Essays in Development and Health Economics

HARVARD UNIVERSITY
Graduate School of Arts and Sciences

DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the Committee for the PhD in Business Economics have examined a dissertation entitled

Essays in Development and Health Economics

Presented by Heather Schofield

candidate for the degree of Doctor of Philosophy and hereby certify that it is worthy of acceptance.

Signature ______________________

Sendhil Mullainathan, Chair

Signature ______________________

Shawn Cole

Signature ______________________

David Cutler

Signature ______________________

Larry Katz

Signature ______________________

Michael R. Kremer

Date: 4/14/14

Essays in Development and Health Economics

A dissertation presented

by

Heather Schofield

To

The Committee on Degrees in Business Economics

in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Business Economics

Harvard University
Cambridge, Massachusetts

May 2014

Dissertation Advisor: Professor Sendhil Mullainathan | Heather Schofield

Essays in Development and Health Economics

Abstract

In this collection of essays I study individuals' health related decision-making and the consequences of those decisions for health and labor market productivity.

The first two essays focus on the impact of low caloric intake on economic productivity in India and find that changes in caloric intake result in substantial and broadly generalizable changes in productivity among malnourished adults. The first of these essays, based on a randomized controlled trial of cycle-rickshaw drivers, finds that an additional 700 calories per day results in improved physical and cognitive function in the laboratory as well as a 10 percent increase in labor supply and earnings by the fifth and final week of the study. The second essay studies the impact of a decline of 700 calories per day caused by Ramadan fasting on agricultural production. The estimated decrease in production implies a 20 to 40 percent loss in productivity per fasting individual.

In both of these studies, the estimated return on investment in calories is relatively high, with point estimates of 75 to 200 percent over a few months. Yet, substantial evidence suggests that liquidity constraints do not meaningfully limit caloric consumption. Hence, the low caloric intake of the majority of Indian adults presents a puzzle. I study choices regarding caloric intake through incentivized surveys and find evidence that inaccurate beliefs about both the returns to calories and the caloric content of foods may play a role in the low caloric consumption observed in India.

The third essay examines the impact of individually oriented, purely altruistic, and a hybrid of competitive and cooperative monetary incentives on older adults' completion of cognitive exercises in the United States. This research finds that all three incentive structures approximately double the number of exercises completed during the six-week active experimental period relative to a no incentive control condition. However, the altruistic and cooperative/competitive designs led to different patterns of participation, with significantly higher inter-partner correlations in software utilization, as well as greater persistence once incentives were removed. Provision of all incentives improved performance on the incentivized exercises but not on an independent cognitive testing battery.

Contents

ACKNOWLEDGEMENTS

I would like first to thank my numerous advisors for their invaluable contributions which I appreciate more than I can say. I am deeply grateful to Prof. Sendhil Mullainathan for his continued guidance which challenged me to think about the world with greater openness and clarity. Thank you to Prof. Michael Kremer for your encouragement to pursue questions which are not always easy to answer, but which push us to view the world around us in a new light. Sincere thanks to Prof. Larry Katz for his ever-thoughtful, constructive, and insightful feedback which helped me to formulate and convey my ideas more effectively. David Cutler, a resounding thank you for your productive comments, unfailing encouragement, and steadfast generosity in discussing both these papers and career development. To Prof. George Loewenstein and Prof. Kevin Volpp, a world of thanks for expanding my horizons and instilling confidence in my own creativity; your faith in me gave me faith in myself. You have all served as examples to aspire to and I am incredibly grateful. I have also benefited greatly from conversations with Profs. David Laibson, Ed Glaeser, Shawn Cole, Rohini Pande, Asim Khwaja, David Bloom, Rick Hornbeck, and Max Bazerman and thank them for their time and keen insights. A hearty thank you to Kate Sturla, Luke Ravenscroft, and all of the research staff in Chennai for their outstanding field assistance. I would also like to thank my parents Janet and Doug and my sisters Alanya and Emily for their warm love that was always there when I needed it most. You've taught and encouraged me to strive for a life with purpose and have given me the tools and encouragement to take on a challenge rather than shy away from it. Thank you for helping me become the person I am today. To Joana Naritomi, Ian Tomb, Raluca Dragusanu, Dan Bjorkegren, Supreet Kaur, Frank Schilbach, Laura Trucco, Clara Zverina, Arash Nekoei, Francisco Queiro, and Anjali Aduika I am forever grateful to my wonderful friends from this PhD program for your laughter, love, support, and the conversations and moments that I will carry with me always.

1 The Economic Costs of Low Caloric Intake: A Randomized Controlled Trial with Cycle-Rickshaw Drivers*

1.1 Introduction

Many of the world's poor consume very few calories.[1] One-seventh of the world's population remains below the level of caloric intake recommended by health professionals (FAO 2011). Of course, the poor consume less of any normal good. However, food differs from many goods in that it is both a consumption good, producing utility directly, and an input into production. A long line of theory literature has modeled intertemporal nutrition choices and their implications for labor productivity and the functioning of labor markets (e.g., Leibenstein 1957; Bliss and Stern 1978; Stiglitz 1976; Dasgupta and Ray 1986). Although theory in this area is extensive and well developed, given the inherent challenges posed by the endogeneity of caloric intake and the measurement of productivity, empirical work is less well developed.[2] Specifically, despite the fact that calories are a broadly available investment good and a basic economic choice made by all people, it remains an open question whether there is an economically significant calorie-productivity gradient at the levels of caloric intake observed among the world's poor today.

Theoretically, there are reasons to believe the answer is no. There is good evidence that liquidity is unlikely to constrain investment in calories.[3] Hence, a revealed preference argument suggests that despite

* I am deeply grateful to my advisors Sendhil Mullainthan, Lawrence Katz, Michael Kremer, and David Cutler for feedback and support. For helpful discussions, I thank David Laibson, Rohini Pande, Asim Khwaja, Rick Hornbeck, Ed Glaeser, George Loewenstein, Kevin Volpp, Shawn Cole, Max Bazerman, David Bloom, Supreet Kaur, Frank Schilbach, Dan Bjorkegren, Ian Tomb, Joana Naritomi, Raluca Dragusanu, Laura Trucco, Anjali Adukia and seminar participants at Harvard University. I thank Kate Strula, Luke Ravenscroft, Manasa Reddy and the research staff in Chennai for outstanding field assistance.

[1] Throughout this document I use the word "calorie" (with a lower case "c") in the commonly used sense of 1 kilocalorie rather than the technical definition in which 1 calorie is 1/1000 of one kilocalorie or 1/1000 of one Calorie.

[2] In addition, the existing empirical evidence is quite mixed. For example, Deolalikar (1988) and Immink and Viteri (1981) find no relationship between nutrition and productivity while Kraut and Muller (1946), Strauss (1986), and Wolgemuth et al. (1982) find evidence in favor of a such a relationship. A more detailed discussion of existing productivity and physiology literatures are provided in Appendix 1B.

[3] Even among extremely poor individuals, food rarely accounts for more than 40 to 70 percent of expenditures and the additional calories necessary to reach recommended caloric intake could typically be purchased for less than five

the apparently low consumption, any productivity gains from additional caloric intake are likely to be relatively small. Specifically, with perfect markets and full information, the Euler equation implies that the return to calories must be less than the discount rate. Yet, there are both structural and behavioral reasons why the revealed preference argument may not hold. For example, principal-agent contracting may reduce consumption below its optimal level if consumption is not readily observable. Or, if individuals are not aware of the returns to calories or the nutritional content of foods, they may consume too little, potentially leaving large economic gains on the table even in the absence of liquidity constraints.

In the first two chapters of this dissertation I study the relationship between caloric consumption and economic output using two distinct empirical methodologies. The first, presented in Chapter 1, is a field experiment (with laboratory components) which examines the impact of increased caloric intake on the labor supply, earnings, and physical and cognitive performance of 211 cycle-rickshaw drivers in Chennai, India.[4] The second analysis, presented in Chapter 2, draws on the quasi-random declines in caloric intake caused by fasting during Ramadan, a month-long Muslim holiday. This analysis utilizes a differences-in-difference approach to assess the impact of reduced caloric intake on agricultural production at the crop-district-year level. These two methodologies are complementary. While field experiments can provide strong causal identification, they often lack external validity. Conversely, natural experiments typically apply to broader populations, but may not establish causality as cleanly. In drawing on both methodologies, I aim to establish a casual mechanism in a controlled setting and to demonstrate that the mechanism has economically meaningful impacts on one of the largest sectors of the Indian economy.

In more detail, the 211 cycle-rickshaw drivers in the five-week randomized controlled trial (RCT) presented in Chapter 1 were randomly assigned to either a control group, receiving cash for their participation, or a treatment group receiving a mixture of cash and food of the same total value. The foods

percent of daily income. Alternatively, substitution to less expensive grains could increase caloric intake 20 percent at no additional cost (Subramanian and Deaton 1996; Banerjee and Duflo 2011). In addition, the elasticity of caloric intake with respect to income is remarkably low in many developing countries, with estimates commonly ranging from 0.1 to 0.5 (Subramanian and Deaton 1996; Behrman and Deolalikar 1987; Bouis and Haddad 1992).

[4] Cycle-rickshaw drivers drive large human-powered tricycles similar to "bicycle taxis" in the United States. The vehicle has a bench large enough for two passengers behind the driver.

provided to treated participants were snacks with little nutritive value beyond 700 calories.[5] In order to minimize crowd out and generate a substantial and sustained increase in caloric intake, participants could choose from a variety of snacks which were consumed daily between meal times in the office.[6]

While only treated participants received food, participants in both experimental conditions visited the study office briefly each day to report their labor supply and earnings. The study also included a battery of physical and cognitive laboratory-based tasks, compensated according to performance, at both enrollment and in the final week of the study ("endline"). These tasks were designed to assess the underlying physical and cognitive skills relevant to labor supply in this population in the absence of the variable demand that is common in this labor market.[7] For example, physical tasks included cycling on an exercise bike in the office and cognitive tasks targeted skills such as persistence, motivation, and planning.

