless, most studies concentrate primarily on one particular aspect—agricultural production—and hence provide only fragmentary evidence on the implications of droughts on food security among poor households.

While a large number of studies have established that low precipitation reduces agricultural yields, in this paper, I examine two other facets of food security that have received much less attention in the existing literature: food utilization (e.g., nutrition) and food access (e.g., prices, income). Although conventional intuition suggests that droughts may negatively im-

[1]For example, see Dai (2013) on increasing drought under global warming and Parry et al. (2004) on climate change and global food production.

pact these two dimensions, it is not immediately clear that such harmful effects may materialize since trade, storage, or savings may offset the effects of a dry rainfall shock. To this end, I focus on the following three questions. First, how do negative rainfall shocks affect household food expenditures, diet, and macronutrient intake? Second, what role do prices play as a channel for these effects? And finally, in what ways do impacts differ between various groups of households (i.e., agricultural vs. non-agricultural, marginalized vs. non-marginalized)?

I investigate these questions in the context of rural India, an environment home to much of the country's food insecure population and where agriculture continues to be largely rain-fed. Rural India provides an interesting empirical setting for this study not only because the vast majority of households depend on agriculture for their primary livelihood, but also due to the tremendous variability of rainfall patterns across both time and space. Moreover, the Indian setting provides remarkably rich microdata on household consumption through the National Sample Survey (NSS). With over 300 food items, these surveys include both market purchases and home production; these data thus provide a relatively exhaustive picture of household consumption, allowing for estimates of a households' total caloric, protein, and fat intake.

As several previous studies have suggested that more rainfall in the Indian setting tends to be beneficial (e.g., Jayachandran 2006), food security becomes a more relevant concern during times of drought. Hence, throughout this paper, I focus on the effects of a "dry shock," defined as the absolute deviation of rainfall, in meters per year, *below* the district's long-run mean. The results show substantial negative effects of low precipitation on household per capita food consumption and nutrition. In particular, I find that a one meter dry shock is associated with 5.6% less household per capita food expenditure, significant at the 1% level.[2] This lower food expenditure likewise corresponds with a decline in household per capita intake of calories and fat of 2.7% and 3.7%, respectively.

To better understand the underlying factors driving the decline in aggregate expenditure and nutrition, I examine the impacts of the dry shock on consumption of each of the following food groups: (1) cereals; (2) pulses, nuts, and oilseeds; (3) vegetables and fruits; (4) meat, fish, and dairy (including seafood, eggs, and milk); (5) sugar, honey, oils, and fats (including cooking oils, butter, and ghee); and (6) miscellaneous food and beverages (e.g., bottled drinks, biscuits). Notably, the disaggregated results show adverse effects of the dry shock on food expenditure essentially across the board. Low precipitation levels are also associated with a decrease in all per capita household nutrition outcomes I consider (i.e., caloric, protein, and fat consumption) across all of the above six categories.

Having established significant negative impacts of low rainfall on household consumption and nutrition, I next investigate the importance of prices as a channel for these effects. Strikingly, in contrast to the typical expectation that prices increase during droughts because of the decline in yields, I am unable to reject the null hypothesis that the effect of the dry shock on prices is zero for almost all food items I study. Hence, the negative impacts of drought on food utilization more likely operate through income rather than prices. The

[2] A one meter deviation of annual rainfall below its long-term mean is roughly equivalent to a little over one standard deviation.

null price effects I observe are also consistent with several previous studies in India (e.g., Zimmermann 2017; Blakeslee and Fishman 2017), which have argued that better market integration and storage have made prices less sensitive to weather conditions.

In line with the important role of storage, the results also provide some suggestive evidence on the heterogeneity in the impact of low rainfall on food prices, depending on the product's perishability. In particular, a one meter dry shock is associated with a statistically significant 3.7% increase in the price of potato, an item that requires a cold facility for long-term storage. In contrast, the effect of the dry shock is neither economically nor statistically significant for pulses, a product with a long shelf-life. Nevertheless, these null effects on prices are not very precisely estimated: the large standard errors preclude any rigorous conclusion as price data from developing countries tend to be quite noisy, with India being no exception. This paper thus underscores the need for better data on prices in developing countries for both for research and policy analysis.

In the final part of the paper, I explore the heterogeneity in the effects of drought on food consumption based on households' main source of income, as one might expect the effects to be larger among households who depend on agriculture. Indeed, I find that for agricultural laborers, a one meter dry shock is associated with a significant decline in food expenditure (−7.7%) and calories (−3.9%) per capita. However, self-employed agricultural households do not experience any such negative effects in the short-run, potentially because such households also have higher savings or access to loans as they own large plots of land. Lastly, I compare the effects of the dry shock between traditionally marginalized and non-marginalized households,[3] and the results show no statistically significant differences between the two. Hence, the broad negative impacts of the dry shock appear to be similar across these social groups.

This paper contributes directly to both the literature on the economic consequences of rainfall shocks as well as the policy discourse on food security. Previous studies have considered the impact of weather on, among others, agricultural production, wages, income, industrial productivity, conflict, and economic growth (see Dell, Jones, and Olken (2014) for a review). This study expands upon the existing knowledge base by considering how the negative impacts of drought on availability and access to food ultimately translate into impacts on food utilization among rural households. Overall, I find that droughts are associated with significant declines not only in household food expenditure, but likewise in caloric and macronutrient intake. These findings provide insights for policy makers towards designing targeted nutrition programs, especially in light of growing concerns on food security amid a changing climate.

[3] As will be explained in Section 3.5, I define the marginalized group as Scheduled Caste, Scheduled Tribe, and Other Backward Classes.

3.2 Background and Literature Review

Food security is a complex and multi-faceted concept. Although many interpretations of the term "food security" exist, in this paper, I adopt the widely-used definition proposed by the FAO (see FAO, IFAD and WFP 2013, p. 50):

> **Food security.** A situation that exists when all people, at all times, have physical, social and economic access to sufficient, safe and nutritious food that meets their dietary needs and food preferences for an active and healthy life.
>
> **Food insecurity.** A situation that exists when people lack secure access to sufficient amounts of safe and nutritious food for normal growth and development and an active and healthy life. It may be caused by the unavailability of food, insufficient purchasing power, inappropriate distribution or inadequate use of food at the household level.

This characterization highlights four dimensions of food (in)security: (1) *availability*, as determined by food production, stock levels, net trade, and food aid; (2) *access*, referring to economic, social and physical factors such as income, purchasing power, and market infrastructure; (3) *utilization*, concerning diet quality, diet diversity, and the nutritional aspects of food consumption; and (4) *stability*, encompassing the ability of individuals, households, and communities to cope with negative shocks, for example, due to low levels of rainfall.

A large literature ranging from the physical to the social sciences investigates the impacts of drought on the above four different aspects of food security. Nevertheless, existing research remains uneven and highly skewed because the vast majority of studies primarily concentrates on food *availability*. For example, in a bibliographical analysis of peer-reviewed journal publications since 1990 on food security and climate variability, Wheeler and Braun (2013) find that 70% of all publications considered food availability. In contrast, the access, utilization, and stability components constituted merely 11.9, 13.9, and 4.2% of total publications, respectively. Current knowledge therefore lacks coverage across all four pillars of food security, and as a consequence, the extant literature provides only incomplete evidence on the implications of rainfall shocks on household welfare.

While the numerous studies on availability cannot be adequately summarized in this paper, the broad-scale adverse effects of low rainfall levels on agricultural production are generally well-understood. Droughts cause significant reductions in yield across rain-fed agricultural systems around the globe, although the magnitude of the effects may differ due to differences in agricultural technologies. In South Africa, Akpalu, Rashid, and Ringler (2011) find that 10% lower rainfall corresponds to 2.57% lower maize yields, while recent estimates by Zipper, Qiu, and Kucharik (2016) for the US suggest that droughts are associated with 13% of maize and soybean yield variability over the last fifty years. Furthermore, consistent with many other analyses of the weather-agriculture relationship in India (e.g., Pathania 2007; Guiteras 2009), Blakeslee and Fishman (2017) report that negative rainfall shocks lead to a 20.8% decline in agricultural output as well as a 9% decrease in yields during the dry season.

Closely linked with research on crop production and food availability is the literature on rainfall shocks and food *access*. Several studies related to this area, particularly in economic development, focus on one particular set of outcomes: wages and household income. To give an example, in the context of India, Jayachandran (2006) shows that low levels of precipitation lead to low agricultural productivity and subsequently, low agricultural wages. Similarly, an early study by Rosenzweig and Binswanger (1993) argues that rainfall variability results in significantly lower average income among Indian farmers. Beyond these impacts on rural livelihoods, however, comprehensive empirical evidence on food access remains relatively scarce. The effects of rainfall on other determinants of *economic* access, such as prices, continue to be less frequently discussed in the literature, as are elements of *social* access, for example among minority groups.

A thorough understanding of food security likewise requires probing into food *utilization*, including household diet and nutritional well-being. In the *Declaration of the World Food Summit on Food Security* in 2009, policy makers heralded nutrition as an integral component of food security. Yet this global affirmation stands in stark contrast with the existing literature, of which few studies consider food utilization.[4] While many studies demonstrate negative effects of low precipitation on health outcomes—and notably, propose nutrition as a potential mechanism (e.g., Maccini and Yang 2009)—most empirical evidence do not measure the nutrition channel directly.[5] Investigating the effect of rainfall variability on food utilization thus represents a critical issue not only for academic researchers, but also for policy makers, governments, and stakeholders worldwide.

This paper aims to fill an important research and policy knowledge gap by focusing on the dimensions of food access and utilization. Specifically, I explore three related questions that so far remain fairly understudied in the literature on rainfall shocks and their impacts on household food security. First, how do low rainfall levels impact household food consumption as well as its composition and nutritional value? Second, what is the relative importance of prices and income for these effects? And third, to what extent do the effects vary across different types of households (e.g., agricultural vs. non-agricultural households, marginalized vs. non-marginalized groups)? These research questions likewise address the fourth pillar of food security, *stability*, as they provide insights on ensuring stable food access and utilization among the poor especially during times of drought.

To investigate the above three research questions, I turn the context of households in rural India. Rural India represents a particularly interesting study setting because rain-fed agriculture accounts for more than 50% of the country's net sown area and 40% of total food production; thus, low levels of precipitation may have significant effects on rural households'

[4]As an example, in an overview of climate change and food security, Lobell and Burke (2010, p. 26) write that "[t]he utilization component of food security is perhaps its murkiest and least well-studied aspect."

[5]A related paper is by Hou (2010) who examines the impact of drought on food consumption and caloric intake. The author finds that droughts decrease total food expenditure but increase total calorie availability because households move away from expensive calories (e.g., vegetables, fruits, animal products) and towards cheaper calories (e.g., grains).

food consumption.[6] Additionally, almost 70% of the country's population live in rural areas, where households derive their earnings primarily from the agricultural sector and where much of the country's food insecure population reside.[7] Finally, rainfall patterns in India vary substantially across both districts and years, as shown in Table 3.1. I discuss in more detail the rainfall data, the measure of rainfall shocks I employ as well as its variation over space and time in the following section.

3.3 Data Sources and Summary Statistics

Examining the effects of rainfall shocks on food utilization and access requires detailed data on precipitation, agricultural production, household consumption, and food prices. In this section, I describe the data sources I use in this study, which consist of both district-level panel data and a repeated cross-section of household expenditure surveys. Moreover, I provide summary statistics of these data to illustrate the variation in rainfall across India as well as the typical everyday diet among rural Indian households.

Data Sources

Rainfall Data. I use precipitation data collected by Willmott and Matsuura (2015) to identify rainfall patterns in each district. Compiled from several sources such as the Global History Climatology Network Dataset and India's National Center for Atmospheric Research, the data contain monthly total precipitation for the years 1900 to 2014 for all of India, gridded at a resolution of 0.5 degrees (approximately 30 miles). I follow several previous studies to match these grid-level data to districts.[8] In particular, I identify the grid point that is closest to a district's centroid and assign that grid point's rainfall to the said district.

Aggregating these monthly precipitation data to the annual level, I define a "dry shock" in a given district to be the absolute deviation of rainfall, in meters per year, below the district's long-run mean over four decades (i.e., between 1970-2012).[9] This definition of a dry rainfall shock at the district-level is similar to that in existing studies such as Sekhri and Storeygard (2014). Notably, while some district borders have changed over time, I use boundaries based on the 2001 Census of India throughout this study. Hence, I consolidate new districts that were formed between 2001 to 2012 back to their parent district.

National Sample Survey (NSS). For my main outcome variables, I employ detailed household data from the NSS Consumption and Expenditure Surveys. These surveys ask

[6] These statistics come from the Government of India's Ministry of Agriculture. See `http://agricoop.nic.in/divisiontype/rainfed-farming-system`

[7] The estimate of the percentage of rural population is based on the figure for the year 2015 from the World Bank. See `http://data.worldbank.org/indicator/SP.RUR.TOTL.ZS`

[8] For example, see Cole, Healy, and Werker (2012), Iyer and Topalova (2014), and Shah and Steinberg (2017), among others.

[9] Since my study period ends in 2012, I chose the four decade period starting from 1970 until 2012.

households to report total expenditure and quantities consumed over the last 30 days for an extensive range of food items such as cereals, pulses, vegetables, fruits, and meat, among others. The survey includes market purchases, in-kind wage payments, and home production, thus providing a full image of household food expenditure. Taking all possible source of food expenditure into account is particularly important in the rural Indian setting, since many individuals may be rural laborers who receive in-kind wages or subsistence farmers.[10]

Central to understanding the effects of rainfall shocks on food utilization is a measure of household nutritional intake. To this end, I employ the nutrition chart based on Gopalan, Sastri, and Balasubramanian (1991) and reported in various NSS publications (e.g., NSS 2001, 2007, 2012, 2014). The nutrition chart reports the energy, protein, and fat content per unit of weight of different foods items in the NSS survey. For example, it indicates that one kilogram of potato typically has 967 calories, 17 grams of protein, and 1 gram of fat. I therefore obtain the nutritional composition of households' food expenditure by multiplying the quantities recorded in the survey with the conversion factors specified in the nutrition chart.