A number of features of cycle-rickshaw drivers made them an advantageous population for a study of this type. First, as full residual claimants on their labor and with flexible labor supply, this population faces strong incentives and has the ability to adjust labor supply and earnings over short time horizons. The rickshaw drivers' mobility also allowed them to provide high frequency data via the daily visit to the study office while minimizing the likelihood of spillovers between conditions. Finally, the work habits, demographic characteristics, and low caloric intake of the cycle-rickshaw drivers studied are relatively representative of typical informal labor market participants.

Using an intent-to-treat estimation strategy, the results from this study suggest that increasing caloric intake increases both labor supply and earnings over time. Participants in the treatment group increase their labor supply relative to the control group by an average of eight percent over the entire study. Consistent with the cumulative effects found in the physiology literature, this increase follows a fairly linear pattern across the weeks of the study, rising from no significant difference in the first week to a net gain of

[5] To determine the effect of calories rather than general "nutrition," the snacks were chosen to contain minimal iron and few micronutrients. Specifically, snacks typically consisted of starches and fried foods like potato chips and samosas.
[6] Multiple outcomes, including weight measured at endline, confirm that despite the inframarginal amount of food provided, caloric intake increased among treated participants. See Section 1.2 for additional details of these measures.
[7] Day-to-day demand for rickshaw services varies substantially with the weather, train schedules, and other factors.

roughly 12 percent in the final week (Keys et al. 1950). Income streams in this population are noisy and have relatively large standard errors; however, earnings show a similar linear trend and are roughly 9 percent higher among treated individuals in the final week of the study. Finally, treated participants also earn approximately 10 percent more on both physical and cognitive laboratory-based tasks at endline.

Although cognitive function has not traditionally been central in the economics literature on adult nutrition, the physiology and psychology literatures suggest that improvements in cognitive function are not only possible, but likely, with changes in caloric intake.[8] If these improvements occur, the returns to improved nutrition would not be limited to physical gains on the margin, but would also include gains from improving inframarginal choices as well as savings and consumption decisions. Hence, changes in earnings would provide a lower bound on the gains from higher caloric intake.[9]

The meaningful productivity gains found in the randomized trial of rickshaw-drivers prompts the question of whether the economic returns to additional caloric investment are positive. Although there is greater uncertainty inherent in estimating the return on investment (ROI), with a point estimate of 75 percent over 6 months, the estimated return is relatively high.[10] Yet, one-third of the population remains underweight and caloric intake in India is both low and declining, generating a puzzle (WHO 2013; Deaton and Dreze 2009).

[8] Despite accounting for only 2 percent of body weight, the brain consumes roughly 20 percent of the energy used by the body; hence, limiting the total energy available is likely to constrain the brain's ability to function as it would any other organ requiring substantial energy inputs (Fonsec-Azevedo and Herculano-Houzel 2012). In addition, a number of diverse sources have demonstrated substantial changes in cognitive function and decision-making as a function of caloric intake (Gailliot et al. 2007; Danzinger et al. 2011; Baumeister and Vohs 2007; US Army 1987).

[9] Additionally, the influence of low caloric intake could be much more broadly applicable and could remain relevant despite the global shifts towards cognitive rather than physical labor.

[10] These ROI calculations use a common staple grain, rice, to determine the cost of investment. Due to the cumulative nature of changes in caloric intake and the corresponding linear increases in labor supply and earnings over the course of the five week RCT, I estimate a six-month ROI for this study to better gauge the likely "long-run" impact of higher intake. The expected gains in earnings beyond the end of the study are, however, capped at the value observed in the final week. A wide variety of alternative assumptions about the cost of calories and productivity changes, and the returns implied by these assumptions, are detailed in Section 1.4. Notably, although the estimated returns are high, given the low cost of investment, the absolute value of the gains is relatively small and accounts for only a modest fraction of income. This feature may make the high returns relatively difficult for individuals to detect in the context of highly variable income streams. Returns to caloric intake in the study of agricultural production presented in Chapter 2 are similarly high.

To explore this puzzle and the forces driving decisions about caloric intake, I briefly describe and assess possible reasons for low caloric intake in this population and conclude that incorrect beliefs about nutrition may play an important role in low intake levels. Evidence of these inaccurate beliefs was generated via an incentivized survey about the returns to increased caloric intake and the caloric densities (calories per unit cost) of foods. Responses to these surveys suggest strikingly wrong beliefs. For example, over three-quarters of the respondents believe that increasing caloric intake by the equivalent of one meal per day would either weakly decrease labor supply and earnings, or have no prior about the consequences of increasing caloric intake. And, in pairwise comparisons between commonly consumed foods, participants identify foods with higher caloric density at rates lower than those generated by guessing at random.[11] Because the existence of incorrect beliefs in equilibrium is unusual, supporting evidence for how these beliefs are developed and maintained is also provided.[12]

The remainder of this paper is divided into five parts. Section 1.2 describes the design of the randomized controlled trial among cycle-rickshaw drivers. The changes in labor supply, earnings, and performance on laboratory tasks resulting from increased caloric intake are presented in Section 1.3. Section 1.4 provides calculations of the returns associated with the productivity changes estimated in Section 1.3. Potential reasons for low caloric intake observed in the presence of high returns are discussed in Section 1.5. In addition, this section provides evidence that incorrect beliefs may play a role in the decision not to invest in greater caloric intake. Section 1.6 concludes and discusses implications of incorrect beliefs regarding caloric intake.

[11] Because calories per rupee would be a difficult metric for semi-numerate individuals, these questions are implemented by asking participants to select which of two food items with the same economic value has more energy. Food items were specifically selected to be commonly consumed and to have large caloric differences. The tradeoffs were presented as life-size photographs with verbal and written descriptions of the food accompanying them. More details on the exact implementation are provided in Section 1.5.

[12] Although unusual, there are other well-documented instances in which incorrect beliefs with significant economic consequences exist in equilibrium (e.g., Jensen 2010; Banerjee and Duflo 2011).

1.2 Experimental Design

There are a number of clear challenges inherent in assessing the productivity consequences associated with caloric intake. Use of a randomized controlled trial among cycle-rickshaw drivers provides solutions to two of the most difficult: 1) the endogeneity of caloric intake and 2) measurement of individual productivity in a context with strong incentives and minimal spillovers. The endogeneity of caloric intake was addressed through randomization and design features intended to minimize crowd out of other food consumption and generate substantial and sustained increases in caloric intake among treated individuals. Selection of a participant group with a variety of advantageous features, described below, aided in the measurement of individual productivity. In addition, the use of laboratory tasks allowed for the assessment of changes in both physical and cognitive ability under controlled conditions with clear and consistent incentives.

1.2.1 Participants

Participants were drawn from the population of cycle-rickshaw drivers in Chennai, India. Eligibility criteria included a BMI less than or equal to 20, age greater than 18 years, expected residence in the city for at least two months, and cycle-rickshaw driving as a primary occupation.[13] Basic demographic characteristics and labor supply information collected at enrollment are presented in Table 1.1 (a more extensive discussion of participants' baseline characteristic and habits is presented in Appendix 1B). This population was chosen because: 1) as full residual claimants on their labor, rickshaw drivers face strong incentives, 2) the work has flexible labor supply (both days and hours) and allows for direct measurement of individual labor supply and earnings, 3) the population is mobile allowing them to visit the study office each day while limiting the possibility of spillovers between treated participants and control participants, and 4) members typically

[13] BMI is a measure of weight for height calculated as weight in kilograms divided by height in meters squared. The WHO describes a BMI under 18.5 as "underweight," 18.5 to 24.9 as "normal," 25 to 29.9 as "overweight," and over 30 as "obese" (WHO 2013). As a benchmark, an individual who is 5 foot 8 inches tall (1.73 m) would have a BMI of 20 at a weight of 131 pounds (59.4 kg).

have low caloric intake. In addition, these individuals were selected because they are similar to many other casual laborers in India, improving external validity.

Table 1.1: Characteristics of RCT Participants at Enrollment

	(1)	(2)	(3)
Experimental Condition	Control	Treatment	Difference
BMI	17.87	17.73	-0.14
	[1.41]	[1.39]	[0.19]
Age	45.63	46.26	0.63
	[10.94]	[9.37]	[1.40]
Number of household members	3.42	3.40	-0.02
	[2.32]	[2.28]	[0.32]
Migrant (binary)	0.29	0.22	-0.07
	[0.46]	[0.42]	[0.06]
Rooms in house	1.41	1.44	0.03
	[0.96]	[1.07]	[0.14]
Number of small appliances	3.77	4.08	0.31
	[2.16]	[2.54]	[0.33]
Ration card (binary)	0.74	0.69	-0.05
	[0.44]	[0.47]	[0.06]
Years as rickshaw driver	20.09	20.64	0.55
	[11.53]	[11.84]	[1.61]
Rent rickshaw (binary)	0.77	0.76	-0.01
	[0.42]	[0.43]	[0.06]
Earned yesterday	183.28	193.53	10.25
	[145.52]	[160.59]	[21.15]
Work yesterday	0.77	0.82	0.04
	[0.42]	[0.39]	[0.06]
Hours, conditional on working yesterday	9.80	9.81	0.01
	[3.61]	[3.40]	[0.56]
Observations	102	109	211

Notes:
1. In Columns (1) and (2), results are presented as mean [standard deviation]. Column (3) is presented as mean [standard error].
2. The hours conditional on working figures are based on smaller sample sizes of N = 74 for Control and N = 82 for Treatment.
3. Rickshaws are typically rented on a weekly or monthly basis.
4. No significant imbalances on baseline covariates were found between participants in each experimental condition.

1.2.2 Experimental Conditions

Participants were randomly assigned to either a control group, receiving only cash compensation with an average value of Rs. 75 per day, or a treatment condition, receiving a mixture of cash and 700 calories worth of food with the same total value. Treated participants were able to choose a variety of snacks from a varied selection each day to encourage consumption. However, the food provided was of little nutritive value beyond the caloric content.[14] Food was provided during periods between meals to minimize crowd out and treated participants consumed the food in the office to ensure that it was not given to others or thrown away.

1.2.3 Experimental Activities and Timing

Each participant was enrolled in the study for five weeks. During this time, each participant was asked to visit the study office each day (except Sunday) to complete a short survey about labor supply and earnings the previous day.[15] In addition, participants reported the number of meals consumed (as a measure of potential crowd out of food consumption) and their energy level on a scale of one to five during this survey. Participants in the treatment group were also provided with food at this time.

All participants also spent one day on the first day of the study ("enrollment") and in the fifth week of the study ("endline") in the study office to complete more extensive surveys and engage in a battery of physical and cognitive tasks. These experimental tasks, compensated according to performance and conducted in a controlled setting, were designed to provide a comprehensive picture of the impact of improved nutrition on both physical and cognitive function relevant to production. For example, cognitive tasks targeted areas such as motivation and planning, both skills relevant to work as a rickshaw driver. The tasks were a mixture of standardized tasks used in the economics and psychology literatures (e.g. crossing out symbols) and original tasks specific to the context (e.g. choices between various rickshaw job offers,

---

[14] Options typically included foods such as potato chips, "mixture" (small bits of fried dough similar to Chex mix or Cheetos), vada (a fried spiced dough), or samosas (fried dough stuffed with potatoes, vegetables, and spices).

[15] Participants reported both Saturday and Sunday labor supply and earnings on Monday.

cycling on an exercise cycle). Finally, anthropometric measurements were taken and participants completed a 24-hour dietary recall survey to permit calculation of changes in participants' caloric intake. A list of tasks in included in Appendix 1A, Table 1A.1.

1.3 Results of the Randomized Controlled Trial

1.3.1 Randomization

The baseline randomization was successful with no significant imbalances on enrollment covariates for either demographics or labor supply and earnings (Table 1.1 above).

1.3.2 Caloric Intake

Dietary recall surveys, daily reports of meals consumed, and an increase in BMI of treated individuals all provide evidence that the treatment was effective at increasing caloric intake. Results of these analyses are presented in Appendix 1A, Table 1A.2 and discussed in Appendix 1B.

1.3.3 Attrition

Overall average attendance was 87 percent in the control group and 76 percent in the treatment group. These average figures shroud important variation in attrition patterns though. Treated individuals were significantly more likely to drop out during the first few days (Figure 1.1). However, conditional on attending more than four days in the first week, daily attendance was 93 percent among control individuals and 91 percent among treated individuals, a difference that is statistically indistinguishable (Table 1.2 and Figure 1.2, $p = 0.51$).

Figure 1.1: Daily Attendance

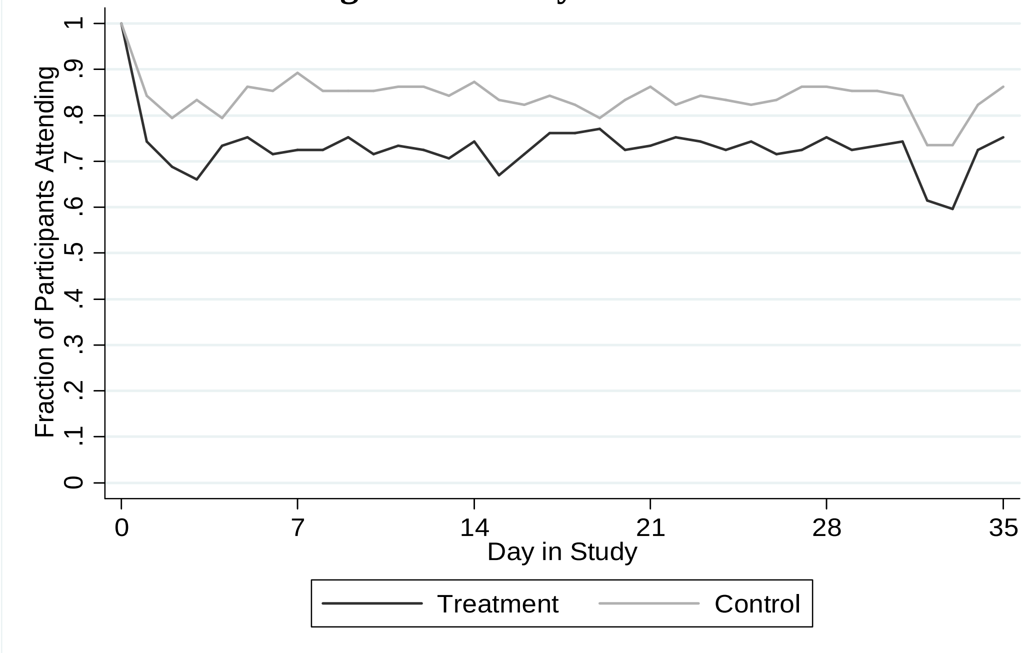
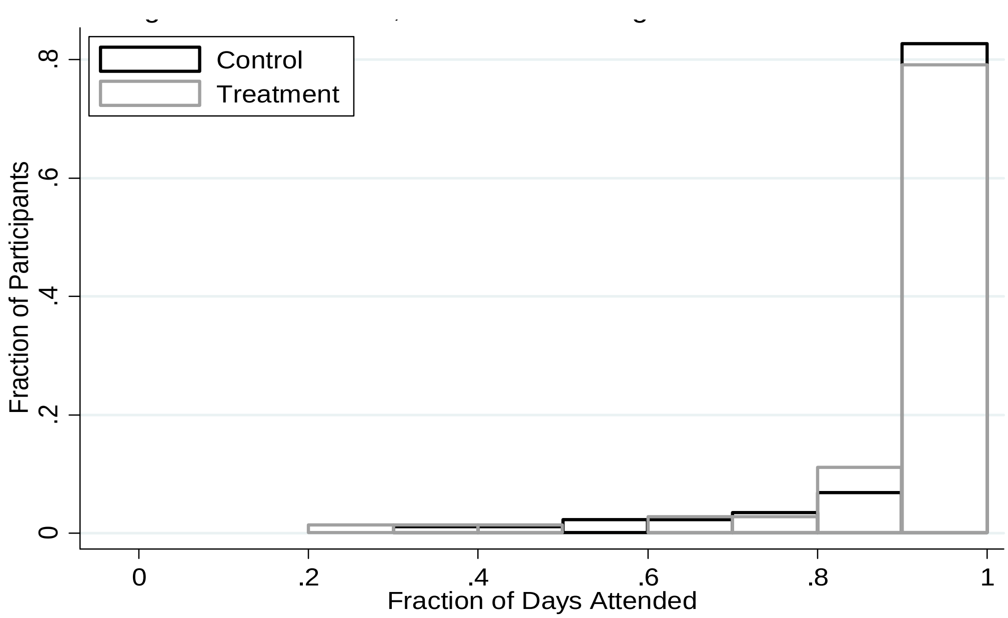


Figure 1.2: Attendance, Conditional on High 1st Week Attendance



Note: High 1st Week Attendance is defined as attending more than 4 days in the first week

Table 1.2: Attendance Levels By Experimental Condition

Experimental Condition	(1) Control	(2) Treatment
Average fraction of days attended	0.87	0.76
Fraction who ever return after enrollment	0.94	0.89
Fraction who attend > 25% of days	0.94	0.83
Fraction who attend > 50% of days	0.89	0.8
Fraction who attend >75% of days	0.82	0.69
Attend endline	0.9	0.81
Average fraction of days attended conditional on attending more than 4 days in the first week	0.93	0.91

Notes:
1. See Figure 1.1 for a graphical representation of attendance over time

Baseline covariates gathered at enrollment are useful in understanding the factors associated with attrition. At enrollment each individual was asked whether he had worked the day before, and if so, his earnings and hours. Of individuals who attrited, control participants reported a mean earnings of Rs. 125 while treated participants reported a mean earnings nearly twice as high, Rs. 237. Similarly, attriters in the control group reported lower labor supply.[16] These differences are statistically significant despite the large standard errors resulting from the small sample (Table 1.3). The direction of the selective attrition suggests that active and high earning individuals assigned to the treatment did not find the monetary compensation worthwhile, resulting in higher attrition rates.[17] The noisy nature of the baseline productivity measures limits the possibility of fully controlling for "type." However, if treatment effects are relatively homogeneous, the differential attrition is likely to bias the results of the study towards the null by lowering the average labor provision and earnings of the treated individuals. Nonetheless, because there may also be selection on unobservable characteristics, I also construct bounded estimates below.

[16] Attending fewer than 4 days in the first week is the measure of attrition used because of its strong predictive power for later attendance and for consistency with the figures above. However, results are similar in magnitude and direction if different measures of "attriter" (e.g. attending fewer than 10 percent of the days over the entire study) are used.
[17] Efforts were also made to contact attriters at their stand and inquire why they no longer wished to participate. It was not possible to locate all individuals who left the study. However, among those we were able to contact, insufficient compensation, driven by the lower value placed on the in-kind compensation, was the most commonly cited reason for leaving the study among treated individuals.

Table 1.3: Baseline Characteristics of Attriters

Dependent Variable	(1) Earn	(2) Work	(3) Hours if > 0	(4) BMI
Treatment	111.96**	0.33**	2.27	0.43
	[52.48]	[0.13]	[1.66]	[0.56]
Mean of dependent variable	138.20	0.79	9.82	17.79
Observations	38	38	30	37
R-squared	0.11	0.16	0.06	0.02

Notes:
1. Attrition is defined as attending fewer than 4 days in the first week as attendance is high (greater than 90% in both conditions) and relatively stable for those who attend at least four days in the first week. However, results are qualitatively similar if other definitions of attrition are used.
2. "Earn" is the total daily earnings as a rickshaw driver in Indian Rupees. "Work" is a binary variable indicating whether the participant worked that day. "Hours if > 0" is the total number of hours worked, conditional on working. Each of these variables was collected at enrollment referencing the day prior to enrollment. "BMI" is body mass index, a measure of weight for height.
3. Robust standard errors are in brackets.
4. *** p<0.01, ** p<0.05, * p<0.1

1.3.4 Labor Supply and Net Earnings

Evidence from the literature on nutritional deprivation demonstrates that the process of recovery from low caloric intake is cumulative and often slow, taking weeks or even months in cases of extreme deprivation (Keys et al 1950). Given the cumulative nature of the recovery process, the impact at later points in the study is likely to better reflect the long-run changes associated with the increased caloric intake. Hence, I begin by dividing the study at the midpoint and assessing the differences between the treatment and control groups in both periods. However, both week by week effects and average treatment effects over the course of the study are also estimated to more specifically gauge the path of the changes as well as the net impact of the intervention during the study.