Although using the nutrition chart to convert households' survey responses to caloric, protein, and fat intake is a relatively straightforward accounting exercise, food expenditure reported in the NSS surveys may not reflect all household members' consumption for two reasons. First, individuals may consume meals outside the home, for instance through employers or schools. The food content of these meals will be difficult to measure and are not captured in the household surveys. Second, households may provide meals to non-household members such as neighbors or helpers. Expenditure for these meals are included in the survey responses but do not contribute to the household's nutritional intake.[11]

I address the above issues by using an "adjustment factor" defined as $(N_h + N_a) \div (N_h + N_o)$, as in NSS (2014) and Eli and Li (2017). Here, N_h is the total number of meals taken by household members at home during the last 30 days, while N_a and N_o are the analogous figures for meals away from home and meals given to others (i.e., non-household members), respectively. This adjustment factor implicitly assumes that meals given and received enter symmetrically and that the nutritional value of both types of meals are directly proportional with a meal at home.[12] Importantly, household-level data on the number of meals taken by household members at home, the number of meals taken by household members away from home, and the number of meals given to others are all available through the NSS surveys.[13]

[10]However, the quantity and value of food consumption from each source is not available in the NSS surveys used in this study, with the exception of NSS Rounds 61, 66, and 68 which asks households to report both total consumption and consumption from home production.

[11]Actual nutritional intake my also depend on how households prepare and cook food. However, it is not possible to make adjustments for cooking methods since this information is not available in the NSS data.

[12]Note that the adjustment factor is greater than one for households that are receiving meals away from home much more than giving meals to others. Similarly, it is less than one for households that are serving many more meals to non-household members than receiving meals away from home.

[13]In this paper, one reason why I use NSS Survey Rounds 60-64, 66, and 68 is because previous rounds do not collect information on the number of meals given by the household to non-household members. Hence, it is not possible to calculate the adjustment factor in other rounds of the NSS surveys.

Rural Price Collection (RPC) Survey. The RPC survey, implemented by India's National Sample Survey Organization, collects rural retail prices every month for a basket of goods, fielded in a fixed set of 603 markets spread over 26 states. The data are therefore at the market-month-product level. The data cover a large number of products consumed by rural households, including food items such as cereals, pulses, vegetables, and fruits. These data also form the basis for the Consumer Price Index for Agricultural and Rural Laborers, compiled and published by India's Ministry of Labor.

Two important issues must be noted regarding the quality of the RPC data. The first is that many observations are missing because some products are not always present in the data while others are indicated with zero prices. The second concern is that no data are available for the months between October 2007 to June 2009. To address these issues and to increase the signal-to-noise ratio, I collapse the data to the district-year level by taking the median price across markets for a given district and year. Furthermore, I drop all data for 2007 and 2009, as these years do not cover the full twelve months and lack data for either the fourth or first quarter of the year, which span an important agricultural season.

Crop Production Statistics. I obtain district-level measures of crop production from the Government of India's Ministry of Agriculture and Farmer Welfare, Directorate of Economics and Statistics. These data include information on total production and area planted across districts for several different types of crops such as grains, pulses, root vegetables, nuts, oilseeds, and sugarcane. Although the crop production statistics are publicly available online from 1997 onwards, I use data from 2003 to 2012 to correspond with the years covered in the NSS Consumption and Expenditure Surveys.

Census of India, 2001. I further augment the above datasets with the 2001 Census of India. This dataset contains district-level demographic variables, for example, population density, literacy rate, and unemployment rate. These data thus provide additional control variables in the regression analysis, which I explain in more detail in Section 3.4.

Summary Statistics

Given this study's objective of investigating the effects of low rainfall levels on household food utilization, in this section, I provide descriptive statistics on two important sets of variables. The first concerns annual precipitation, and in particular, my measure of a dry shock. Table 1 reports the mean and standard deviation (SD) of these variables. As can be seen in this table, there is substantial variation both across space and over time in total rainfall levels and the dry shock. In any given year, for example, the SD of total annual precipitation across districts lies between 0.78 and 0.98 meters, with a mean of around 1.2 to 1.4 meters. Furthermore, from 2004 and 2012, the percentage of districts experiencing a dry shock (i.e., total annual rainfall that is below the long-term mean) ranged from 36% to 81%.

The second set of key variables in this study involve consumption, diet composition, and nutrition among rural Indian households. For each NSS survey round, Table 3.2 shows averages, and in brackets, the SD, of monthly household per capita food expenditure (in nominal Rupees) as well as daily per capita intake of calories (in Kcal), protein (in grams), and fat (in grams), all adjusted using the factors described in the previous section. Total average household food expenditure in nominal terms has generally increased over time, and the same is true for fat consumption. Importantly, over the years, mean caloric intake lies between only 2,150 to 2,250 Kcals per day. Since 2,400 calories per capita per day is often cited as the "minimum requirement" in India and forms the basis of the country's poverty line threshold,[14] these statistics suggest that the average rural household fails to meet basic subsistence levels of energy intake.

To further understand the quality and composition of rural households' diet, I then disaggregate household food and macronutrient consumption into six different food groups, These food categories consist of the following: (1) cereals (such as rice, wheat, maize, millet, jowar, and barley); (2) pulses, nuts, and oilseeds (such as gram, urad, peas, arhar and tur, groundnut, and walnut); (3) vegetables and fruits (such as potato, onion, cauliflower, spinach, and bananas); (4) meat, fish, and dairy, (including prawns, crab, and other seafood; milk and milk products; eggs); (5) sugar, honey, oils, and fats (including cooking oils, margarine, butter, and ghee); and (6) miscellaneous food and beverages (such as coffee, tea, cooking spices, bottled drinks, biscuits, and sauces).[15]

Table 3.2 presents summary statistics for each of the above six food groups using the same set of expenditure and nutrition variables as before, but with the addition of the budget shares of each category in total food expenditure. Three interesting patterns stand out from this table. First, we observe that cereals represent a significant component of diet among rural households. In particular, across all NSS Survey years, the highest share of the average households' food budget is always devoted to cereals, at around 25 to 33% of total food spending. This large expenditure share is likewise reflected in calorie consumption. Indeed, for the typical rural household, the greatest source of calories again comes from cereals, which provides an average of 1,300 to 1,400 calories per person per day and amounts to more than 50% of total daily caloric intake.

Second, the biggest category of food expenditure, next to cereals, is meat, fish, and dairy. In this food group, households spend on average about one-fifth of their total food budget. Meat, fish, and dairy products, in turn, contribute substantial amounts to per capita intake of protein and fat. For instance, in the NSS Survey for 2011-12, about 17% (10.2 grams) of total protein and 27% (12.6 grams) of total fat intake come from meat, fish, and dairy. Notably, while this food category is the second largest source of protein for the average household, the majority of protein intake comes from cereals, providing almost 40 grams per day. Hence, although the protein content of one kilogram of cereals is much lower than one

[14] See for example, Deaton and Drèze (2009) and National Institute of Nutrition (2011).

[15] These categories of food as well as the composition of each food category comes from various NSS publications (see NSS 2001, 2007, 2012, 2014).

kilogram of meat, households still derive much of their protein from cereals given that they purchase and consume more of such products.

Lastly, Table 3.2 also demonstrates that sugar, honey, oils, and fats play an important role in supplying calories for rural households. Specifically, rural individuals obtain an average of about 300 calories per day from sugars and fats, and this category delivers many more calories than any other food group except for cereal. Cooking oils and ghee are likewise the most important contributors to daily per capita fat intake, providing almost twice as much grams of fat than meat, fish, and dairy in the 2011-2012 survey. At the same time, households spend relatively little buying sugary and fatty foods, with a budget share of only 12.4% and mean monthly per capita expenditure level of Rs. 87 (US$ 1.3) in the most recent survey. In this context, sugars, oils and fats therefore serve not only as conduits for increasing the taste and palatability of food, but also as inexpensive and rich sources of energy for rural households.

3.4 Empirical Method

I exploit random year-to-year deviations of precipitation from its long-run mean as a measure of exogenous local rainfall shocks. Throughout this paper, I focus on *negative* rainfall shocks in a given district—where rainfall is *below* the district's long-run mean—rather than positive shocks given that food insecurity is a much greater concern during times of drought. Indeed, several empirical studies have argued that more rainfall in the Indian context tends to generally be favorable (e.g., Jayachandran 2006; Duflo and Pande 2007).[16] My basic empirical framework then investigates how both district- and household-level outcomes respond to low levels of precipitation in a given year.

I take advantage of the spatial and temporal variation in rainfall across districts to estimate two sets of log-linear regressions. The first uses district-level panel data to estimate a fixed effects model given by

$$ln(y_{dt}) = \beta DryShock_{dt} + \gamma_d + \lambda_t + \delta_t \mathbf{X}_d + \epsilon_{dt}. \tag{3.1}$$

For a district d in year t, the left-hand side variable $ln(y_{dt})$ represents the natural logarithm of food security-related outcomes. These outcomes consist of indicators of food availability (i.e., agricultural yields for cereals, pulses, and other crops) as well as food access (i.e., the median price of a particular food item across markets in a given district-year).

The right-hand side of equation (3.1) includes the following variables: $DryShock_{dt}$, the absolute deviation of rainfall in a given district below its long-run mean (i.e., between 1970 and 2012) in meters per year; γ_d, district fixed effects to capture time-invariant district characteristics; λ_t, time fixed effects to control for changes over time that are common to all districts; and $\delta_t \mathbf{X}_d$, year-interacted district characteristics from the 2001 Census. The

[16] In the robustness checks to follow in Section 3.5, I examine the impact of a wet shock on household food consumption and nutritional intake. The results show no systematic effects, suggesting that food security is a much more pressing and relevant issue during a dry shock.

coefficient of interest in this equation is β. Since I implement a log-linear regression, $100 \cdot \hat{\beta}$ measures the average percentage change in outcomes when total annual rainfall in the district is one meter below its long-term mean (henceforth, a "one meter dry shock").

In comparison to equation (3.1) which considers outcomes at the district level, the second regression framework I employ examines the impacts of drought at the household level. To this end, I use the household-level analogue of equation (3.1) with the form

$$ln(y_{idt}) = \beta DryShock_{idt} + \gamma_d + \lambda_t + \delta_t \mathbf{X}_d + \alpha \mathbf{H}_{idt} + \epsilon_{idt}. \tag{3.2}$$

Here, the dependent variable $ln(y_{idt})$ represents the natural logarithm of per capita food expenditure or nutritional intake for household i in district d at time t. I estimate this equation by pooling together multiple rounds of the NSS consumption surveys, and the regression does not include household fixed effects as the NSS data is not a panel of households.

The basic features of equation (3.2) are similar to that in equation (3.1) but with two notable differences. First, while the $DryShock$ variable in the latter is measured at the district level, in the former, it is at the household level. All households in the NSS surveys are asked to report their food consumption and expenditure over the last 30 days, but because fieldwork is divided into four quarters of the year, the interviews happen at different dates across the full sample. Hence, to align the precipitation shocks with the timing of the survey, $DryShock_{idt}$ in equation (3.2) represents the absolute deviation from the district's long-term mean of total rainfall *in the past 12 months* prior to household i's survey date.

Second, in addition to controlling for year-interacted district characteristics $\delta_t \mathbf{X}_d$ as in equation (3.1), equation (3.2) also controls for *household* characteristics with the vector $\mathbf{H}_{idt}$. Since the outcome variables I consider in equation (3.2) pertain to food expenditures and nutritional intake, it will be critical to control for characteristics—for example, household gender-age composition—that may influence the types and quantities of food items that households consume. The vector $\mathbf{H}_{idt}$ enables me to account for such factors, as it consists of dummies for the household's religion, social group (i.e., ST, SC, or OBC), and the fraction of household members in different male/female age cells.[17]

Using equation (3.2), I explore the effects of a dry shock on household consumption both in total across all types of food as well as broken down by food groups. Since the regressors are expressed in logarithmic form, it is important to note that a small number of households report zero consumption for some categories of food, at a frequency of less than 2% across all households and survey years; the dependent variable will thus be undefined for these households. To address this issue, I adopt the same method as in Blakeslee and Fishman (2017) and Pakes and Griliches (1980): I replace the dependent variable with zero for all zero values of consumption and include a dummy variable in the regression for this data transformation.

Finally, standard errors in both equations (3.1) and (3.2) are clustered at the district-level. This approach is similar to that in previous studies on the economic effects of weather in India, such as Burgess et al. (2017) and Sekhri and Storeygard (2014), who have argued

[17]Following various NSS publications (e.g., NSS 2001, 2007, 2012, 2014), I use the following age bins: 0-1, 2-3, 4-5, 7-9, 10-12, 13-15, 16-19, 20-39, 40-49, 50-59, 60-69, and 70+.

the measurement errors are likely to be correlated within districts over time. Nevertheless, I explore robustness of the results to the potential spatial correlation of rainfall shocks by clustering at the state-year level and by implementing spatially correlated errors as modeled in Conley (1999). The results from these alternative standard errors show similar patterns, as will be discussed in the robustness checks in Section 3.5.

3.5 Empirical Results

Effects on Agricultural Yields

As a point of departure, I investigate the impacts of droughts on agricultural yields. This analysis therefore speaks directly to the *availability* dimension of food security, which concerns the supply side of food and the sufficiency of food production. Because the effects of drought on food availability have already been widely studied—and as explained earlier in Section 3.2, a large number of studies have demonstrated negative effects on agricultural production—in this section, I aim primarily to replicate the findings from the existing literature. Doing so likewise serves as a test to confirm the validity of my measure for a dry shock, which is a main regressor of interest throughout this study.

Table 3.3 shows results from estimating equation (3.1) for the outcome variable log agricultural yield. Here, agricultural yield is defined as production (measured in tonnes) divided by area planted (measured in hectares). In this table, I include five different crops, namely, two types of grains (rice and wheat), two types of pulses (gram and urad) as well as potato. I have chosen these particular crops because they possess important characteristics that are necessary for this empirical study: they contain the most comprehensive coverage in the crop production data both over time and across space, and moreover, they represent relatively substantial components of diet and food expenditure among rural Indian households.

Consistent with earlier studies on the relationship between weather and agricultural production in India, Table 3.3 highlights a stark negative effect of negative rainfall shocks on crop yields. In particular, the estimates suggest that when annual rainfall is one meter below the long-term average, rice and wheat yields fall by 31% and 20% respectively, significant at the 1% level. Similarly, during the same dry shock, yields for pulses see a decline of 24% for gram and 12% for urad, while potato yields decrease by 11%. These findings are in line with the fact that much of agriculture in rural India continues to be rain-fed. Additionally, the magnitude of these coefficients are quite similar to those in the extant literature, thus echoing previous results on the detrimental impacts of poor rainfall on food availability.

Yet knowledge about the ramifications of precipitation on crop yields, in and of itself, provides an inadequate depiction of food security during negative rainfall shocks. Although agricultural production is a critical component of food security, it nevertheless represents only a small piece of a big picture because food security is a wide-ranging issue with many interrelated concepts and complexities. Hence, for the rest of this paper, I investigate the impacts of drought on household outcomes that relate to two other pillars of food security—that is, utilization and access. These factors have thus far received only little attention in the ongoing research and policy discourse, but they remain critical to building a broader understanding of the repercussions of negative rainfall shocks on food security among poor households.