Equation (1.1) estimates the impact of increased caloric intake on average labor supply and earnings, including earnings on enrollment and endline days, in the first and second halves of the study. In this equation, l_{in} denotes a measure of average earnings or labor supply for individual i in period n, T_i is an indicator for assignment to the treatment group, P_n is an indicator for the second half of the study, and X_i is a vector of controls for baseline work habits. The regression also includes fixed effects for stand location, θ_d, and enrollment month, γ_m, because enrollment was conducted on a rolling basis.

$$(1.1) \qquad l_{in} = \beta_0 + \beta_1 T_i + \beta_2 P_n + \beta_3 T_i{*}P_n + \alpha X_i + \theta_d + \gamma_m + \varepsilon_{in}$$

Results of these regressions are presented in Table 1.4. The first three columns examine labor supply outcomes. As can be seen in Column (1), there is an 8 percentage point increase in the average fraction of days worked in the second half of the study among treated individuals relative to control individuals. The point estimate for hours conditional on working is positive for treated individuals in the second half of the study. However, it is only marginally significant (Column (2)). The combined metric of total hours is presented in Column (3). Driven largely by the increase in days worked, there is an average increase of roughly one hour per day, or about 15 percent, for treated participants in the second half of the study. This substantial increase is particularly notable given the fairly long baseline work weeks of nearly 50 hours and the physical difficulty of the labor.

Although earnings are significantly noisier, the trends are similar. While only significant at the 10 percent level, Columns (4) and (5) show an increase in treated participants' average daily earnings of roughly 20 Rupees per day, or approximately 11 percent, in the second half of the study.[18]

[18] One potential concern about the gains among treated individuals relative to control individuals is that control participants could be decreasing labor supply because they have greater cash payments for coming to the study office each day. However, it is possible to test for reduced labor supply among controls by comparing baseline earnings reported at enrollment with earning during the study in this group. These measures are not significantly different, suggesting that labor supply and earnings are not reduced among control individuals and that the estimated differences are indeed increases among the treated individuals.

Table 1.4: Average Labor Supply and Earnings By Study Period

Dependent variable	(1) Average days worked	(2) Average hours if > 0	(3) Average hours	(4) Average earnings	(5) ln(Average earnings)
Treatment (first half of study)	0.02 [0.03]	-0.29 [0.32]	-0.13 [0.36]	-1.92 [9.16]	0.01 [0.06]
Second half of study	0.02 [0.02]	-0.35* [0.20]	0.09 [0.23]	3.61 [6.63]	0.00 [0.05]
Treatment*Second half of study	0.08*** [0.03]	0.73* [0.28]	1.07*** [0.33]	19.61* [10.88]	0.11* [0.06]
Mean of dependent variable	0.77	8.72	6.81	179.78	5.10
Observations	370	364	370	383	383
R-squared	0.14	0.24	0.19	0.15	0.15

Notes:
1. This table provides the results of regressions testing for changes in labor supply and earnings, including enrollment and endline day earnings, among treated participants. The unit of observation is the participant-period (first or second half of the study). Outcome variables are averaged over all observations the period for each participant. "Work" captures the average fraction of days worked in that period. "Hours if >0" is the average number of hours worked, conditional on working. "Total hours" is the number of hours worked per day where hours are zero if the participant did not work. "Earn" is the total daily earnings as a rickshaw driver summed with enrollment or endline earnings, as applicable to the period, in Indian Rupees. This variable takes the value zero if the participant reported not working. The exchange rate is roughly 45 Rupees to 1 US Dollar. "ln(Average earnings)" is the log of average earnings where earnings is defined as previously described.
2. All regressions include stand location and enrollment month fixed effects as well as controls for baseline work habits as detailed in Appendix 1B. Results are less precisely estimated, but qualitatively similar if these controls are omitted.
3. Robust standard errors clustered by individual are in brackets.
4. *** p<0.01, ** p<0.05, * p<0.1

1.3.5 Trends in Labor Supply and Earnings

In order to better understand the temporal trends and the forces driving the overall gains in earnings, I next examine the more finely grained temporal path of the changes and decompose the effect between earnings as a rickshaw driver and earnings in the laboratory.

To examine trends outside of the laboratory, I estimate Equation (1.2) which examines the week by week impact of treatment on labor supply and earnings as a rickshaw driver.[19] The dependent variables,

[19] In addition to this non-parametric specification, I also estimate a simple linear trend by day. Results of these regressions are included in Appendix 1A, Table 1A.3. Linear trends are significant at the 5 percent level for all outcomes presented in Table 1.5 except hours conditional on working, which is significant at the 10 percent level.

measures of labor supply and earnings, are denoted by l_{it} and are measured at the level of the participant-day beginning the day after enrollment (which was spent in the study office). T_i is an indicator for assignment to the treatment group, χ_W denotes the week in the study, X_i is a vector of controls for baseline work habits, and ω_S and θ_d are calendar week and stand location fixed effects, respectively.

(1.2) $$l_{it} = \beta_0 + \beta_1 T_i + \beta_{2w} T_i{*}\chi_W + \alpha X_i + \theta_d + \omega_S + \chi_W + \varepsilon_{it}$$

As can be seen in Table 1.5, Column (1), the increase in days worked among treated participants follows a fairly smooth trend over the course of the study, with an insignificant gain of 2 percent in the first week rising to a significant increase of 11 percent by the final week. Hours per day conditional on working rises more rapidly in the first few weeks of the study but then levels off with a slightly lower point estimate in the final week of the study (Column (2)). These two margins of adjustment are combined into a single metric of total hours per day, which takes the value zero if the participant did not work that day, in Column (3). This metric shows no significant change in the first week, but rises to a 12 percent net increase in total hours in fifth week relative to the control group. Figure 1.3 displays the change in days and total hours worked by treated individuals relative to control individuals by week in the study.

Earnings, once again, follow a similar pattern of increase, but begin with a slightly negative, although insignificant, point estimate. In fact, the initial negative point estimates drive a near zero impact on earnings over the duration of the study among treated individuals (See Appendix 1A, Table 1A.4 for overall estimates). However, these initial negative point estimates may be caused by the differential attrition among high earning individuals in the treatment group during first week.[20] This supposition is supported by Table 1A.5 in Appendix 1A, limiting the sample to individuals with at least 90 percent attendance (results are also presented graphically in Figure 1.4). Among these individuals, there is no

[20] In fact, the difference in earnings for treated individuals in the first week is similar to the difference expected based on the differential earnings between control and treated attriters and the attrition rates. While the baseline controls would be expected to account for these differences, the baseline measures are only moderately correlated with average outcomes during the study (r = 0.1 to r = 0.5) and hence may not adequately control for differences.

impact of treatment in the first week and an increase by the fifth week (Rs. 20 per day) that is slightly larger than that found in the non-restricted sample.

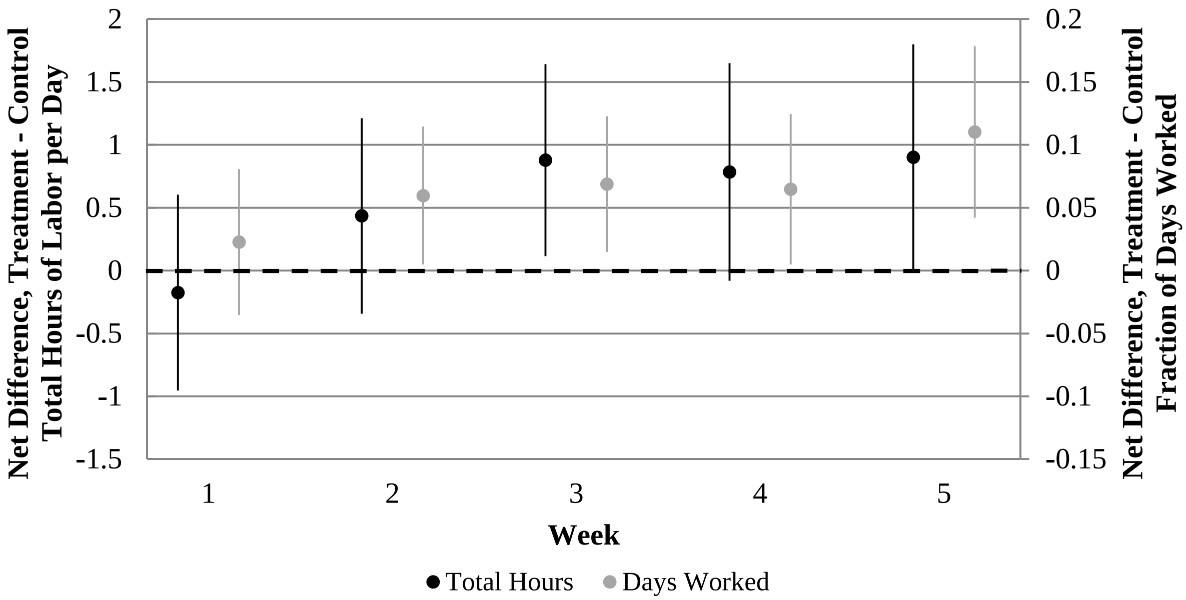
Table 1.5: Rickshaw Driver Labor Supply and Earnings by Week