Effects on Food Consumption and Nutritional Intake

This section explores the impact of a negative rainfall shock on food *utilization*, particularly the diversity, quality, and nutritional value of households' day-to-day diet. My empirical analysis uses data from the NSS Surveys, which includes the quantity (e.g., in kilograms) and value (in Rupees) of consumption from both market purchases and home production, thus providing a complete picture of household food expenditure. In particular, I estimate regression equation (3.2) where the outcome variables are per capita food expenditure, caloric intake, protein consumption, and fat intake. All of these outcomes are measured in logarithm at the household level, and I examine them both in aggregate across all food groups as well as disaggregated by each food category.

Table 3.4 reports the results on aggregate food and macronutrient consumption. While the negative effects of droughts on agricultural production have been well established, the results of this table highlight a related but distinct finding: the lower crop yields due to a dry shock are likewise accompanied by lower household food expenditure and nutrition. I find that on average, a one meter dry shock corresponds to 5.6% less per capita spending on food, statistically significant at the 1% level. In addition, the nutritional value of food consumption falls, where the strongest effects of the dry shock can be seen in calories (-2.7%) and fat (-3.7%). The same dry shock negatively impacts protein intake as well with a magnitude of 1.9%, although this estimate is not statistically significant.

Next, I delve into the dynamics underlying the reduction in household's *aggregate* food consumption by *disaggregating* the effects into different groups of food. As explained earlier in Section 3.3, the six food categories I consider are the following: (1) cereals; (2) pulses, nuts, and oilseeds; (3) vegetables and fruits; (4) meat, fish, and dairy; (5) sugar, honey, oils, and fats; and (6) miscellaneous food and beverages. The latter group consists primarily of packaged and processed foods that are typically consumed by richer and/or more urban households (e.g., bottled drinks, fruit juices, biscuits, sauces) together with spices for cooking, tea, and coffee. Hence, for rural households, it is reasonable to assume that this category consists of food items with an income elasticity above one.

The regression results broken down by the above six food groups are reported in Table 3.5. The estimates in this table show a negative impact of low rainfall levels essentially across *all* food groups and likewise across *all* expenditure and nutrition outcomes. For cereals in particular, which is typically the biggest component of household consumption, I find that a one meter dry shock is associated with a decline in per capita expenditure of 5.2% as well as in both caloric and protein intake per capita of about 2%. Notably, the size of these coefficients are much smaller in comparison to that of other food groups. These results therefore suggests that rural households continue to rely heavily on cereals and cereal products for energy and protein even during times of drought.

Whereas the cereals food group exhibit the smallest effects, the largest impacts of the dry shock in percentage terms come from the category "miscellaneous food and beverages," for which a one meter dry shock corresponds to 7.6% less expenditure. This substantial negative effect is most likely because spending on food items in this category exhibits an income elasticity greater than one, so that households significantly reduce their expenditure for such high-income elastic goods when experiencing low levels of precipitation. Furthermore, the drought-induced decline in expenditure for miscellaneous food and beverages parallels with a considerable drop in nutritional intake of calories (−13.7%), protein (−10.6%), and especially fat (−20%), all of which are most sizable in magnitude across the six different food groups.

Following miscellaneous food and beverages, the second biggest effect on expenditure can be seen in the food group "sugar, honey, oils, and fat." In this category, a one meter dry shock is associated with 6.9% less spending, indicating that poor rainfall may lead rural households to substitute away from products that increase the palatability of food. But in addition to adding flavor, products such as cooking oil and ghee represent an important source of fat in rural households' diet, as described previously in Section 3.3. Indeed, the reduced expenditure in these food items translates to a 7.1 % decrease in consumption of fat per capita, significant at 1% level. Droughts may therefore lead to a considerable decline in the availability of fat, a macronutrient that provides twice as much energy than carbohydrates, particularly among households close to subsistence levels of food consumption.

Finally, I examine the impact of low levels of precipitation on the remaining three food groups, namely, pulses, nuts and oilseeds; vegetables and fruits; and meat, fish and dairy. In each of these categories, the dry shock is associated with a decline in expenditure as well as in nutritional intake of calories, protein, and fat, with particularly large effects for animal products. Although the coefficients on all of these outcomes are always negative, these estimates attain statistical significance in only one-third of all cases. This may be due to attenuation bias arising from substantial measurement error in the data, as respondents in the survey are asked to report their households' consumption values and quantities for over 600 items, all with a 30-day recall period.

Effects on Prices

Having established the negative impact of droughts on household food expenditure and nutrition, I next investigate the role of prices as a mechanism behind these effects. Food prices may significantly influence households' food *access*—that is, their ability to acquire food—through two avenues: consumption and production. On the consumption side, access to food depends in part on how much food costs, which is especially important for households who obtain their provisions primarily from market purchases, such as landless laborers. On the production side, any change in price directly impacts the agricultural income that farm households earn, particularly for those who are net producers of food products.

Table 3.6 reports estimates of equation (3.1) using the log rural *retail* prices of wheat, rice, gram, urad, potato, spinach, and banana as outcome variables. With the addition of the latter two items, Table 3.6 includes the same set of products as in Table 3.3, which demonstrated a significant decline in agricultural yields due to droughts. As before, I have selected the food items in Table 3.6 given their extensive spatial and temporal coverage in the retail price data as well as their relative importance in rural households' diet and food expenditure. All retail prices in this table come from the RPC survey data and are defined as the median price across surveyed markets in a given district-year.

Across almost all products, Table 3.6 shows a positive effect of low rainfall levels on prices. A one meter dry shock, for instance, corresponds to a 2% and 0.7% hike in the price of rice and urad, respectively. Yet these estimates are by and large not statistically significant, and I am unable to reject the hypothesis that the true population coefficient is zero. These results therefore suggest that prices may be unresponsive to precipitation in India, so that the negative impacts of drought on food consumption and nutrition more likely comes through *income* rather than prices. Furthermore, these null effects on prices indicate that the earnings of farm households from agricultural production may potentially fall, since crop yields decrease with no offsetting effect on income through a price increase.

The zero effects I find, however, are not very precisely estimated. Indeed, the relatively large standard errors in Table 3.6 make it difficult to draw strong conclusions, given that price data from developing countries tend to be very noisy. Nevertheless, several previous studies in India have similarly noted a lack of impact of rainfall shocks on prices. Specifically, using data on daily wholesale prices, Zimmermann (2017) reports that storage and re-optimization across space makes rice prices less sensitive to weather conditions in the country. Additionally, Blakeslee and Fishman (2017) find little within-year price variability across Indian districts, arguing that agricultural markets have become increasingly integrated over time due to advent of transport infrastructure such as railroads.

The price effects I observe are not only consistent with earlier studies, but also point to a relatively new insight. While Zimmermann (2017) focuses mainly on grains, I find suggestive evidence that the impact of rainfall on food prices may be heterogenous depending on the product's perishability. In line with the important role of storage, I find that for potatoes, which require a cold facility when stocked away for long periods of time, a one meter dry shock is associated with a price increase of 3.7%, statistically significant at the 5% level. The result for spinach, also a highly perishable good, is of similar magnitude at 3.2%. In contrast, the estimates on gram and urad, both of which have a long shelf life, are much smaller and very close to zero. These differences may thus arise potentially because pulses can be stored and transported much more easily than root and green vegetables, which spoil more quickly.

Heterogeneity of Effects among Households

This section investigates the heterogenous effects of negative precipitation shocks on food consumption and nutrition across the population, focusing on two sets of comparisons: (1) different subgroups of households based on their main source of income, and (2) socially marginalized versus non-marginalized households. The following discussion explains these analyses in more detail.

Effects by Household's Main Source of Income

Because the impacts of drought on food security depends not only on prices but also on how households earn their income, I explore the heterogeneity of effects of the dry shock across subsets of households according to the sector of their primary livelihood. As discussed previously in Section 3.5, the statistically insignificant effects of low precipitation levels on prices suggests that the negative effects on household food expenditure and nutrition most likely operate through a reduction in income. Consequently, one might also expect the impact of a dry shock to be larger in magnitude among households who rely heavily on agricultural earnings, such as unskilled farm laborers or self-employed farmers.

I divide households into the following five groups depending on their main income source, following the categories defined in the NSS Surveys: (1) self-employed in agriculture; (2) self-employed in non-agriculture; (3) laborers primarily in agriculture; (4) laborers who depend both on agriculture and non-agricultural work; and (5) other income sources. An important limitation with specifying these groups is that as per the NSS, households who do not belong in the first four groups automatically fall into the fifth category. Hence, this fifth category is somewhat difficult to interpret, as it includes salaried non-manual workers as well as households who depend primarily on pensions, remittances, and cash transfers.[18]

With the above caveat in mind, for these five group of households, Table 3.7 reports the impact of the dry shock on *total* monthly per capita expenditures, while Table 3.8 shows estimates on monthly per capita *food* expenditure and nutrition. Results from both tables demonstrate substantial heterogeneity across households and point to several notable patterns. First, consider households primarily self-employed in agriculture. For this subset, the average effect of the dry shock on both total expenditure (Table 3.7, Column 2) and food expenditure (Table 3.8, Column 1) are very close to zero and not statistically significant. These households likewise do not experience any significant change neither in caloric nor in nutrient intake of proteins and fat during droughts.

That the dry shock has little short-term impacts on food utilization among households self-employed in agriculture may be because almost all of such households are land-owners. Indeed, the data shows that among those self-employed in agriculture, 99% own any land. The median landholding within this group is 1.28 hectares, much larger than that for those *not* self-employed in agriculture, at 0.16 hectares. Despite the importance of rainfall for the income of self-employed farmers, these households most likely have either sufficient savings or borrowing capacity that allow them to smooth the income fluctuations caused by the dry shock in the short-run. The results therefore highlights the possibility that droughts may not always lead to lower consumption and nutrition, depending on the type of household.

Whereas negative rainfall shocks have no visible effects on self-employed agricultural households, the opposite is true among agricultural laborers who typically own very little land.[19] The estimates show that a one meter dry shock is associated with 4.6% less total

[18] See the discussion by Rawal (2014) for more information on the issues regarding the classification of rural labour households in the NSS Surveys.

[19] Among this group, the median land holding is 0.04 hectares

expenditure (Table 3.7, Column 4) and 7.7% less food expenditure (Table 3.8, Panel A, Column 3) among households who rely primarily on agricultural labor. Importantly, the lower spending on food likewise corresponds with 3.9% fewer calories per capita (Table 3.8, Panel B, Column 3). These findings are consistent with an argument whereby negative rainfall shocks reduce the demand for agricultural labor, thus putting a downward pressure on wages and employment (e.g., Jayachandran 2006). As a result, households who derive their main livelihood from agricultural labor earn less income, with severe negative implications for their food and energy intake.

Finally, I consider the effects on non-agricultural households: those self-employed in non-agriculture (e.g., carpenters) or who derive income from "other sources" (e.g., salaried non-manual laborers, remittances). For both types of households, the results in Table 3.7 show a significant negative effect of dry shocks on total expenditure. Similarly, both groups experience statistically significant declines in food consumption as well as in intake of calories, protein, and fat. Given data limitations, it is hard to precisely pinpoint the mechanisms at play behind these substantial negative effects. Nevertheless, these effects may potentially be due to a reduction in demand for non-agricultural labor due to general equilibrium effects. More comprehensive data is therefore necessary to investigate such channels, and it remains an interesting avenue for future research.

Effects by Social Group

While the previous analyses examined the *economic* aspect of food access by exploring the impacts of droughts on prices as well as on the consumption of households with different types of income sources, this section focuses on the *social* dimension of food access. In particular, I compare the effects of the dry shock on the consumption of marginalized and non-marginalized households. Here, "marginalized" is defined as those households from traditionally disadvantaged groups in India—that is, Scheduled Castes (SC), Scheduled Tribe (ST), and Other Backward Classes (OBC). These three groups have faced a long history of discrimination, and despite the Indian government's affirmative action policies, many still face discrimination even in the present day.

Table 3.9 reports the effect of the dry shock on total monthly per capita expenditure for both marginalized and non-marginalized households. I find that a one meter dry shock is associated with 3.5% and 5.2% less total spending for marginalized and non-marginalized groups, respectively. Although the magnitude of the estimate is larger for the latter group, the difference between the two coefficients is not statistically significant at convention levels. Hence, in terms of total expenditure, the negative effects of low precipitation appear to be shared equally across both the marginalized and the non-marginalized sectors.

The similar effects of low precipitation across the two groups is perhaps more apparent in Table 3.10, which shows effects on food consumption and nutritional intake. In particular, in the regressions on food expenditure (Panel A), the coefficient of the dry shock is almost identical for both groups, with an estimate of −5.7% for the marginalized and −5.9% for the non-marginalized, both individually significant at the 1% level. Furthermore, none of the differences in the coefficients for nutritional intake of calories (Panel B), protein (Panel C), and fat (Panel D) per capita are statistically significant between the two groups. Overall, these findings provide evidence that low precipitation has similar impacts on food utilization among marginalized and non-marginalized sectors of society.

Robustness Checks

Alternative Definition of Dry Shock. Thus far, I have investigated the effects of a negative rainfall shock measured in levels: the *Dry Shock* variable is defined as the absolute deviation in meters per year of total yearly rainfall below the district's long-term annual mean. Nevertheless, the impact of a one meter dry shock may not be comparable across agro-climactic zones, as some areas may have drier or wetter climates than others. Hence, to allow for rainfall shocks that are more comparable in magnitude across regions, I likewise consider an alternative definition of dry shock that is measured in units of standard deviation.

Table 3.11 shows results for the effects of this alternative definition of negative rainfall shocks on household consumption and nutrition outcomes. In this table, the regressor is *Standardized Dry Shock*, obtained by dividing the previously defined *Dry Shock* variable with the district's long-term standard deviation of annual rainfall. Notably, the patterns in this table are qualitatively similar to that in Table 3.4, where the independent variable is *Dry Shock*. The main difference appears to be in the results for fat, which does not attain statistical significance when using the standardized rainfall measure.

Intensity of Rainfall Shocks. I next consider the possibility that effects vary with the intensity of the rainfall shock. Specifically, I regress household consumption and nutrition outcomes on dummy variables representing bins of the standardized rainfall measure, including both positive and negative deviations from the long-term precipitation average. This analysis enables me to investigate two key ideas: first, whether there may be non-linearities in the impacts of rainfall, and second, whether positive shocks are indeed beneficial for households in the rural Indian context, as explained earlier in Section 3.4.