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	Work	Hours if >0	Total hours	Earn	ln(earn+1)	IHS(earn)
Treated (Week 1)	0.02	-0.54	-0.18	-4.37	0.08	0.10
	[0.03]	[0.33]	[0.40]	[9.88]	[0.16]	[0.18]
Treated*Week 2	0.04	0.30	0.61*	-3.71	0.13	0.16
	[0.03]	[0.28]	[0.37]	[9.77]	[0.17]	[0.19]
Treated*Week 3	0.05	0.82**	1.05**	1.93	0.21	0.24
	[0.04]	[0.32]	[0.41]	[10.53]	[0.19]	[0.21]
Treated*Week 4	0.04	0.78**	0.96**	1.03	0.19	0.22
	[0.03]	[0.38]	[0.41]	[10.58]	[0.18]	[0.20]
Treated*Week 5	0.09**	0.41	1.08**	18.80*	0.43**	0.49**
	[0.04]	[0.38]	[0.45]	[11.11]	[0.21]	[0.24]
Mean of dependent variable	0.81	8.88	7.21	163.57	4.21	4.77
Observations	5,675	4,604	5,669	5,675	5,675	5,675
R-squared	0.05	0.13	0.09	0.08	0.05	0.05

Notes:
1. This table provides the results of regressions testing for changes in labor supply and earnings as a rickshaw driver among treated participants. The participant-day is the unit of observation. "Work" is a binary variable indicating whether the participant worked that day. "Hours if >0" is the total number of hours worked, conditional on working. "Total hours" is the number of hours worked per day where hours are zero if the participant did not work. "Earn" is the total daily earnings as a rickshaw driver in Indian Rupees. This variable takes the value zero if the participant did not work. The exchange rate is roughly 45 Rupees to 1 US Dollar. "ln(earn+1)" is the log of earnings + 1 where earnings is defined as previously described. "IHS(earn)" is the inverse hyperbolic sine of earnings.
2. All regressions include stand location, calendar week, and study week fixed effects as well as controls for baseline work habits which are detailed in Appendix 1B. Results are less precisely estimated, but qualitatively similar if these controls are omitted.
3. Robust standard errors clustered by individual are in brackets.
4. *** p<0.01, ** p<0.05, * p<0.1

Despite the negative point estimate in early portion of the study in the full sample, and consistent with the trend in labor supply, earnings follow an increasing and significant trend across the weeks of the study (Table 1.5, Column (4) and Table 1A.3, Column (4)). Although earnings lag labor supply slightly in the middle portion of the study, the 95 percent confidence intervals for estimated earnings and earnings predicted according to labor supply changes overlap significantly (Figure 1.5). Even with very noisy income streams, all specifications show an increase in earnings significant at the 10 percent level by the fifth week.[21] Of note, the point estimate of a daily net gain of approximately Rs. 15 by week five corresponds to a 9 percent increase in average earnings, a figure similar to the increase in labor supply.[22]

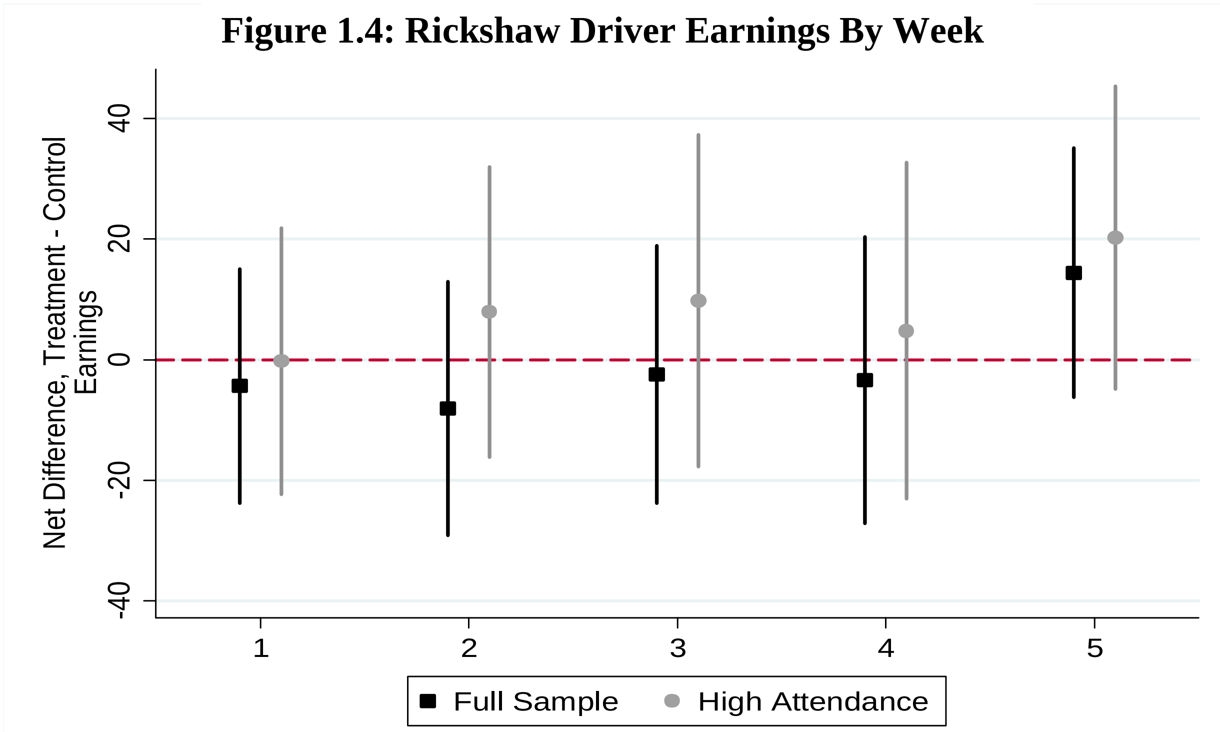
Figure 1.3: Rickshaw Driver Labor Supply By Week



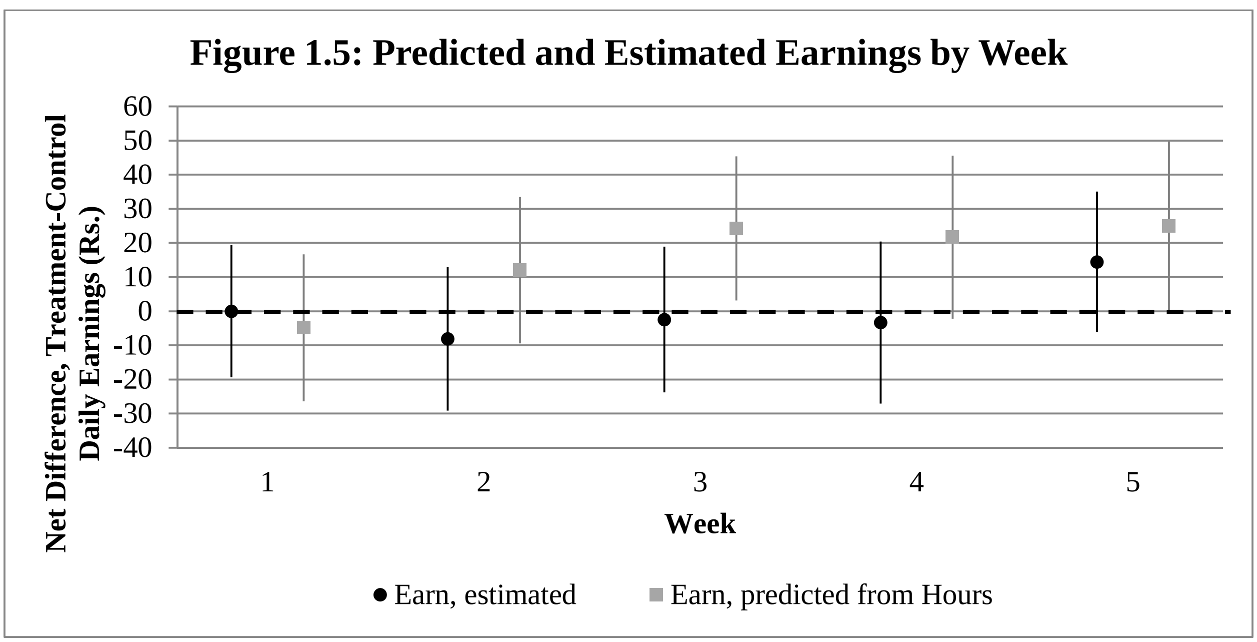
Notes: This figure is drawn from the results in Table 1.5, Columns (1) and (3). "Days Worked" measures the fraction of days worked. "Total Hours" is the number of hours worked per day. If an individual does not work, this variable takes the value zero. Error bars measure 95 percent confidence intervals.

[21] The differences in the point estimates in the regressions by week and by period (first and second half) of the study are driven by two effects: 1) the period regressions also include laboratory earnings while week regressions do not, 2) a compositional effect driven by the unit of observation (person-period averages in the period regressions which weight individuals equally, and person-day observations in the weekly regressions which implicitly weight by attendance). Given the higher attrition among the high-earning treated individuals, regressions using the person-day as the unit of observation will result in lower estimated effects if baseline controls cannot fully control for "type."

[22] This estimate is less than the estimated coefficient in the log specification because that specification takes logs of earnings + 1 to account for extensive margin changes in labor supply.



Notes: This figure is drawn from the results in Table 1.5, Column (4) and Appendix 1A, Table 1A.5, Column (2). High attendance is defined as attending at least 90 percent of the days in the study. Error bars measure 95 percent confidence intervals.



Notes: Estimated changes in earnings are in Rupees and are based on Equation (1.2). Results of this regression are displayed in Table 1.5, Column (4). Predicted earnings are generated by multiplying estimated changes in total hours (Table 1.5, Column (3)) among treated individuals by average hourly earnings in the control group. Error bars measure 95 percent confidence intervals.

1.3.6 Enrollment and Endline Tasks

As described previously, in addition to reporting daily earnings and labor supply participants were asked to spend a full day in the study offices at enrollment and during the last week of their participation. The tasks completed at enrollment and endline were designed to be as directly related to potential productivity as possible. For example, physical tasks included choosing whether or not to take actual rickshaw journeys at different distances, weights, and payoffs. This design allows assessment of the productive ability of the participants in the absence of demand fluctuations and other sources of variability (e.g. measurement error). Task payments were a direct function of performance. Additional information regarding the tasks is provided in Appendix 1A, Table 1A.1.

The primary specification to examine the impact of treatment on task earnings is shown in Equation (1.3) where e_{it} is earnings on for individual i at time t (where t = {enrollment, endline}), T_i is a binary variable indicating treatment status, and End_t indicates whether the measurement was taken at Endline.

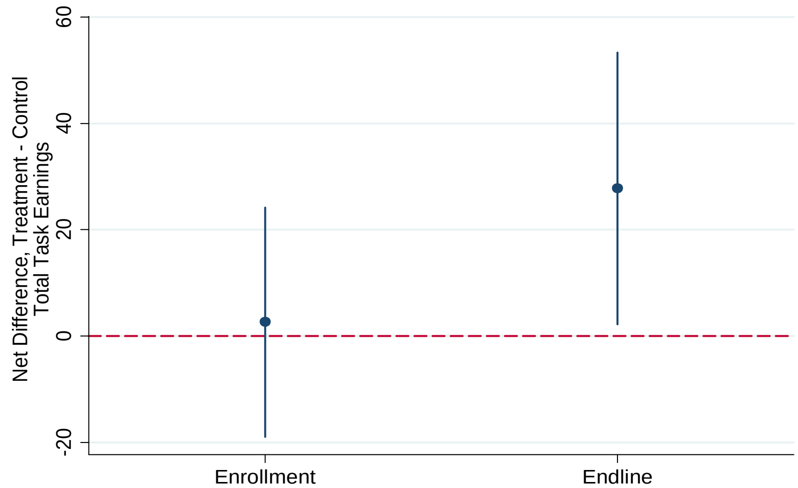
(1.3) $$e_{it} = \beta_0 + \beta_1 T_i + \beta_2 End_t + \beta_3 T_i * End_t + \varepsilon_{it}$$

As can be seen in Figure 1.6 and Table 1.6, Column (1), treated participants earn roughly 9 percent (Rs. 25) more than control participants on the tasks at endline. Overall earnings are divided into cognitive and physical tasks in Columns (2) and (3) and Figure 1.7. Cognitive tasks show a virtually immediate benefit of the increased consumption with 12 percent higher earnings at enrollment among the treated individuals (who received an additional 700 calories on the enrollment day as well as throughout the study) relative to control individuals.[23] This immediate change in cognitive function is consistent with evidence from a variety of other sources.[24] The improvement in cognitive outcomes is maintained at approximately the same level at endline. In contrast, gains in physical performance associated with increased caloric intake

[23] On the enrollment and endline days all participants were provided with a lunch. In addition, treated participants received snacks amounting to 700 calories shortly after enrolling. The snacks could all be consumed immediately or eaten incrementally throughout the day.
[24] For example, in a double blind experiment, Gailliot et al. (2007) demonstrate that higher blood glucose levels increase self-control over the course of just a few hours and Danziger et al. (2011) provide evidence that judges produce much lighter sentences when they have eaten recently.

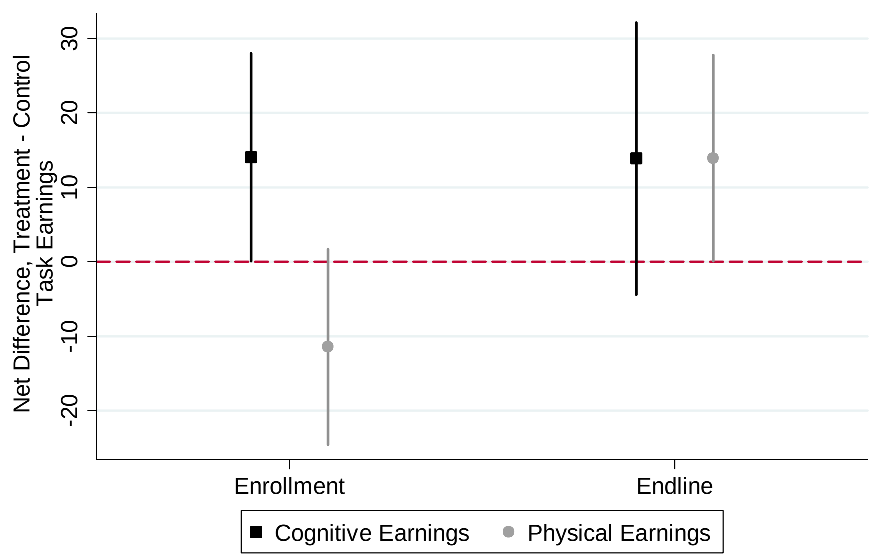
accrue over time with no significant effect at enrollment but a positive and significant gain of over 10 percent at endline relative to the control group. These results correspond well with both basic physiology and the results of experiments measuring changes in physical performance as a function of caloric intake over longer time horizons (e.g. Keys et al. 1950).

Figure 1.6: Earnings on Laboratory Tasks



Notes: Earnings (in Rs.) are aggregated across all laboratory tasks. Tasks were compensated according to performance. This figure is based on the regressions in Table 1.6, Column (1). Control participants earned an average of Rs. 279 at enrollment. Error bars measure 95 percent confidence intervals.

Figure 1.7: Earnings on Laboratory Tasks, By Type



Notes: Earnings (in Rs.) are aggregated across all laboratory tasks within a category. Tasks were compensated according to performance. This figure is based on the regressions displayed in Table 1.6, Columns (2) and (3). Control participants earned an average of Rs. 279 at enrollment. Error bars measure 95 percent confidence intervals.

Table 1.6: Earnings From Enrollment and Endline Tasks

Dependent Variable	(1) Total Pay	(2) Cognitive pay	(3) Physical pay	(4) Total Pay	(5) Cognitive pay	(6) Physical pay
Treatment (Enrollment)	0.29	13.72*	-13.43	21.04	21.20*	-0.16
	[11.62]	[7.03]	[8.19]	[17.49]	[10.97]	[11.84]
Endline	12.64	30.18***	-17.45**	33.68**	31.73***	2.15
	[9.39]	[4.45]	[8.22]	[14.60]	[7.26]	[12.50]
Low BMI				15.90	1.33	14.57
				[17.47]	[9.82]	[13.18]
Treatment*Endline	24.95**	-0.67	25.53**	-0.74	2.76	-3.70
	[11.68]	[6.49]	[9.90]	[17.95]	[9.82]	[14.81]
Low BMI*Treatment				-33.42	-11.75	-21.67
				[23.44]	[14.35]	[16.28]
Low BMI*Endline				-35.54*	-2.62	-33.12**
				[18.84]	[9.20]	[16.32]
Low BMI*Treatment*Endline				41.25*	-5.17	46.62**
				[23.46]	[12.99]	[19.57]
Mean of dependent variable	278.6	112.88	165.72	269	112.07	156.93
Observations	389	390	389	387	388	387
R-squared	0.03	0.08	0.02	0.04	0.08	0.03

Notes:
1. This table presents results of regressions of participant earnings on laboratory tasks as a function of time in the study (enrollment or endline day) and treatment status. Tasks were compensated according to performance. "Total Pay" is the total earned payment from participation in the enrollment (endline) tasks in Indian Rupees. "Cognitive pay" is the total earned payment from participation in the subset of tasks during enrollment (endline) which were cognitive. Similarly, "Physical pay" is the earned payment on physical tasks. The exchange rate is roughly 45 Rupees to 1 US Dollar. The tasks and their designation as Cognitive or Physical is included in Appendix 1A, Table 1A.1. The participant day is the unit of observation.
2. "Low BMI" is defined as BMI < 18.5, the WHO cutoff for "underweight."
3. Robust standard errors clustered by individual are in brackets.
4. *** p<0.01, ** p<0.05, * p<0.1

Although all participants have a BMI under 20, treatment may have differential effects for particularly low BMI individuals. To examine possible heterogeneity in the treatment effect, I estimate

Equation (1.4) in which an indicator for low BMI, denoted by L_i, is interacted with the previous specification.[25]

$$(1.4) \quad e_{it} = \beta_0 + \beta_1 T_i + \beta_2 End_t + \beta_3 L_i + \beta_4 L_i{*}T_i + \beta_5 L_i{*}End_t + \beta_6 T_i{*}End_t + \beta_7 L_i{*}T_i{*}End_t + \varepsilon_{it}$$

These results, presented in Table 1.6 Columns (4) thru (6) suggest that while most of the physical benefits appear to accrue to the lower-BMI individuals, the cognitive benefits apply equally to all treated participants.[26]

1.3.7 Bounded Estimates

The estimated effects are relatively large and show consistent trends across labor supply, earnings on the job, and earnings in the laboratory. However, to address concerns that the above treatment effects may be driven by the attrition patterns described previously, I also construct bounded estimates of labor supply and earnings both as a rickshaw driver and in the laboratory which account for the differential attrition. The bounding process follows the method proposed and described by Lee (2003), with minor adjustments to account for the structure of the data. This method relies on the trimming of extreme observations in the condition with lower attrition to bound the estimated treatment effect. Specifically, after calculating the relative difference in attrition to determine the extent of the trimming necessary, the individuals with the highest (lowest) average values of the outcome are removed from the dataset and the treatment effect is estimated on the trimmed dataset. Additional details about this procedure are included in Appendix 1B.