The regression results are reported in Table 3.12. Here, each bin is 0.75 standard deviations wide, and the bin centered at 0 is the omitted category in the regression. The effects are generally consistent with those in Table 3.4. The magnitude of the negative effects become larger as the dry shock becomes more severe, and much of the negative impacts of poor rainfall are concentrated with the driest shocks. Importantly, the results also show that the effects of the wet shock are by and large zero, with small coefficients showing no clear pattern across all bins and outcome variables.[20] These findings therefore support the argument that food security is a greater concern during times of drought.

[20]Wet shocks that are 1.875 standard deviations above the mean are associated with a decline in food expenditure, but a closer inspection on the effects of positive rainfall shocks on yields and prices suggest that this is likely due to lower food prices as a result of the outward shift in supply from abundant rainfall.

Spatial Correlation of Rainfall Shocks. Throughout this paper, I have cluster standard errors at the district level following Burgess et al. (2017) and other studies who have argued that measurement errors are likely correlated within districts over time. But to address robustness of the results to the spatial correlation of rainfall shocks, I implement two additional analyses: (1) clustered standard errors at the state-year level; and (2) spatially correlated errors following Conley (1999) using a temporal lag of 0 to 2 years and a spatial lag of 100 to 200 kilometers.[21] In both cases, I find that the results are very similar as in Table 3.4, with the exception that the standard errors for fat intake double, thus rendering the coefficient estimate statistically insignificant.

3.6 Conclusion

This paper shows that dry rainfall shocks have severe negative consequences on food consumption and nutrition among households in rural India. In particular, the results show that when annual rainfall is one meter below its long-term average, household per capita food expenditures fall by 5.6%. This decrease in food expenditures, in turn, translates into lower nutritional intake of calories, protein, and fat. Exploring these aggregate effects across different categories of food, I find that rural households continue to rely heavily on cereals for energy and protein even during times of drought. They also move away from highly income elastic goods that are typically consumed by richer and/or more urban households, such as processed foods, biscuits, and bottled beverages.

This paper likewise provides suggestive evidence that negative impacts of drought on food security likely operate through effects on income rather than food prices. For almost all food items I study, I am unable to reject the hypothesis that the effect of rainfall on prices is zero, though these null effects are not very precisely estimated. While households who derive income primarily from self-employment in agriculture are not impacted by low rainfall, likely because they are land-owners who have large savings or access to credit, the opposite is true for households who rely on agricultural labor. This finding is consistent with earlier studies for India which have argued that low levels of rainfall lead to low crop yields, and consequently, low agricultural wages (e.g. Jayachandran 2006).

[21] To estimate the Conley (1999) spatially correlated errors, I use the implementation by Hsiang (2010).

Table 3.1: Summary Statistics on Rainfall

	Rainfall Level		Prop. of Districts	Dry Shock	
Year	Mean	SD	w/ Dry Shock	Mean	SD
2004	1.284	0.947	0.690	0.211	0.131
2005	1.364	0.792	0.535	0.203	0.180
2006	1.327	0.831	0.585	0.230	0.197
2007	1.414	0.931	0.491	0.167	0.153
2008	1.446	0.801	0.365	0.158	0.163
2009	1.187	0.810	0.806	0.244	0.170
2010	1.477	0.984	0.355	0.253	0.178
2011	1.371	0.778	0.387	0.249	0.304
2012	1.326	0.868	0.596	0.166	0.135

Notes: This table shows rainfall variation measured in meters per year. Columns 1 and 2 show the mean and standard deviation of total annual rainfall across districts. Column 3 shows the proportion of districts in a given year which experience total annual rainfall below the long-run annual mean. The long-run mean is calculated over the years 1970 to 2012. Columns 4 and 5 show the mean and standard of the dry shock, defined as the absolute deviation of rainfall below the long-run mean. Data on rainfall come from Willmott and Matsuura (2015).

Table 3.2: Summary Statistics for Household Food Consumption and Nutrition

Variable	Food Group	2004-05	2005-06	2006-07	2007-08	2009-10	2011-12
Food Exp. (Rs.)	All categories	354.0 [163.1]	404.2 [190.2]	429.8 [200.3]	487.5 [220.7]	577.4 [260.8]	726.7 [336.9]
	Cereals	104.6 [38.7]	113.4 [43.9]	120.1 [48.1]	135.5 [49.7]	153.6 [66.2]	163.0 [71.1]
	Pulses, nuts, and oilseeds	25.2 [17.7]	30.6 [21.6]	34.5 [24.8]	35.8 [22.9]	48.0 [31.7]	58.9 [38.1]
	Vegetables and fruits	47.4 [27.6]	55.8 [33.1]	61.9 [35.2]	71.9 [41.3]	80.5 [47.5]	94.8 [57.3]
	Meat, fish, and dairy	76.4 [73.2]	94.7 [85.4]	94.9 [84.7]	115.7 [101.4]	133.3 [120.0]	185.1 [156.7]
	Sugar, honey, oils, and fats	46.2 [26.8]	50.4 [30.7]	49.0 [29.3]	51.2 [32.5]	67.9 [40.1]	86.6 [50.2]
	Misc. food and beverages	51.9 [45.2]	56.7 [49.7]	66.6 [58.4]	74.8 [60.2]	90.7 [70.9]	133.8 [104.5]
Share in Food Exp.	Cereals	32.7 [12.4]	31.2 [12.4]	25.6 [12.3]	30.7 [11.7]	27.3 [11.8]	24.6 [10.5]
	Pulses, nuts, and oilseeds	7.2 [3.6]	7.6 [3.8]	6.9 [4.2]	7.5 [3.6]	8.0 [4.3]	8.3 [4.0]
	Vegetables and fruits	13.5 [4.8]	13.9 [4.9]	12.3 [5.5]	14.9 [5.1]	13.3 [5.4]	13.2 [5.2]
	Meat, fish, and dairy	18.9 [12.1]	20.8 [12.0]	17.4 [11.7]	21.1 [11.7]	19.9 [12.1]	23.0 [12.3]
	Sugar, honey, oils, and fats	13.3 [5.2]	12.7 [5.3]	9.9 [5.1]	10.6 [4.9]	11.4 [5.1]	12.4 [5.2]
	Misc. food and beverages	14.0 [7.7]	13.5 [7.6]	12.9 [8.8]	15.0 [8.2]	14.8 [8.6]	18.1 [9.3]
Calories (kcal)	All categories	2159.9 [559.6]	2191.8 [577.8]	2197.2 [606.4]	2256.7 [578.1]	2172.6 [550.3]	2251.6 [547.8]
	Cereals	1412.8 [380.5]	1387.6 [388.2]	1403.0 [402.0]	1395.0 [370.8]	1343.0 [367.4]	1332.4 [350.7]
	Pulses, nuts, and oilseeds	109.9 [70.6]	120.5 [78.6]	120.1 [83.8]	116.2 [72.4]	109.5 [72.3]	129.0 [77.4]
	Vegetables and fruits	111.5 [66.8]	113.8 [67.3]	118.2 [71.1]	125.4 [70.1]	102.7 [59.8]	110.9 [63.0]
	Meat, fish, and dairy	172.5 [188.1]	198.7 [200.9]	182.7 [184.9]	214.9 [200.6]	182.9 [180.4]	193.3 [178.2]
	Sugar, honey, oils, and fats	277.3 [145.1]	294.5 [154.1]	289.5 [151.9]	278.3 [147.9]	304.6 [147.5]	331.3 [149.4]
	Misc. food and beverages	61.1 [65.6]	63.9 [66.3]	71.8 [78.9]	113.8 [135.6]	121.0 [149.2]	147.5 [162.7]
Protein (gm)	All categories	58.5 [17.9]	59.4 [18.2]	58.6 [18.6]	62.2 [18.1]	57.9 [16.8]	60.4 [16.9]
	Cereals	38.1 [11.6]	37.1 [11.4]	36.5 [11.5]	38.1 [11.2]	36.3 [10.9]	36.1 [10.6]
	Pulses, nuts, and oilseeds	6.5 [3.8]	6.9 [4.0]	6.7 [4.2]	6.8 [3.8]	6.2 [3.7]	7.4 [4.0]
	Vegetables and fruits	3.1 [1.8]	3.1 [1.7]	3.2 [1.8]	3.4 [1.8]	2.8 [1.5]	2.9 [1.6]
	Meat, fish, and dairy	8.8 [8.4]	10.3 [9.1]	10.1 [8.9]	10.8 [8.9]	9.4 [7.9]	10.2 [8.1]
	Sugar, honey, oils, and fats	0.0 [0.0]	0.0 [0.0]	0.0 [0.0]	0.0 [0.0]	0.0 [0.0]	0.0 [0.0]
	Misc. food and beverages	1.6 [1.5]	1.7 [1.5]	1.9 [1.8]	2.7 [2.9]	2.8 [3.2]	3.5 [3.4]
Fat (gm)	All categories	39.2 [21.5]	42.6 [23.0]	41.7 [22.7]	42.5 [23.1]	43.7 [21.3]	47.1 [21.5]
	Cereals	4.6 [2.9]	4.3 [2.6]	4.1 [2.8]	4.6 [2.7]	4.2 [2.4]	4.1 [2.2]
	Pulses, nuts, and oilseeds	2.2 [3.5]	2.7 [4.1]	3.0 [4.5]	2.3 [3.4]	2.4 [3.7]	2.8 [3.9]
	Vegetables and fruits	0.5 [0.3]	0.5 [0.3]	0.5 [0.3]	0.5 [0.3]	0.4 [0.2]	0.4 [0.2]
	Meat, fish, and dairy	11.3 [13.1]	13.0 [14.0]	11.8 [12.8]	14.2 [14.0]	12.0 [12.6]	12.6 [12.5]
	Sugar, honey, oils, and fats	18.5 [9.8]	19.9 [10.5]	20.0 [10.6]	17.8 [11.6]	21.7 [10.9]	23.7 [11.0]
	Misc. food and beverages	1.8 [1.7]	1.8 [1.7]	2.0 [2.0]	2.6 [2.7]	2.6 [2.8]	3.2 [3.0]

Notes: This table reports means (and in brackets, the standard deviation) for the following variables: food expenditure per capita per month (in Rupees); budget share of each food group in total food expenditure; calories per capita per day (in kcal); protein per capita per day (in grams); and fat per capita per day (in grams). Data come from the NSS Consumer Expenditure Surveys.

Table 3.3: Effects on Log Agricultural Yield

	Cereals		Pulses		Other Crops
	(1) Rice	(2) Wheat	(3) Gram	(4) Urad	(5) Potato
Dry shock	–0.307***	–0.199***	–0.242***	–0.124**	–0.108*
	(0.046)	(0.036)	(0.054)	(0.055)	(0.057)
Year FEs	Yes	Yes	Yes	Yes	Yes
District FEs	Yes	Yes	Yes	Yes	Yes
Adj. R-squared	0.765	0.856	0.609	0.628	0.773
Districts	540	478	483	473	463
Observations	4590	3982	3733	3697	3268

Notes: The dependent variable is ln(yield) at the district-level, where yield is production (measured in tonnes) divided by area planted (measured in hectares). Dry shock is the absolute deviation of rainfall below the long-run mean (i.e., 1970-2012) in meters per year. All regressions include year-interacted district characteristics from the 2001 Census, namely, percent SC, percent literate, percent employed, and total population. Data on agricultural yields cover 2003-2012 and come from the Directorate of Economics and Statistics, Ministry of Agriculture. Standard errors are clustered at the district level. *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table 3.4: Effects on Household Per Capita Food Consumption and Nutrition

	(1) Log Total Exp	(2) Log Food Exp	(3) Log Calories	(4) Log Protein	(5) Log Fat
Dry shock	–0.042**	–0.056***	–0.027**	–0.019	–0.037**
	(0.018)	(0.016)	(0.011)	(0.013)	(0.018)
Survey Round FEs	Yes	Yes	Yes	Yes	Yes
District FEs	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
Adj. R-squared	0.540	0.538	0.230	0.322	0.448
Observations	246565	246565	246565	246565	246565

Notes: The dependent variables, indicated in the column titles, are expressed in natural log and measured per capita at the household level. All expenditures include the value in Rupees of both market purchases and home production. Dry shock is the absolute deviation of total rainfall (measured in the past 12 months prior to the household's survey date) from the district's long-term mean of annual precipitation (i.e., from 1970-2012). All regressions include household-level controls, namely, dummies for religion, dummies social group (SC/ST/OBC), and the fraction of household members in each male/female age cell (0-1, 1-3, 4-5, 7-9, 10-12, 13-15, 16-19, 20-39, 40-49, 50-59, 60-69, 70+). Furthermore, all regressions include year-interacted district characteristics from the 2001 Census, namely, percent SC, percent literate, percent employed, and total population. Data on household food consumption and nutritional intake come from the NSS Consumer Expenditure Surveys, Rounds 60-68 (January 2004-June 2012). Standard errors are clustered at the district level. *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table 3.5: Effects on HH Per Capita Food Consumption and Nutrition, by Food Group

	Cereals	Pulses, nuts, and oilseeds	Vegetables and fruits	Meat, fish, and dairy	Sugar, honey, oils, and fats	Misc. food and beverages
Panel A: Budget Share of Total Food Consumption						
	(1)	(2)	(3)	(4)	(5)	(6)
Dry shock	–0.052**	–0.069**	–0.029	–0.044	–0.069***	–0.076**
	(0.022)	(0.029)	(0.031)	(0.030)	(0.024)	(0.034)
Panel B: Log Calories Per Capita						
	(1)	(2)	(3)	(4)	(5)	(6)
Dry shock	–0.024*	–0.020	–0.054*	–0.068**	–0.011	–0.137***
	(0.013)	(0.027)	(0.029)	(0.034)	(0.020)	(0.049)
Panel C: Log Protein Per Capita						
	(1)	(2)	(3)	(4)	(5)	(6)
Dry shock	–0.025*	–0.021	–0.049	–0.036	0.026	–0.106**
	(0.013)	(0.028)	(0.030)	(0.028)	(0.031)	(0.046)
Panel D: Log Fat Per Capita						
	(1)	(2)	(3)	(4)	(5)	(6)
Dry shock	–0.005	–0.068	–0.025	–0.089**	–0.071**	–0.200***
	(0.016)	(0.050)	(0.033)	(0.044)	(0.029)	(0.060)
Observations	246565	246565	246565	246565	246565	246565
Survey Round FEs	Yes	Yes	Yes	Yes	Yes	Yes
District FEs	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variables are expressed in natural logarithm and are measured per capita at the household level. Food expenditure (Panel A) includes both market purchases and home production. Dry shock is the absolute deviation of total rainfall (measured in the past 12 months prior to the household's survey date) from the district's long-term mean of annual precipitation (i.e., from 1970-2012). All regressions include household-level controls, namely, dummies for religion, dummies social group (SC/ST/OBC), and the fraction of household members in each male/female age cell (0-1, 1-3, 4-5, 7-9, 10-12, 13-15, 16-19, 20-39, 40-49, 50-59, 60-69, 70+). Furthermore, all regressions include year-interacted district characteristics from the 2001 Census, namely, percent SC, percent literate, percent employed, and total population. Data on household food consumption and nutritional intake come from the NSS Consumer Expenditure Surveys, Rounds 60-68 (January 2004-June 2012). Standard errors are clustered at the district level. *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table 3.6: Effects on Log Prices of Cereals, Pulses, Vegetables, and Fruits