The observed labor supply increases are relatively robust (Table 1.7a, Columns (1) to (4) and Figure 1.8). Regardless of the extent or type of trimming, the fraction of days worked is higher among treated individuals at the end of the study. Although the most extreme lower bound does not result in a statistically

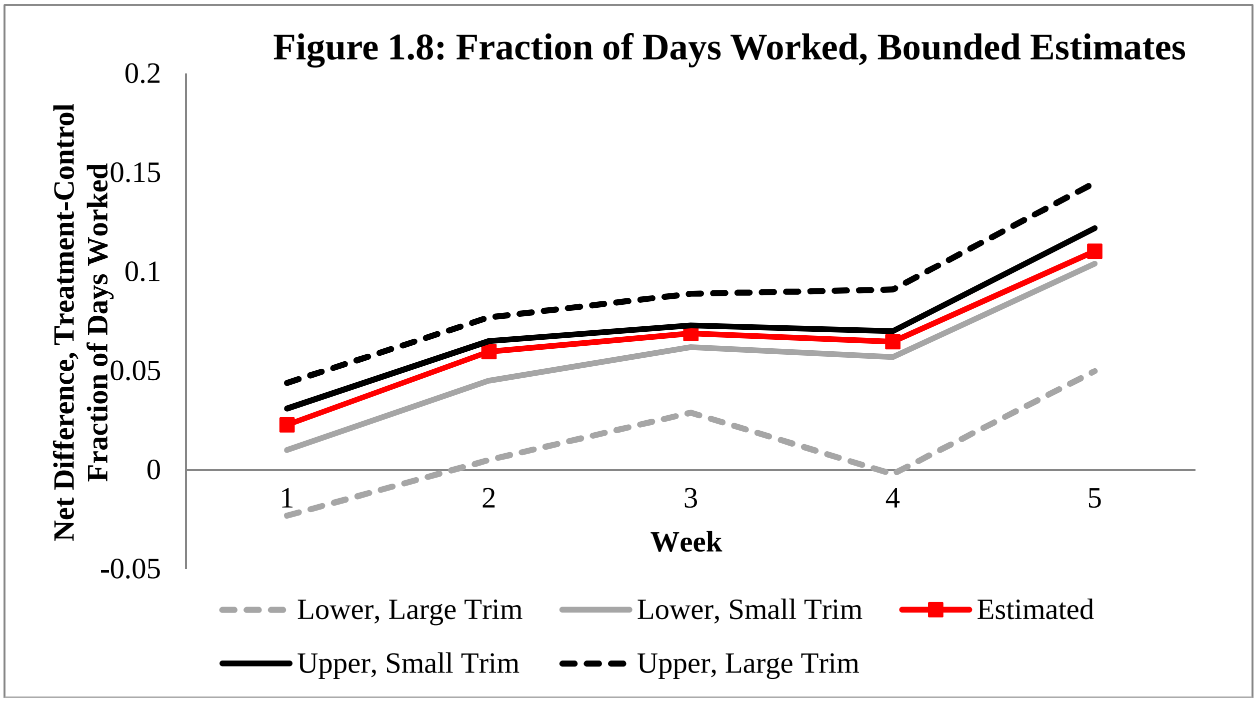
[25] "Low BMI" is defined as a BMI under 18.5, the WHO cutoff for "underweight," at enrollment. Sixty-five percent of participants are under this threshold. The distribution of participants' BMIs is provided in Figure 1A.1.
[26] Labor supply and earnings follow a pattern more closely matched to the physical tasks. Despite this similarity, the heterogeneity by BMI does not reach statistical significance for these outcomes (See Appendix 1A, Table 1A.6).

significant net effect in the fifth week, the point estimate is positive and still shows a marginally significant increase from the first week. In addition, the more moderate lower bound does estimate a significant net gain in the fifth week and at 5 to 10 percent these lower bound increases in labor supply are economically significant.

Table 1.7a: Bounds, Rickshaw Driver Labor Supply and Earnings

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	Days Worked				Earn			
Bound	Lower	Lower	Upper	Upper	Lower	Lower	Upper	Upper
Trimming percentage	15.9	5.3	5.3	15.9	15.9	5.3	5.3	15.9
Treated (Week 1)	-0.02	0.01	0.03	0.04	-19.03**	-10.25	-0.98	8.14
	[0.03]	[0.03]	[0.03]	[0.03]	[9.52]	[9.65]	[9.19]	[9.02]
Treated*Week 2	0.03	0.04	0.03	0.03	-2.90	-3.72	2.48	4.65
	[0.03]	[0.03]	[0.03]	[0.04]	[9.87]	[9.67]	[9.49]	[9.29]
Treated*Week 3	0.05	0.05	0.04	0.05	4.81	3.90	6.22	7.33
	[0.03]	[0.04]	[0.04]	[0.04]	[10.64]	[10.41]	[9.99]	[10.32]
Treated*Week 4	0.02	0.05	0.04	0.05	-0.90	2.20	9.29	10.94
	[0.03]	[0.03]	[0.04]	[0.04]	[10.75]	[10.53]	[9.94]	[10.22]
Treated*Week 5	0.07*	0.09**	0.09**	0.10**	20.08*	20.61*	20.02*	22.41**
	[0.04]	[0.04]	[0.04]	[0.04]	[11.35]	[10.94]	[10.88]	[11.17]
Mean of dependent variable	0.80	0.81	0.82	0.84	151.01	157.32	165.36	170.96
Observations	5,291	5,602	5,501	5,199	5,294	5,602	5,523	5,196
R-squared	0.05	0.05	0.05	0.05	0.10	0.09	0.07	0.07

Notes:
1. This table provides the bounded estimates of changes in labor supply and earnings for treated participants over time. The participant-day is the unit of observation. "Work" is a binary variable indicating whether the participant worked that day. "Earn" is the total daily earnings as a rickshaw driver in Indian Rupees.
2. Attrition was greater in the treated condition, hence observations were trimmed from the Control condition. Trimming percentages were calculated in according with Lee (2003). Because attrition can be measured in multiple ways given the panel nature of the data, five measures of attrition (never return after the first day, do not attend endline, and attend less than X% of the days where X is 25, 50, or 75) were created and the smallest and largest trimming percentages (5.3% and 15.9%) were used as outer bounds.
3. "Bound" indicates whether observations were removed from the top of the control distribution, resulting in an upper bound on the treatment effect ("Upper"), or bottom of the control distribution, leading to a lower bound on the treatment effect ("Lower").
4. The metric used for the distribution of outcomes was the average days worked (earnings) over all observations for an individual. All individuals have at least one observation because all participants provided labor supply information at enrollment.
5. All regressions include stand location, calendar week, and study week fixed effects as well as controls for baseline work habits. Results are less precisely estimated, but qualitatively similar if controls are omitted.
6. Robust standard errors clustered by individual are in brackets.
7. *** p<0.01, ** p<0.05, * p<0.1



Notes: This figure is drawn from the results in Table 1.5, Column (1), which provides the point estimates of the changes in days worked without any corrections for attrition ("Estimated") and Table 1.7a, Columns (1) thru (4) which provides the bounded estimates of the changes in days worked (the "Trimmed" estimates).

Earnings estimates have much wider bounds. These regressions also show significant gains in the fifth week relative to the first among treated individuals. However, although there are gains among treated individuals over time, as discussed previously, the initial point estimates in week one are negative in some specifications. Given these negative initial point estimates, I cannot not rule out the possibility that there are no net gains in cycle-rickshaw driver earnings in week five in the bounded specifications.

Finally, Table 1.7b presents bounded estimates for laboratory task payments. Bounds on these estimates are relatively wide and cannot rule out no net effect of treatment at endline, but are still generally supportive of a positive impact of treatment on laboratory earnings. In Column (1), the lower bound has a negative point estimate for treated individuals at enrollment given the censoring of high earning control individuals. However, the interaction between treatment and endline remains positive and significant. The upper bound estimates a significant treatment effect of Rs. 24 at enrollment and an additional, although non-significant, gain of Rs. 9 at endline for treated individuals (Column (2)).

Table 1.7b: Bounds, Earnings From Enrollment and Endline Tasks

	(1)	(2)
Dependent Variable	Total Pay	
Bound	Lower	Upper
Treatment	-19.30*	23.79**
	[11.22]	[10.34]
Endline	6.51	28.22***
	[10.16]	[8.64]
Treatment*Endline	31.09**	9.38
	[12.30]	[11.08]
Mean of dependent variable	298.19	255.1
Observations	363	365
R-squared	0.03	0.07

Notes:
1. This table provides the bounded estimates of changes in earnings in laboratory tasks by treatment status and day in the study (enrollment or endline). Tasks were compensated according to performance. "Total Pay" is the total earned payment from participation in the enrollment (endline) tasks in Indian Rupees.
2. Attrition was greater in the treated condition, hence observations were trimmed from the Control condition. Trimming percentages were calculated in according with Lee (2003) with attrition in each experimental condition calculated as the percentage of individuals not attending endline.
3."Bound" indicates whether observations were removed from the top of the control distribution, resulting in an upper bound on the treatment effect ("Upper"), or bottom of the control distribution, leading to a lower bound on the treatment effect ("Lower").
4. The metric used for the distribution of outcomes was the total earnings on all enrollment tasks for an individual.
5. Robust standard errors clustered by individual are in brackets.
6. *** p<0.01, ** p<0.05, * p<0.1

1.3.8 Summary of RCT Results

Increased caloric intake among treated individuals resulted in gains in both laboratory tasks designed to measure ability in domains relevant to their work as well as direct measures of labor supply. Earnings present a somewhat more mixed picture due to high variance in daily earnings and the initial negative point estimates early in study which may be driven by selected attrition, with high earning treated individuals being more likely to leave the study in the first few days. However, consistent with evidence from the physiology literature, effects of increased intake do appear to be cumulative with clear trends in both labor supply and earnings over time that do not appear to plateau within the five week period studied. These trends result in significant and economically meaningful net increases in labor supply and earnings (in both laboratory tasks and as a rickshaw driver) among treated individuals of approximately 10 percent by the fifth and final week.

1.4 Randomized Controlled Trial Return on Investment

The estimated changes in productivity in the RCT suggest that changes in caloric intake cause economically meaningful changes in productivity. Although these changes are of interest in and of themselves given the importance of labor productivity in development, they do raise the additional question of whether these changes in productivity simply "subsidize" food consumption, or whether they not only pay for themselves but also generate positive returns. Hence, I next provide calculations of the return on investment.