	Cereals		Pulses		Vegetables and Fruits		
	(1) Wheat	(2) Rice	(3) Gram	(4) Urad	(5) Potato	(6) Spinach	(7) Banana
Dry shock	0.085	0.020	–0.001	0.007	0.037**	0.032	0.019
	(0.083)	(0.013)	(0.014)	(0.018)	(0.018)	(0.045)	(0.040)
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Market FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R-squared	0.764	0.910	0.860	0.946	0.840	0.717	0.817
Markets	353	373	361	357	373	361	373
Districts	3175	4699	4204	3914	4761	4205	4730

Notes: The dependent variable is ln(price) at the district-level. Prices are in nominal terms and are based on the median price across markets within each district-year. Dry shock is the absolute deviation of rainfall below the long-run mean (i.e., 1970-2012) in meters per year. All regressions include year-interacted district characteristics from the 2001 Census, namely, percent SC, percent literate, percent employed, and total population. Data on prices cover the years 2003-2006 and 2010-2011 and come from the NSS Rural Price Collection Survey. Standard errors are clustered at the district level. *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table 3.7: Effects on Log Household Total Monthly Expenditure Per Capita, by Main Income Source

	All HHs	Self-Employed		Laborer		Other Source
	(1)	(2) Ag	(3) Non-Ag	(4) Primarily Ag	(5) Both Ag/Non-Ag	(6)
Dry shock	–0.042**	–0.006	–0.054**	–0.046*	–0.020	–0.095***
	(0.018)	(0.022)	(0.025)	(0.027)	(0.027)	(0.032)
Survey Round FEs	Yes	Yes	Yes	Yes	Yes	Yes
District FEs	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R-squared	0.540	0.512	0.560	0.705	0.585	0.506
Observations	246565	73148	52637	51229	32071	37387

Notes: The dependent variable is ln(total monthly per capita expenditure), including both food and non-food expenditure. In column 1, the sample includes all households. In columns 2 to 6, the regression includes different subsamples of households depending on their main source of income. Dry shock is the absolute deviation of total rainfall (measured in the past 12 months prior to the household's survey date) from the district's long-term mean of annual precipitation (i.e., from 1970-2012). All regressions include household-level controls, namely, dummies for religion, dummies social group (SC/ST/OBC), and the fraction of household members in each male/female age cell (0-1, 1-3, 4-5, 7-9, 10-12, 13-15, 16-19, 20-39, 40-49, 50-59, 60-69, 70+). Furthermore, all regressions include year-interacted district characteristics from the 2001 Census, namely, percent SC, percent literate, percent employed, and total population. Data on household expenditures come from the NSS Consumer Expenditure Surveys, Rounds 60-68 (January 2004-June 2012). Standard errors are clustered at the district level. *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table 3.8: Effects on HH Food Consumption and Nutrition, by Main Income Source

	Self-Employed		Laborer		Other
	Ag	Non-Ag	Primarily Ag	Ag/Non-Ag	
Panel A: Log Food Expenditure					
	(1)	(2)	(3)	(4)	(5)
Dry shock	0.002	–0.081***	–0.077***	–0.042	–0.093***
	(0.020)	(0.022)	(0.023)	(0.026)	(0.026)
Panel B: Log Calories Per Capita					
	(1)	(2)	(3)	(4)	(5)
Dry shock	0.009	–0.041***	–0.039**	–0.050***	–0.040**
	(0.015)	(0.014)	(0.018)	(0.018)	(0.016)
Panel C: Log Protein Per Capita					
	(1)	(2)	(3)	(4)	(5)
Dry shock	0.011	–0.028*	–0.016	–0.038*	–0.041**
	(0.016)	(0.016)	(0.019)	(0.020)	(0.017)
Panel D: Log Fat Per Capita					
	(1)	(2)	(3)	(4)	(5)
Dry shock	0.016	–0.059**	–0.038	–0.070**	–0.083***
	(0.022)	(0.024)	(0.030)	(0.030)	(0.027)
Observations	73148	52637	51229	32071	37387
Survey Round FEs	Yes	Yes	Yes	Yes	Yes
District FEs	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variables are expressed in natural logarithm and are measured per capita at the household level. Food expenditure (Panel A) includes both market purchases and home production. Dry shock is the absolute deviation of total rainfall (measured in the past 12 months prior to the households survey date) from the districts long-term mean of annual precipitation (i.e., from 1970-2012). All regressions include household-level controls, namely, dummies for religion, dummies social group (SC/ST/OBC), and the fraction of household members in each male/female age cell (0-1, 1-3, 4-5, 7-9, 10-12, 13-15, 16-19, 20-39, 40-49, 50-59, 60-69, 70+). Furthermore, all regressions include year-interacted district characteristics from the 2001 Census, namely, percent SC, percent literate, percent employed, and total population. Data on household food consumption and nutritional intake come from the NSS Consumer Expenditure Surveys, Rounds 60-68 (January 2004-June 2012). Standard errors are clustered at the district level. *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table 3.9: Effects on Log HH Total Monthly Expenditure Per Capita, by Social Group

	All HHs	Social Group	
	(1)	(2) Marginalized	(3) Non-Marginalized
Dry shock	–0.042**	–0.035*	–0.052**
	(0.018)	(0.020)	(0.025)
Survey Round FEs	Yes	Yes	Yes
District FEs	Yes	Yes	Yes
Controls	Yes	Yes	Yes
Adj. R-squared	0.540	0.519	0.526
Observations	246565	174817	71786

Notes: The dependent variable is ln(total monthly per capita expenditure), including both food and non-food expenditure. In column 1, the sample includes all households. In columns 2 and 3, the regression includes different subsamples of households depending on their social group. The marginalized group consists of SC, ST, and OBC. Dry shock is the absolute deviation of total rainfall (measured in the past 12 months prior to the household's survey date) from the district's long-term mean of annual precipitation (i.e., from 1970-2012). All regressions include household-level controls, namely, dummies for religion, and the fraction f household members in each male/female age cell (0-1, 1-3, 4-5, 7-9, 10-12, 13-15, 16-19, 20-39, 40-49, 50-59, 60-69, 70+). Furthermore, all regressions include year-interacted district characteristics from the 2001 Census, namely, percent SC, percent literate, percent employed, and total population. Data on household expenditures come from the NSS Consumer Expenditure Surveys, Rounds 60-68 (January 2004-June 2012). Standard errors are clustered at the district level. *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table 3.10: Effects on HH Per Capita Food Consumption and Nutrition, by Social Group

	Social Group	
	Marginalized	Non-marginalized
Panel A: Log Food Expenditure		
	(1)	(2)
Dry shock	–0.057***	–0.059***
	(0.018)	(0.022)
Panel B: Log Calories Per Capita		
	(1)	(2)
Dry shock	–0.033**	–0.010
	(0.013)	(0.015)
Panel C: Log Protein Per Capita		
	(1)	(2)
Dry shock	–0.023	–0.004
	(0.014)	(0.016)
Panel D: Log Fat Per Capita		
	(1)	(2)
Dry shock	–0.022	–0.061***
	(0.020)	(0.022)
Observations	174817	71748
Survey Round FEs	Yes	Yes
District FEs	Yes	Yes

Notes: The dependent variables are expressed in natural logarithm and are measured per capita at the household level. Food expenditure (Panel A) includes both market purchases and home production. Dry shock is the absolute deviation of total rainfall (measured in the past 12 months prior to the household's survey date) from the district's long-term mean of annual precipitation (i.e., from 1970-2012). All regressions include household-level controls, namely, dummies for religion, dummies social group (SC/ST/OBC), and the fraction of household members in each male/female age cell (0-1, 1-3, 4-5, 7-9, 10-12, 13-15, 16-19, 20-39, 40-49, 50-59, 60-69, 70+). Furthermore, all regressions include year-interacted district characteristics from the 2001 Census, namely, percent SC, percent literate, percent employed, and total population. Data on household food consumption and nutritional intake come from the NSS Consumer Expenditure Surveys, Rounds 60-68 (January 2004-June 2012). Standard errors are clustered at the district level. *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table 3.11: Robustness to Alternative Definition of Rainfall Shocks: Standardized Measure

	(1) Log Food Exp	(2) Log Calories	(3) Log Protein	(4) Log Fat
Standardized dry shock	–0.015***	–0.009**	–0.006	–0.004
	(0.005)	(0.004)	(0.004)	(0.005)
Survey Round FEs	Yes	Yes	Yes	Yes
District FEs	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Adj. R-squared	0.538	0.230	0.322	0.448
Observations	246565	246565	246565	246565

Notes: The dependent variables, indicated in the column titles, are expressed in natural log and measured per capita at the household level. All expenditures include the value in Rupees of both market purchases and home production. Standardized dry shock is the absolute deviation of total rainfall (measured in the past 12 months prior to the household's survey date) below the district's long-term mean of annual precipitation (i.e., from 1970-2012) divided by the standard deviation. All regressions include household-level controls, namely, dummies for religion, dummies social group (SC/ST/OBC), and the fraction of household members in each male/female age cell (0-1, 1-3, 4-5, 7-9, 10-12, 13-15, 16-19, 20-39, 40-49, 50-59, 60-69, 70+). Furthermore, all regressions include year-interacted district characteristics from the 2001 Census, namely, percent SC, percent literate, percent employed, and total population. Data on household food consumption and nutritional intake come from the NSS Consumer Expenditure Surveys, Rounds 60-68 (January 2004-June 2012). Standard errors are clustered at the state-year level. *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table 3.12: Effects on Food Consumption and Nutrition by Rainfall Shock Intensity

	(1) Log Food Exp	(2) Log Calories	(3) Log Protein	(4) Log Fat
Dry shock: > 2.625 SD below	–0.121***	–0.080***	–0.079***	–0.173***
	(0.014)	(0.010)	(0.011)	(0.014)
Dry shock: 1.875-2.625 SD below	–0.088***	–0.059***	–0.065***	–0.066**
	(0.020)	(0.016)	(0.017)	(0.026)
Dry shock: 1.125-1.875 SD below	–0.008	–0.006	–0.003	0.003
	(0.007)	(0.005)	(0.005)	(0.007)
Dry shock: 0.375-1.125 SD below	–0.006	–0.003	–0.001	0.002
	(0.005)	(0.003)	(0.003)	(0.005)
Wet shock: 0.375-1.125 SD above	–0.001	–0.001	–0.004	0.008
	(0.005)	(0.003)	(0.004)	(0.005)
Wet shock: 1.125-1.875 SD above	–0.010	0.003	–0.001	0.018**
	(0.007)	(0.005)	(0.005)	(0.008)
Wet shock: 1.875-2.625 SD above	–0.029***	–0.008	–0.014*	0.002
	(0.010)	(0.007)	(0.008)	(0.012)
Wet shock: > 2.625 SD above	–0.054**	–0.008	–0.010	–0.018
	(0.021)	(0.015)	(0.015)	(0.024)
Survey Round FEs	Yes	Yes	Yes	Yes
District FEs	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Adj. R-squared	0.538	0.230	0.322	0.448
Observations	246565	246565	246565	246565

Notes: The dependent variables, indicated in the column titles, are expressed in natural log and measured per capita at the household level. All expenditures include the value in Rupees of both market purchases and home production. The independent variables are dummy variables representing bins of standard deviations of total rainfall (measured in the past 12 months prior to the household's survey date) from the district's long-term mean of annual precipitation (i.e., from 1970-2012). The omitted category is the bin for deviations within 0.375 standard deviations of the mean. All regressions include household-level controls, namely, dummies for religion, dummies social group (SC/ST/OBC), and the fraction of household members in each male/female age cell (0-1, 1-3, 4-5, 7-9, 10-12, 13-15, 16-19, 20-39, 40-49, 50-59, 60-69, 70+). Furthermore, all regressions include year-interacted district characteristics from the 2001 Census, namely, percent SC, percent literate, percent employed, and total population. Data on household food consumption and nutritional intake come from the NSS Consumer Expenditure Surveys, Rounds 60-68 (January 2004-June 2012). Standard errors are clustered at the district level. *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Bibliography

Agarwal, Sumit, Souphala Chomsisengphet, Neale Mahoney, and Johannes Stroebel. 2014. "Regulating consumer financial products: Evidence from credit cards." *Quarterly Journal of Economics* 130 (1): 111–164.

Akpalu, Wisdom, Hassan M. Rashid, and Claudia Ringler. 2011. "Climate variability and maize yield in the Limpopo region of South Africa: Results from GME and MELE methods." *Climate and Development* 3 (2): 114–122.

Angrist, Joshua, Daniel Lang, and Philip Oreopoulos. 2009. "Incentives and services for college achievement: Evidence from a randomized trial." *American Economic Journal: Applied Economics* 1 (1): 136–163.

Angrist, Joshua, and Victor Lavy. 2009. "The effects of high stakes high school achievement awards: Evidence from a randomized trial." *American Economic Review* 99 (4): 1384–1414.

Ashraf, Nava. 2009. "Spousal control and intra-household decision making: An experimental study in the Philippines." *American Economic Review* 99 (4): 1245–1277.

Bandura, Albert, and Daniel Cervone. 1983. "Self-evaluative and self-efficacy mechanisms governing the motivational effects of goal systems." *Journal of Personality and Social Psychology* 45 (5): 1017–1028.

Barnes, Douglas F., Kerry Krutilla, and William Hyde. 2004. *The Urban Household Energy Transition: Energy, Poverty, and the Environment in the Developing World.* Washington, DC: Resources for the Future.

Barron, Manuel, and Maximo Torero. 2017. "Household Electrification and Indoor Air Pollution." Unpublished.

Bartels, Daniel, and Abigail Sussman. 2015. "The influence of goal-setting on credit card payment decisions: A first look." Paper presented at the IPA Research Gathering on Financial Inclusion, New Haven, CT.

Berg, Gunhild, and Bilal Zia. 2017. "Harnessing emotional connections to improve financial decisions: Evaluating the impact of financial education in mainstream media." *Journal of the European Economic Association:* Forthcoming.

Bettinger, Eric P. 2012. "Paying to learn: The effect of financial incentives on elementary school test scores." *Review of Economics and Statistics* 94 (3): 686–698.

Blakeslee, David S., and Ram Fishman. 2017. "Weather Shocks, Agriculture, and Crime: Evidence from India." *Journal of Human Resources:* forthcoming.

Bruhn, Miriam, Gabriel Lara Ibarra, and David McKenzie. 2014. "The minimal impact of a large-scale financial education program in Mexico City." *Journal of Development Economics* 108:184–189.

Bruhn, Miriam, Luciana de Souza Leão, Arianna Legovini, Rogelio Marchetti, and Bilal Zia. 2016. "The impact of high school financial education: Evidence from a large-scale evaluation in Brazil." *American Economic Journal: Applied Economics* 8 (4): 256–295.

Bryan, Judith F., and Edwin A. Locke. 1967. "Goal-setting as a means of increasing motivation." *Journal of Applied Psychology* 51 (3): 274–277.

Burgess, Robin, Olivier Deschenes, Dave Donaldson, and Michael Greenstone. 2017. "The Unequal Effects of Weather and Climate Change: Evidence from Mortality in India." Working Paper.

Burgess, Robin, and Rohini Pande. 2005. "Do rural banks matter? Evidence from the Indian social banking experiment." *American Economic Review* 95 (3): 780–795.

Burlig, Fiona, and Louis Preonas. 2016. "Out of the Darkness and Into the Light? Development Effects of Rural Electrification." Energy Institute at Haas Working Paper No. 268.

Carpena, Fenella, Shawn Cole, Jeremy Shapiro, and Bilal Zia. 2015. "Unpacking the causal chain of financial literacy." Working Paper.

Chakravorty, Ujjayant, Kyle Emerick, and Majah-Leah Ravago. 2016. "Lighting up the last mile: The benefits and costs of extending electricity to the rural poor." Unpublished.

Chong, Alberto, Dean Karlan, and Martin Valdivia. 2010. *Using radio and video as a means for financial education in Peru.* `http://www.poverty-action.org/study/financial-education-delivered-through-radio-and-videos-among-low-income-households-cuzco-peru`. Accessed: 2016-05-30.

Cole, Shawn, Xavier Giné, Jeremy Tobacman, Petia Topalova, Robert Townsend, and James Vickery. 2013. "Barriers to household risk management: Evidence from India." *American Economic Journal: Applied Economics* 5 (1): 104–135.

Cole, Shawn, Andrew Healy, and Eric Werker. 2012. "Do voters demand responsive governments? Evidence from Indian disaster relief." *Journal of Development Economics* 97:167–181.

Cole, Shawn, Thomas Sampson, and Bilal Zia. 2011. "Prices or knowledge? What drives demand for financial services in emerging markets?" *Journal of Finance* 66 (6): 1933–1967.

Collins, Daryl, Jonathan Morduch, Stuart Rutherford, and Orlanda Ruthven. 2009. *Portfolios of the Poor: How the World's Poor Live on $2 a Day.* Princeton, NJ: Princeton University Press.

Collins, J Michael. 2013. "The impacts of mandatory financial education: Evidence from a randomized field study." *Journal of Economic Behavior & Organization* 95:146–158.

Collins, J Michael, and Collin M. O'Rourke. 2010. "Financial education and counseling—Still holding promise." *Journal of Consumer Affairs* 44 (3): 483–498.

Conley, Tim G. 1999. "GMM estimation with cross sectional dependence." *Journal of Econometrics* 92:1–45.

Dai, Aiguo. 2013. "Increasing drought under global warming in obsevations and models." *Nature Climate Change* 3 (1): 52–58.

Dalal, Aparna, and Jonathan Morduch. 2010. "The psychology of microinsurance: Small changes can make a surprising difference." *ILO Microinsurance Paper No. 5.*

Deaton, Angus. 1990. "Price Elasticities from Survey Data: Extensions and Indonesian Results." *Journal of Econometrics* 44 (33): 281–309.

———. 1997. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy.* Washington, DC: World Bank.

———. 1998. "Quality, Quantity, and Spatial Variation of Price." *American Economic Review* 78 (3): 418–430.

Deaton, Angus, and Jean Drèze. 2009. "Food and Nutrition in India: Facts and Interpretations." *Economic and Political Weekly* XLIV (7): 42–65.

Dell, Melissa, Benjamin F. Jones, and Benjamin A. Olken. 2014. "What do we learn from the weather? The new climate–economy literature." *Journal of Economic Literature* 52 (3): 740–798.

Dinkelman, Taryn. 2011. "The effects of rural electrification on employment: New evidence from South Africa." *American Economic Review* 101 (7): 3078–3108.

Doi, Yoko, David McKenzie, and Bilal Zia. 2014. "Who you train matters: Identifying combined effects of financial education on migrant households." *Journal of Development Economics* 109:39–55.

Drexler, Alejandro, Greg Fischer, and Antoinette Schoar. 2014. "Keeping it simple: Financial literacy and rules of thumb." *American Economic Journal: Applied Economics* 6 (2): 1–31.

Duflo, Esther, and Rohini Pande. 2007. "Dams." *Quarterly Journal of Economics* 122 (2): 601–646.

Duflo, Esther, and Emmanuel Saez. 2003. "The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment." *Quarterly Journal of Economics* 118 (3): 815–842.

Dupas, Pascaline, and Jonathan Robinson. 2013. "Savings constraints and microenterprise development: Evidence from a field experiment in Kenya." *American Economic Journal: Applied Economics* 5 (1): 163–192.

Earley, P. Christopher, and Brian C. Perry. 1987. "Work plan availability and performance: An assessment of task strategy priming on subsequent task completion." *Organizational Behavior and Human Decision Processes* 39 (3): 279–302.

Eli, Shari, and Nicholas Li. 2017. "Caloric Requirements and Food Consumption Puzzles of the Poor." WORKING PAPER.

FAO, IFAD and WFP. 2013. *The State of Food Insecurity in the World: The Multiple Dimensions of Food Security.* Rome: FAO.

Fernandes, Daniel, John G. Lynch Jr., and Richard G. Netemeyer. 2014. "Financial literacy, financial education, and downstream financial behaviors." *Management Science* 60 (8): 1861–1883.

Financial Literacy Foundation. 2007. *Financial literacy: Australians understanding money.* Parkes ACT: Commonwealth of Australia.

Fryer, Roland G, Jr. 2011. "Financial incentives and student achievement: Evidence from randomized trials." *Quarterly Journal of Economics* 126 (4): 1755–1798.

Gartner, Kimberly, and Richard M. Todd. 2005. "Effectiveness of online 'early intervention' financial education programs for credit-card holders." Working Paper.

Goerg, Sebastian J., and Sebastian Kube. 2012. "Goals (th)at Work: Goals, Monetary Incentives, and Workers' Performance." Working Paper.

Gopalan, C., B.V. Rama Sastri, and S. C. Balasubramanian. 1991. *Nutritive Value of Indian Foods.* Hyderabad: National Institute of Nutrition, Indian Council of Medical Research.

Guiteras, Raymond. 2009. "The Impact of Climate Change on Indian Agriculture." Unpublished Manuscript, Department of Economics, University of Maryland, College Park.

Harding, Matthew, and Alice Hsiaw. 2014. "Goal setting and energy conservation." *Journal of Economic Behavior & Organization* 107:209–227.

Hastings, Justine S., Brigitte C. Madrian, and William L. Skimmyhorn. 2013. "Financial literacy, financial education, and economic outcomes." *Annual Review of Economics* 5:347–373.

Heath, Chip, Richard P. Larrick, and George Wu. 1999. "Goals as reference points." *Cognitive Psychology* 38 (1): 79–109.

Hou, Xiaohui. 2010. "Can drought increase total calorie availability? The impact of drought on food consumption and the mitigating effects of a conditional cash transfer program." *Economic Development and Cultural Change* 58 (4): 713–737.

Hsiang, Solomon. 2010. "Temperatures and cyclones strongly associated with economic production in the Carribean and Central America." *Proceedings of the National Academy of Sciences* 107:15367–15372.

Hsiaw, Alice. 2013. "Goal-setting and self-control." *Journal of Economic Theory* 148 (2): 601–626.

International Energy Agency. 2016. *World Energy Outlook.* Paris: IEA.

Iyer, Lakshmi, and Petial Topalova. 2014. "Poverty and Crime: Evidence from rainfall and trade shocks in India." Working Paper.

Jain, Abishek, Sudatta Ray, Karthik Ganesan, Michael Aklin, Chao-Yo Cheng, and Johannes Urpelainen. 2015. *Access to Clean Cooking Energy and Electricity.* New Delhi: Council on Energy, Environment, and Water.

Jamison, Julian C., Dean Karlan, and Jonathan Zinman. 2011. "Financial education and access to savings accounts: Complements or substitutes? Evidence from Ugandan youth clubs." *NBER Working Paper No. 17020.*

Jayachandran, Seema. 2006. "Selling Labor Low: Wage Responses to Productivity Shocks in Developing Countries." *Journal of Political Economy* 114 (3): 538–575.

Kantor, Paula, and Padmaja Nair. 2003. "Risks and responses among the urban poor in India." *Journal of International Development* 15 (8): 957–967.

Karlan, Dean, Margaret McConnell, Sendhil Mullainathan, and Jonathan Zinman. 2016. "Getting to the top of mind: How reminders increase saving." *Management Science* 62 (12): 3393–3411.

King, Abby C., Robert Friedman, Bess Marcus, Cynthia Castro, Melissa Napolitano, David Ahn, and Lawrence Baker. 2007. "Ongoing physical activity advice by humans versus computers: The Community Health Advice by Telephone (CHAT) trial." *Health Psychology* 26 (6): 718–727.

Koch, Alexander K., and Julia Nafziger. 2011. "Goals and psychological accounting." *IZA Working Paper No. 5802.*

Kremer, Michael, Edward Miguel, and Rebecca Thornton. 2009. "Incentives to learn." *Review of Economics and Statistics* 91 (3): 437–456.

Laibson, David. 1997. "Golden eggs and hyperbolic discounting." *Quarterly Journal of Economics* 112 (2): 443–478.

LaPorte, Ronald E., and Raghu Nath. 1976. "Role of performance goals in prose learning." *Journal of Educational Psychology* 68 (3): 260–264.

Lerman, Caryn, Edward Lustbader, Barbara Rimer, Mary Daly, Suzanne Miller, Colleen Sands, and Andrew Balshem. 1995. "Effects of individualized breast cancer risk counseling: a randomized trial." *Journal of the National Cancer Institute* 87 (4): 286–292.

Leuven, Edwin, Hessel Oosterbeek, and Bas Klaauw. 2010. "The effect of financial rewards on students' achievement: Evidence from a randomized experiment." *Journal of the European Economic Association* 8 (6): 1243–1265.

Lipscomb, Molly, Mushfiq A. Mobarak, and Tania Barham. 2013. "Development effects of electrification: Evidence from the topographic placement of hydropower plants in Brazil." *American Economic Journal: Applied Economics* 5 (2): 200–231.

Lobell, David B., and Marshall Burke. 2010. "Climate Effects on Food Security: An Overview." Chap. 2 in *Climate Change and Food Security: Adapting Agriculture to a Warmer World,* edited by David B. Lobell and Marshall Burke, 37:13–30. Advances in Global Change Research. Dordrecht: Springer.

Locke, Edwin A., and Gary P. Latham. 2002. "Building a practically useful theory of goal setting and task motivation: A 35-year odyssey." *American Psychologist* 57 (9): 705–717.

Lusardi, Annamaria. 2015. "Financial Literacy Skills for the 21st Century: Evidence from PISA." *Journal of Consumer Affairs* 49 (3): 639–659.

Lusardi, Annamaria, and Olivia S. Mitchell. 2009. "How ordinary consumers make complex economic decisions: Financial literacy and retirement readiness." *NBER Working Paper No. 15350.*

———. 2014. "The economic importance of financial literacy: Theory and evidence." *Journal of Economic Literature* 52 (1): 5–44.

Maccini, Sharon, and Dean Yang. 2009. "Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall." *American Economic Review* 99 (3): 1006–1026.

Milkman, Katherine L., John Beshears, James J. Choi, David Laibson, and Brigitte C. Madrian. 2011. "Using implementation intentions prompts to enhance influenza vaccination rates." *Proceedings of the National Academy of Sciences* 108 (26): 10415–10420.

Miller, Margaret, Julia Reichelstein, Christian Salas, and Bilal Zia. 2015. "Can you help someone become financially capable? A meta-analysis of the literature." *The World Bank Research Observer* 30 (2): 220–246.

Morduch, Jonathan. 1995. "Income smoothing and consumption smoothing." *Journal of Economic Perspectives* 9 (3): 103–114.

Mullainathan, Sendhil, and Eldar Shafir. 2009. "Savings policy and decision-making in low-income households." In *Insufficient Funds: Savings, Assets, Credit, and Banking among Low-income Households,* edited by Michael S. Barr and Rebecca M. Blank, 121–145. New York, NY: Russell Sage Foundation Press.

———. 2013. *Scarcity: Why Having Too Little Means So Much.* London: Macmillan.

National Institute of Nutrition. 2011. *Dietary Guidelines for Indians: A Manual.* Hyderabad: Indian Council of Medical Research.

Nickerson, David W., and Todd Rogers. 2010. "Do you have a voting plan? Implementation intentions, voter turnout, and organic plan making." *Psychological Science* 21 (2): 194–199.

NSS. 2001. *Nutritional Intake in India, 1990-2001.* 471(55/1.0/9). New Delhi: National Sample Survey Office, Ministry of Statistics / Program Implementation.

———. 2007. *Nutritional Intake in India, 2004-2005.* 513(61/1.0/6). New Delhi: National Sample Survey Office, Ministry of Statistics / Program Implementation.

———. 2012. *Nutritional Intake in India, 2009-2010.* 540(66/1.0/2). New Delhi: National Sample Survey Office, Ministry of Statistics / Program Implementation.

———. 2014. *Nutritional Intake in India, 2011-2012.* 560(68/1.0/3). New Delhi: National Sample Survey Office, Ministry of Statistics / Program Implementation.

Pakes, Ariel, and Zvi Griliches. 1980. "Patents and R and D at the Firm Level: A First Report." *Economics Letters* 5 (4): 377–381.

Parry, Martin L., Cynthia Rosenzweig, Ana Iglesias, Matthew Livermore, and Gunther Fischer. 2004. "Effects of climate change on global food production under SRES emissions and socio-economic scenarios." *Global Environmental Change* 14 (1): 53–67.

Pathania, Vikram. 2007. "The Long Run Impact of Drought at Birth on Height of Women in Rural India." WORKING PAPER.

Programme Evaluation Organisation. 2014. *Evaluation Report on Rajiv Gandhi Grameen Vidyutikaran Yojana* (PEO Report No. 224). New Delhi: Government of India.

Proper, Karin I., Vincent H. Hildebrandt, Allard J. Van der Beek, Jos W.R. Twisk, and Willem Van Mechelen. 2003. "Effect of individual counseling on physical activity fitness and health: A randomized controlled trial in a workplace setting." *American Journal of Preventive Medicine* 24 (3): 218–226.

Rawal, Vikas. 2014. *On Identification of Rural Labour Households in NSS Surveys.* `http://www.indianstatistics.org/2014/04/02/rural-labour-households.html`.

Rosenzweig, Mark R., and Hans P. Binswanger. 1993. "Wealth, Weather Risk and the Composition and Profitability of Agricultural Investments." *Economic Journal* 103 (416): 56–78.

Rothkopf, Ernst Z., and M.J. Billington. 1979. "Goal-guided learning from text: Inferring a descriptive processing model from inspection times and eye movements." *Journal of Educational Psychology* 71 (3): 310–327.

Salisbury, Linda Court. 2014. "Minimum payment warnings and information disclosure effects on consumer debt repayment decisions." *Journal of Public Policy & Marketing* 33 (1): 49–64.

Salop, Steven C. 1979. "Monopolistic Competition with Outside Goods." *The Bell Journal of Economics* 10 (1): 141–156.

Sekhri, Sheetal, and Adam Storeygard. 2014. "Dowry deaths: Response to weather variability in India." *Journal of Development Economics* 111:212–223.

Shah, Manisha, and Bryce Millet Steinberg. 2017. "Drought of Opportunities: Contemporaneous and Long-Term Impacts of Rainfall Shocks on Human Capital." *Journal of Political Economy* 125 (2): 527–561.

Shilts, Mical Kay, Marcel Horowitz, and Marilyn S. Townsend. 2004. "Goal setting as a strategy for dietary and physical activity behavior change: A review of the literature." *American Journal of Health Promotion* 19 (2): 81–93.

Soman, Dilip, and Amar Cheema. 2011. "Earmarking and partitioning: Increasing saving by low-income households." *Journal of Marketing Research* 48 (SPL): S14–S22.

Soman, Dilip, and Min Zhao. 2011. "The fewer the better: Number of goals and savings behavior." *Journal of Marketing Research* 48 (6): 944–957.

Standing Committe on Energy. 2009. *Implementation of Rajiv Gandhi Grameen Vidyutikaran Yojana* (31st Report, Fourteenth Lok Sabha). New Delhi: Lok Sabha Secretariat.

———. 2013. *Implementation of Rajiv Gandhi Grameen Vidyutikaran Yojana* (41st Report, Fifteenth Lok Sabha). New Delhi: Lok Sabha Secretariat.

Thaler, Richard H. 1999. "Mental accounting matters." *Journal of Behavioral Decision Making* 12 (3): 183–206.

Ülkümen, Gülden, and Amar Cheema. 2011. "Framing goals to influence personal savings: The role of specificity and construal level." *Journal of Marketing Research* 48 (6): 958–969.

United Nations. 2017. *UNDP support to the implementation of sustainable development goal 7: Affordable and clean energy.*

Wheeler, Tim, and Joachim von Braun. 2013. "Climate Change Impacts on Global Food Security." *Science* 341 (6145): 508–513.

Willis, Lauren E. 2008. "Against financial literacy education." *Iowa Law Review* 94:197–286.

———. 2011. "The Financial Education Fallacy." *American Economic Review Papers and Proceedings* 101 (3): 429–434.

Willmott, C.J., and K. Matsuura. 2015. *Terrestrial Precipitation: 1900-2014 Gridded Monthly Time Series (Version 4.01).*

World Bank. 2003. *India: Access of the Poor to Clean Household Fuels.* Energy Sector Management Assistance Program, 263/03. Washington, DC: World Bank.

Xu, Lisa, and Bilal Zia. 2012. "Financial literacy around the world: an overview of the evidence with practical suggestions for the way forward." *World Bank Policy Research Working Paper No. 6107.*

Zimmermann, Laura. 2017. "Inside the Black Box of Rainfall Shocks: Rainfall and Market Prices in India." Poster session presented at the ASSA Annual Meeting, Chicago, IL.

Zipper, Samuel C., Jiangxiao Qiu, and Christopher J. Kucharik. 2016. "Drought Effects on US Maize and Soybean Production: Spatiotemporal Patterns and Historical Changes." *Environmental Research Letters* 11 (9): 094021.

Appendix A

The Adverse Effects of Electrification: Evidence from India

A.1 Proof of Lemma 1

Lemma 1 *When electricity is available, the following prices p^* and number of retailers n^* constitute an equilibrium in the kerosene market, given that $p_e + \gamma > \frac{t}{2n^*} + p^* > p_e > p^*$:*

(1) $p^* = c + \sqrt{\frac{tF}{(2-\lambda)M}}$

(2) $n^* = \frac{\lambda t\sqrt{(2-\lambda)M}}{\sqrt{tF}(3-2\lambda)-2(p_e-c)(1-\lambda)\sqrt{(2-\lambda)M}}$

In this equilibrium, some *type L and* no *type H consumers adopt electricity.*

Proof Let seller i be located at 0 and seller j be located at $1/n$, denoting p_i and $\bar{p}$ as seller i's price and all other sellers' prices, respectively. An H-type consumer located at $\hat{x}$ between i and j is indifferent between the two sellers if both provider her with the same utility: $v - t|\hat{x}| - p_i = v - t|\frac{1}{n} - \hat{x}| - \bar{p}$, where $\hat{x} = \frac{\bar{p}-p_i+t/n}{2t}$. Similarly, an L-type consumer located at $\hat{y}$ between i and j is indifferent between electricity and kerosene from seller if both options provide her with the same utility: i if $v - t|\hat{y}| - p_i = v - p_e$, where $\hat{y} = \frac{p_e-p_i}{t}$.

Since all H-type consumers living in $[0, \hat{x}]$ and all L-type consumers living in $[0, \hat{y}]$ purchase kerosene from i, seller i's demand curve is given by

$$D(p_i) = 2 \cdot \left[\frac{p_e - p_i}{t}(1-\lambda)M + \frac{p_j - p_i + t/n}{2t}\lambda M\right].$$

Thus, seller i's profit maximization problem is

$$\max_{p_i}\ (p_i - c)\frac{2M}{t}\left[(p_e - p_i)(1-\lambda) + \frac{p_j - p_i + t/n}{2}\lambda\right] - F.$$

Taking the FOC with respect to p_i and imposing symmetry $p_i = \bar{p} = p^*$ yields $(p_e - p_i^*)(1 - \lambda) + \frac{\lambda t}{2n} = \left(1 - \frac{\lambda}{2}\right)(p^* - c)$. This equation, combined with the zero profit condition, results in equilibrium price

$$p^* = c + \sqrt{\frac{tF}{(2-\lambda)M}}.$$

Substituting p^* into the zero profit condition obtains the equilibrium number of sellers

$$n^* = \frac{\lambda t\sqrt{(2-\lambda)M}}{\sqrt{tF}(3-2\lambda) - 2(p_e - c)(1-\lambda)\sqrt{(2-\lambda)M}}.$$

For the above p^* and n^* to be supported as an equilibrium where only some L-types adopt electricity, the following conditions must hold. First, so that no H-types connect to electricity, the type H consumer traveling the longest distance to a kerosene seller must receive higher utility from electricity than kerosene, i.e. $p_e + \gamma > t/2n^* + p^*$. This condition implies that the total cost of buying electricity for H-types is greater than the cost of buying kerosene for the H-type consumer living farthest away from a kerosene seller.

Second, so that some L-types continue to use kerosene but others switch to electricity, the type L traveling the longest distance to a kerosene seller must prefer electricity, but the type L consumer living exactly at a kerosene seller's location must prefer kerosene. These conditions require $p_e < p^* + t/2n^*$ and $p_e > p^*$. In other words, the electricity price must be lower than the cost of buying kerosene for the L-type consumer living farthest away from a kerosene seller, but higher than the kerosene price. Putting all of these conditions together yields the inequalities $p_e + \gamma > \frac{t}{2n^*} + p^* > p_e > p^*$. ∎

A.2 Proof of Lemma 2

Lemma 2 *When electricity is available, the following prices p^* and number of retailers n^* constitute an equilibrium in the kerosene market, given that $p_e + \gamma > \frac{t}{n^*} + p^*$ and $p_e < c + \sqrt{\frac{tF}{(2-\lambda)M}}$:*

(1) $p^* = c + \sqrt{\frac{tF}{\lambda M}}$

(2) $n^* = \sqrt{\frac{t\lambda M}{F}}$

In this equilibrium, all *type L and* no *type H consumers adopt electricity.*

Proof Notice that if all type L consumers switch to electricity and type H consumers do not, the market size for kerosene is λM (i.e., the fraction of consumers that are type H). Consequently, p^* and n^* are analogous to the benchmark case in Lemma 0, where the market size is λM instead of M. Following the same solution process as in Lemma 0, we obtain kerosene prices $p^* = m + \sqrt{tF/\lambda M}$ and number of kerosene sellers $n^* = \sqrt{t\lambda M/F}$.

For this p^* and n^* to be supported as an equilibrium, we need three conditions. First, for all type L consumers to adopt electricity, it must be that the type L consumer living closest to a kerosene retailer (in other words, at the same location as the kerosene retailer) receives higher utility from electricity than kerosene. This requirement amounts to $p^* > p_e$, so the price of kerosene must be higher than the price of electricity.

Second, for all type H consumers to remain in the kerosene market, the type H consumer living farthest away from a kerosene seller must receive higher utility from kerosene than electricity, i.e., $p_e + \gamma > \frac{t}{2n^*} + p^*$. This inequality means that the total cost of electricity for H-types is larger than the total cost of buying kerosene for the consumer traveling the longest distance to a kerosene store. This condition is similar to Case 1, wherein no H-types adopt electricity, as in Case 2.

Third, we must ensure that no kerosene retailer can undercut electricity by setting a price $\hat{p} \leq p_e$ to attract type L consumers back into the kerosene market. Given a kerosene price $\hat{p} < p_e$, the kerosene retailer's profit maximization problem is

$$\hat{\Pi}(\hat{p}) = \max_{\hat{p} \leq p_e} (\hat{p} - c)\frac{2M}{t}\left[(p_e - \hat{p})(1-\lambda) + \frac{p^* - \hat{p} + t/n^*}{2}\lambda\right] - F.$$

Taking the derivative with respect to $\hat{p}$, we find that the value $\tilde{p} = \frac{p_e(1-\lambda)+(p^*+t/n^*)(\lambda/2)+m(1-\lambda/2)}{2-\lambda}$ solves the FOC. Denote $\hat{p}^*$ as the price that maximizes $\hat{\Pi}(\hat{p})$, and consider the following cases.

(1) Assume $\tilde{p} \geq p_e$. Since the objective function is concave, we have a corner solution so that $\hat{p}^* = p_e$. However, note that $\hat{\Pi}(p_e) < 0$ because setting $\hat{p} = p^*$ maximizes $(\hat{p} - c)\frac{2M}{t}\left[\frac{p^* - p + t/n^*}{2}\lambda\right] - F$ and yields zero profits. Hence, selling kerosene at price p_e is not profitable for any retailer.

(2) Assume $\tilde{p} < p_e$. Setting $\hat{p}^* = \tilde{p}$ maximizes the profits from undercutting electricity, and $\hat{\Pi}(\tilde{p}) = (\tilde{p} - c)^2(M/t)(2-\lambda) - F$. For this to be an unprofitable deviation, we must require $\hat{\Pi}(\tilde{p}) < 0$, which results in $\tilde{p} < c + \sqrt{\frac{tF}{(2-\lambda)M}}$. Since $\tilde{p} < p_e$, a sufficient condition for the latter to hold is for $p_e < \sqrt{\frac{tF}{(2-\lambda)M}}$, so that no kerosene retailer can profitably deviate.

Taking all the above three conditions, we have (i) $p^* > p_e$, (ii) $p_e + \gamma > \frac{t}{2n^*} + p^*$, and (iii) $p_e < c + \sqrt{\frac{tF}{(2-\lambda)M}}$. However, because condition (iii) implies condition (i), the equilibrium can be supported with only conditions (ii) and (iiii). ∎

Appendix B

The ABCs of Financial Education: Experimental Evidence on Attitudes, Behavior, and Cognitive Biases

B.1 Content of Financial and Health Literacy Videos

This appendix describes the content of the video-based interventions.

Financial Literacy Videos

Session 1: Budgeting. Budgeting is the building block of household financial planning and management and the video aims at making the audience appreciate the need of keeping a household budge today in order to plan and save for a better tomorrow. The video trains participants in making a household budget and tries to dissociate the utility of keeping a budget with the nature of the income; it being a common belief among people that only those with regular and surplus income can keep a household budget. Instead the video brings out how budgeting can be especially useful for those with small incomes to bring down unnecessary expenditure and meet unforeseen expenses.

Session 2: Savings. Building on the previous session, the savings session begins by introducing the audience to the plight of Ramiben, a vegetable vendor who is caught in a debt trap given her spendthrift habits and inability to appreciate the need of accumulating small savings. Apart from educating the audience on the need of savings, the session dwells on the merit of saving in a bank vis a vis home and is a comprehensive guide on the various savings instruments present in the market.

Session 3: Loans. The session addresses three primary questions, namely: what to take a loan for (productive versus unproductive reasons), where to take a loan from (bank and

MFI versus moneylenders) and how to cost a loan (comparing interest rate versus comparing interest rates, accounting for hidden costs like documentation fee and other terms of the loan that are likely to impact its cost).

Session 4: Insurance. The session begins by introducing the audience to various kinds of risks one faces in life and how insurance can act as a “shield,” protecting us against life’s uncertainties. The video further talks about the various types of insurance available in the market and the companies one can approach to purchase insurance. Insurance is a complex product and choosing a policy that is best suited to one’s needs can be baffling given the dizzying array of options available in the market. The video therefore attempts to explain the design an insurance product; it cost components and factors that can affect the latter.

Session 5: A Concluding Video. “If my neighbor or friend can do it, so can I”—this is the essence of the last video which seeks to instill confidence in the audience’s ability to put in practice the lessons in financial management and planning taught in the last month. Interviewing people from the slums who practice budgeting and savings, who exercise discretion at the time of taking out a loan and who hold an insurance policy, the video aims to highlights how these people have been able to improve their lives with better financial management despite small and erratic incomes.

Health Education Videos

Session 1: Cleanliness & Hygiene. It talks about the essentials of cleanliness and hygiene like washing hands with soap, drinking portable filtered water; using toilets for defecation. We put these forth as basic but extremely crucial aspects of personal health and hygiene that can help keep households disease-free.

Session 2 & 3: Dai (midwife) & Maternal and Child Health. Session two and three discuss various issues relating to maternal heath and child care, both during pregnancy and after childbirth. It highlights the importance of monitoring pregnancy through regular ante-natal check-ups (ANC) and using the services of a trained midwife or a doctor for delivering the baby. Breastfeeding and immunization are described as crucial aspects of child health. Besides, since diarrhea is a common problem faced by infants, the sessions educate them on how to deal with it and what are the immediate remedies.

Session 4: Condoms, AIDS & Syphilis. This session focuses mainly on sexually transmitted diseases like AIDS & Syphilis. It gives detailed information on how these diseases can be contracted and what precautions need to be taken in order to protect ones family. The video also touches upon various myths related to sexually transmitted diseases and attempts to sensitize audience to the need of seeking informed advice on the subject. Reference is also made on how newborns can be saved from these diseases, in spite of their mothers suffering from the same. Towards the end, use of condoms as a means to prevent these diseases is stressed upon.

Session 5: Night Blindness. Night Blindness is common amongst children and pregnant women. The session cautions the audience against the various myths about the disease and suggests how simple measures like a regular diet rich in Vitamin A and iron can cure the ailment.

B.2 Financial Knowledge Survey Questions

Financial Numeracy Skills

1. Assume you purchased a health insurance policy on the 1st of January and you suffer an insurable loss of Rs. 1000 on 31st December due to an accident. Would you be better off if you had purchased an insurance policy with

   A. Rs. 3,000 cover and Rs. 950 premium

   B. Rs. 2,000 cover and Rs. 900 premium

2. If you had the choice, would you prefer to

   A. Receive Rs. 70 in cash 10 months from now

   B. Save Rs. 50 at 5 percent interest per month for 10 months

3. Suppose you had Rs. 50 to save. You could either save this for 1 month in an account which earns 14 percent interest per month, or save it for 1 month in an account that earns 2 percent interest per week. Which would you choose, 14 percent per month or 2 percent per week?

4. Assume you have purchased a medical insurance policy and suffer an accident which results in Rs. 3500 of hospital fees. Would you be better off if you had purchased an insurance policy with

   A. Rs. 3,000 cover and Rs. 950 premium

   B. Rs. 2,800 cover and Rs. 800 premium

5. We would like to tell you a short story about the income and expenditures of a tailor. We would then like you to use this sheet (give worksheet) to determine if in a month, this tailor is saving money or if his monthly expenditures exceed his monthly income. Jerembhai is a tailor in Vasna. Each week he makes Rs. 1500 from his work. He also sells the scraps from his work, for this he earns Rs. 200 each week. Each month Jerembhai must pay Rs. 1000 for the rent of his shop. He also spends Rs. 200 per week on his food and household goods. In addition to this he spends about Rs. 50 per

week on tea and snacks. He must pay Rs. 500 each month for the education expenses of his children. Some time age, Jerembhai took a loan to purchase his sewing machine. He pays an installment of Rs. 250 each week for this loan. He also pays Rs. 150 per month for a life insurance policy.

Basic Financial Awareness

1. Shantiben is preparing a budget for her household. Which of the following needs to be included in the budget?

   A. Income only

   B. Expenses only

   C. Both

2. Do you think you can open a savings account in a bank with amount as low as Rs 50 or 100?

3. Sukhiben's expenses are more than her income. Her friend Najmabanu tells her that writing a budget can help bring down her unnecessary expenses. Do you agree with Najmabanu or not?

4. Suppose I have a savings account in a bank and the bank closes down for some reason, will I get my money back?

5. Nileshbhai recently bought accident insurance with Rs 10,000 cover. The next day, he met with an accident and had to be hospitalized. He incurred Rs. 5,000 in hospital fees. How much do you think the medical insurance policy will pay for?

6. Iqbalbhai is 20 years old and Ashokbhai is 30 years old. If they were to buy life insurance of Rs 1 lakh for 20 years, who between the two to your mind will have to pay higher premium?

7. Manojbhai recently borrowed some money from a local moneylender. He wanted to buy some clothes for his children for Diwali (festival). What do you think about Manojbhai's loan?

Financial Attitudes and Perceptions

1. Rameshbhai does plastering on tall buildings. It is a dangerous job and he is worried that if he gets injured his family's income will become inadequate to meet their needs. If Rameshbhai comes to you for advice what would you suggest?

   A. Quit job

   B. Purchase health/life/ accident insurance

   C. Increase savings

2. Vimlaben has a very bright child who is currently in secondary school, but will probably do well in university. She is worried how her family will pay for the child's education. If Vimlaben comes to you for advice what would you suggest?

   A. Buy child life insurance policy

   B. Borrow money from a moneylender

   C. Open a savings account in a Bank

   D. Save at home

   E. Discontinue education

   F. Other

3. Kashiben has two sons. Her husband and two sons are earning members of the household and contribute towards household income. However Kashiben does not know what is the household's total income and expenditure. How do you think Kashiben can track her income and expenditure?

   A. Open a savings account

   B. Start making a household budget

   C. Buy life insurance for her husband and sons

4. Nareshbhai currently drives a rented auto rickshaw. He wants to purchase his own auto rickshaw but does not have the money and is considering taking out a loan for the same. If Nareshbhai comes to you for advice what will you suggest? Should he take out a loan or should he not?

5. Sajidbhai recently got married. He and his wife are considering buying a TV. They do not have enough savings and will need to take out a loan. Sajidbhai has two options: (1) He can take a loan from the moneylender and a relative and get a bigger amount in loan to buy a big TV, or (2) He can take a loan only from a relative and buy a smaller TV. What would you advise Sajidbhai and his wife?

B.3 Additional Regression Results

Table B.1: Short-term Impact on Financial Numeracy, Individual Questions

	Rs. 3000 cover, Rs. 950 premium vs. Rs. 2000 cover, Rs. 900 premium (1)	Rs. 70 ten months from now vs. Rs. 50 at 5% per month for 10 mos. (2)	14% per month vs. 2% per week (3)	Rs. 3000 cover, Rs. 950 premium vs. Rs. 2800 cover, Rs. 800 premium (4)	Wrote budget correctly (5)
Financial Education	-0.033	-0.013	0.046	-0.051	0.009
	(0.035)	(0.037)	(0.041)	(0.037)	(0.031)
Pay for Performance	-0.012	0.087*	-0.034	-0.016	-0.022
	(0.038)	(0.044)	(0.034)	(0.047)	(0.029)
Interaction of Financial Education and Pay for Performance	0.053	-0.041	-0.021	0.017	0.021
	(0.055)	(0.054)	(0.053)	(0.058)	(0.039)
R-squared	0.133	0.150	0.136	0.135	0.237
Number of Observations	1256	1256	1256	1256	1256
Mean of Dependent Variable in Control Group	0.422	0.686	0.701	0.686	0.735
F-test p-value: Financial Education + Interaction = 0	0.591	0.164	0.549	0.481	0.296

Notes: This table presents regression results on individual questions on financial numeracy from a survey conducted three weeks after the conclusion of the financial education program. The table shows intention-to-treat effects. *Financial Education* is a dummy equal to 1 for an individual who was invited to the financial education treatment. *Pay for Performance* is an orthogonal treatment and is a dummy equal to 1 for an individual who was offered a monetary incentive for correct answers to financial knowledge questions. Results are reported with robust standard errors clustered at the wave-class level. All regressions include monthly discount rate at baseline as well as strata dummies, where strata are defined by gender, chali (neighborhood), and microfinance borrower status. *** indicates statistical significance at the 1% level, ** at the 5% level, * at the 10% level.

Table B.2: Short-term Impact on Financial Awareness, Individual Questions

	Knows to include both income and expenses in HH budget	Knows can open an account with as low as Rs. 50	Knows about bank processing fees	Agrees that budgeting can help decrease unnecessary expenditure	Knows will get money back if bank closes	Knows insurance cover	Knows older person pays higher life insurance premium	Knows borrowing for Diwali is unproductive loan
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Financial Education	0.035	0.156***	0.131***	0.009	0.007	-0.020	0.058	0.196***
	(0.030)	(0.040)	(0.037)	(0.018)	(0.047)	(0.035)	(0.043)	(0.038)
Pay for Performance	-0.011	-0.024	0.078	-0.031	0.005	0.025	-0.026	0.014
	(0.046)	(0.039)	(0.063)	(0.024)	(0.043)	(0.055)	(0.063)	(0.049)
Interaction of Financial Education and Pay for Performance	0.034	0.026	-0.081	0.054*	0.004	0.069	-0.010	-0.008
	(0.051)	(0.047)	(0.067)	(0.028)	(0.054)	(0.070)	(0.072)	(0.061)
R-squared	0.134	0.137	0.163	0.095	0.116	0.113	0.124	0.200
Number of Observations	993	993	993	993	993	993	993	993
Mean of Dependent Variable in Control Group	0.843	0.669	0.614	0.958	0.705	0.554	0.566	0.620
F-test p-value: Financial Education + Interaction = 0	0.047	0.000	0.286	0.002	0.758	0.360	0.416	0.001

Notes: This table presents regression results on indivudal questions on financial awareness from a survey conducted three weeks after the conclusion of the financial education program. The table shows intention-to-treat effects. *Financial Education* is a dummy equal to 1 for an individual who was invited to the financial education treatment. *Pay for Performance* is an orthogonal treatment and is a dummy equal to 1 for an individual who was offered a monetary incentive for correct answers to financial knowledge questions. Results are reported with robust standard errors clustered at the wave-class level. All regressions include monthly discount rate at baseline as well as strata dummies, where strata are defined by gender, chali (neighborhood), and microfinance borrower status. *** indicates statistical significance at the 1% level, ** at the 5% level, * at the 10% level.

Table B.3: Short-term Impact on Financial Attitudes, Individual Questions

	Would suggest purchasing insurance to construction worker friend	Would suggest opening a bank account to friend w/ bright child	Would suggest making HH budget	Would suggest taking out a loan to friend who rents an auto	Would suggest taking out one loan and buy smaller TV
	(1)	(2)	(3)	(4)	(5)
Financial Education	0.104**	0.034	0.196***	0.057	0.019
	(0.045)	(0.038)	(0.066)	(0.044)	(0.028)
Pay for Performance	-0.016	0.031	-0.082	-0.015	-0.003
	(0.063)	(0.071)	(0.076)	(0.059)	(0.039)
Interaction of Financial Education and Pay for Performance	-0.021	-0.011	0.047	-0.007	-0.003
	(0.071)	(0.078)	(0.088)	(0.059)	(0.043)
R-squared	0.189	0.126	0.200	0.134	0.134
Number of Observations	591	591	591	591	591
Mean of Dependent Variable in Control Group	0.762	0.851	0.515	0.921	0.950
F-test p-value: Financial Education + Interaction = 0	0.109	0.693	0.001	0.151	0.608

Notes: This table presents regression results on individual questions on financial attitudes from a survey conducted three weeks after the conclusion of the financial education program. The table shows intention-to-treat effects. *Financial Education* is a dummy equal to 1 for an individual who was invited to the financial education treatment. *Pay for Performance* is an orthogonal treatment and is a dummy equal to 1 for an individual who was offered a monetary incentive for correct answers to financial knowledge questions. Results are reported with robust standard errors clustered at the wave-class level. All regressions include monthly discount rate at baseline as well as strata dummies, where strata are defined by gender, chali (neighborhood), and microfinance borrower status. *** indicates statistical significance at the 1% level, ** at the 5% level, * at the 10% level.

Table B.4: Longer-term Impact on Financial Knowledge, Individual Questions

	Financial Numeracy		Financial Awareness				Financial Attitudes		
	(1) Financial return comparison	(2) Interest rate calculation	(3) Knows to include both income and expenses in HH budget	(4) Knows can open an account with as low as Rs. 50	(5) Knows will get money back if bank closes	(6) Knows borrowing money for Diwali is unproductive	(7) Would suggest purchasing insurance to construction worker friend	(8) Would suggest opening bank account to friend w/ bright child	(9) Would suggest making HH budget
Financial Education	-0.022	-0.019	0.071***	0.160***	0.045	0.141***	0.058**	-0.011	0.240***
	(0.040)	(0.035)	(0.023)	(0.038)	(0.037)	(0.047)	(0.025)	(0.036)	(0.046)
Pay for Performance	-0.084	0.026	-0.036	0.001	-0.012	-0.052	0.060*	-0.086**	-0.048
	(0.054)	(0.041)	(0.031)	(0.045)	(0.036)	(0.060)	(0.035)	(0.033)	(0.044)
Interaction of Financial Education and Pay for Performance	0.090	-0.029	0.055	0.013	0.040	0.096	-0.077*	0.103**	0.045
	(0.066)	(0.050)	(0.037)	(0.054)	(0.051)	(0.074)	(0.040)	(0.043)	(0.058)
R-squared	0.141	0.132	0.164	0.217	0.149	0.162	0.119	0.176	0.238
Number of Observations	972	972	972	972	972	972	972	972	972
Mean of Dependent Variable in Control Group	0.655	0.786	0.851	0.625	0.702	0.548	0.815	0.821	0.565
F-test p-value: Financial Education + Interaction = 0	0.168	0.208	0.000	0.000	0.052	0.000	0.562	0.008	0.000

Notes: This table presents regression results on individual questions on financial numeracy, awareness, and attitudes from an endline survey conducted ten months after the conclusion of the financial education program. The table shows intention-to-treat effects. *Financial Education* is a dummy equal to 1 for an individual who was invited to the financial education treatment. *Pay for Performance* is an orthogonal treatment and is a dummy equal to 1 for an individual who was offered a monetary incentive for correct answers to financial knowledge questions. Results are reported with robust standard errors clustered at the wave-class level. All regressions include monthly discount rate at baseline as well as strata dummies, where strata are defined by gender, chali (neighborhood), and microfinance borrower status. *** indicates statistical significance at the 1% level, ** at the 5% level, * at the 10% level.