Seven hundred calories of a staple grain such as rice or wheat can be purchased for approximately 6.5 Rupees.[27] Accounting for the initial negative point estimates in the early weeks and capping the gains at Rs. 14.4 per day (the estimated net treatment effect on rickshaw driver earnings in the fifth week), a rickshaw driver would expect to break even after approximately two months of increased consumption. Over a six month horizon, the estimated return on investment is approximately 75 percent.[28] Because there is greater inherent uncertainty in these estimates than in the estimates of changes in productivity, additional estimates of the ROI are included in Table 1.8 varying assumptions about both the cost of calories and earnings over time. While the bounds are large, the average estimated 6-month ROI is 134 percent, excluding any gains from improved cognitive function that are not captured in earnings or labor supply as a rickshaw driver.

[27] The food provided as a part of the study cost roughly Rs. 20 per person per day. This figure is substantially higher than the Rs. 6.5 for staple grains because the study provided snack foods to minimize the likelihood of crowding out other food consumption. Although it is not possible to exclude the possibility that the consumption of grains could produce different results, neither grains nor snack foods contain substantial nutritive value beyond calories and both consist primarily of basic starches, minimizing the chance of such a difference.

[28] As described in Chapter 2, the estimated returns to increased caloric intake in agricultural production are even higher at roughly 200% over one month. One potential concern about these calculations is that basal metabolic rates could adjust, requiring greater caloric inputs to maintain the same benefits. However, calories are sufficiently inexpensive that even with changes larger than would be expected given the results of the Minnesota semi-starvation experiment, the return on investment would still be positive and relatively large (Keys et al. 1950).

Table 1.8: Return on Investment for Increased Caloric Consumption Among Rickshaw Drivers

	Sample	Earnings during study	Earnings after study	6-month ROI		
	Cost of 700 calories			3.75	6.5	8.82
(1)	full	average	average	-122	-112	-109
(2)	full	by period (with lab tasks)	2nd period	321	143	79
(3)	full	weekly	5th week	207	77	30
(4)	high attendance	average	average	125	30	-4
(5)	high attendance	weekly	5th week	380	177	104
(6)	full	average labor supply	average labor supply	320	142	78
(7)	full	weekly labor supply	5th week labor supply	515	255	161
		Average		**250**	**102**	**49**

Notes:
1) Earnings values are based on results in Tables 1.4 (by period estimates), 1.5 (weekly estimates) and 1A.4 (average estimates). Values generated from changes in labor supply multiply changes in total hours found in the same tables by the average hourly wage in the control group.
2) Three cost estimates are used. The least expensive, 3.75 Rs. per calorie, assumes consumption of subsidized grains through the Public Distribution System (PDS). The medium cost, 6.5 Rs. per calorie, was assessed by visiting local shops and bargaining to purchase rice, a local staple grain, at prevailing market prices. There are less expensive staple grains. However, rice was used as a baseline because the other grains are less commonly consumed. The third estimate, 8.82 Rs. per calorie, was obtained by calculating the average price of a calorie in the National Sample Survey (NSS) Round 64 and adjusting the price for inflation using the consumer price index.
3) Returns are calculated as the sum of the earnings during the study period and from the end of the study period to 6 months minus the cost of 700 calories per day throughout the six month period, all divided by the same cost. No discounting is used because of the relatively short time horizons.
4) Given the cumulative nature of the response to increased intake, the preferred specification uses the direct earnings results as estimated each week and the intermediate cost of calories (6.5 Rs). This specification is located in Row (3).

1.5 Why Does Caloric Intake Remain Low Despite High Returns?

Despite the greater uncertainty relative to the estimates for productivity changes, the estimated return on investment, 75 percent over a horizon of only six months, is relatively high. In addition, at roughly four percent of income, the cost of investment in an additional 700 calories is low. Given this low cost and high return, it is puzzling that investment in sufficient caloric intake remains low, with over half of the population of India consuming less than 1900 calories per day and one quarter consuming less than 1625 calories per day (Deaton and Dreze 2009). I discuss and briefly evaluate five possible reasons for low investment below. Evidence regarding these potential factors influencing food consumption is drawn from a variety of sources including relevant literature, experimental tasks, and surveys of hundreds of individuals from diverse geographic and demographic backgrounds designed specifically for this purpose. The general design of the surveys is described briefly below and additional details are provided in each of the relevant sections.

Reasons for Low Caloric Intake - Surveys

This survey was conducted to investigate food habits and potential reasons for low caloric intake among individuals similar to the populations in the analyses examining the returns to increased caloric intake. To accomplish this, participants were drawn from both urban and rural areas in India, and were screened to be at least 18 years old, have a BMI under 20, and be an active participant in the labor market. Demographic characteristics of the 222 survey respondents are presented in Appendix 1A, Table 1A.7. Participants were paid a flat fee for completing the survey. However, to ensure that participants considered their responses carefully and that the elicitation was incentive compatible, respondents were also eligible to earn additional compensation for correct responses to a subset of factual questions regarding the caloric density of foods and the returns to increased caloric intake.

1.5.1 Credit

In canonical examples of nutrition-based poverty traps income (or credit) is the constraint limiting investment (e.g., Leibenstein 1957; Bliss and Stern 1978; Dasgupta and Ray 1986). However, there is strong evidence that neither income nor credit is likely to be a binding constraint faced by individuals in India today when making investment decisions about caloric consumption. Even the poorest typically do not spend more than 40 to 70 percent of their income of food, could consume approximately 20 percent more calories at the same cost by purchasing less expensive staple grains, and have a relatively low elasticity of calories with respect to income (Subramanian and Deaton 1996; Banerjee and Duflo 2011; Behrman and Deolalikar 1987). Further, specific to the context of the studies presented in this paper, an increase of roughly 700 calories per day would come at a cost of less than 5 percent of income and could be repaid over a short horizon. Finally, when survey respondents were asked to list all factors influencing their level of food consumption, fewer than 10 percent included lack of money as a limiting factor. In short, liquidity does not appear to be a binding constraint restricting investment in higher caloric intake for many malnourished individuals in India.

1.5.2 Structural Features of the Economy

There are a number of structural features common to developing economies which could potentially limit investment in calories. For example, principal-agent contracting with unobserved caloric intake or strong complementarities in production could reduce intake below optimal levels. However, there are a number of reasons to believe that these features are not likely to be driving the low caloric consumption. First and foremost, while caloric intake is not directly observable, observable characteristics such as weight (BMI) would allow individuals to clearly signal potential productivity. In addition, low caloric intake is quite prevalent among individuals who are the residual claimant on their labor and for whom no complementarities in production with others exist. For example, 65 percent of the cycle-rickshaw drivers in Chennai have a BMI under 20.[29] These facts suggest that the reasons for low investment are likely to lie elsewhere.

1.5.3 Utility Returns

Although the monetary returns to additional caloric intake are high in the populations studied in this paper, it is possible that the utility returns are not. The decision faced for each individual considering an investment in food is between the status quo and a bundle including additional food, (potentially) additional work, and additional income. Hence, there are two possible ways in which utility returns could be lower than monetary returns. First, people could potentially experience disutility from consuming additional food. Although this

[29] This figure is the result of randomly sampling 100 cycle-rickshaw drivers in a 2.5 square kilometer area where the vast majority of rickshaws operate in Chennai.

is possible, it seems quite unlikely given the low level of consumption, the positive income elasticities found in other research, and the ready acceptance of additional food by participants in the RCT (Subramanian and Deaton 1996; Behrman and Deolalikar 1987; Bouis and Haddad 1992).

The second and more plausible concern is that higher caloric intake could increase labor supply, which comes with a utility cost. However, there are three reasons to believe labor disutility is not a factor limiting investment in calories. First, it is not clear that net utility costs of labor will increase given the lack of change in labor supply in the agricultural analysis and the fact that increased consumption may improve energy and reduce the inframarginal disutility of labor, a topic explored further below. In addition, if this disutility were substantial, then one would expect the participants of the RCT to have maintained both their original food consumption (by crowding out consumption outside of the lab) and their original labor supply. Yet, this was not the observed result. Finally, in addition to the direct revealed preference argument, it is possible to provide a rough calibration of the tradeoff observed in the RCT to assess the likelihood that labor disutility is driving low caloric consumption among the rickshaw drivers. This calibration, detailed in Appendix 1B, relies on four sources of data: 1) changes in earnings and labor supply in the RCT, 2) an experimental task in the RCT which elicited willingness to accept values for marginal trips in an incentive compatible manner, 3) additional measures of willingness to accept trips generated from 150 actual offered rickshaw trips with randomly assigned offered fares, 4) purchases of staple grains to determine their cost on the open market.

This calibration suggests that even ignoring any direct utility from food consumption there would be positive utility returns to higher caloric intake for roughly 85 percent of drivers. Further, the calibration is conservative in that it ignores the fact that increased caloric intake may decrease the disutility of inframarginal labor by improving energy levels and making physical exertion less aversive. Although far from conclusive, there are two pieces of suggestive evidence from the RCT that support the hypothesis that the utility cost of inframarginal labor may decline with higher caloric intake. First, treated individuals reported significantly higher energy levels than control participants on the daily surveys (Table 1.9, Column (1)). Second, one of the experimental tasks during enrollment and endline days was comprised of decisions about whether or not to take actual rickshaw journeys. Participants were offered six different tradeoffs of this sort and one of the choices was randomly selected and the participant's choice carried out. In this task, treated participants were significantly more likely to accept the offered trips (Column (2)). In addition, treated participants increased their willingness to take journeys from 86 percent at enrollment to 93 percent at endline. Both the higher level and the trend suggest the cost of labor provision may be decreasing with higher caloric intake. Hence, although these facts do not provide sufficient information about the specific shape of the cost curve for labor supply to provide conclusive evidence, they do suggest that the overall disutility of labor could remain constant or decline rather than increase with additional caloric intake.

Table 1.9: Rickshaw Drivers' Disutility of Labor

Dependent Variable	(1) Energy	(2) Number of Trips Accepted
Treated	0.07*	0.39*
	[0.04]	[0.20]
Mean of dependent variable	3.24	5.18
Observations	5,945	391
R-squared	0.05	0.36

Notes:
1. This table reports the effect of treatment on participants' self-reported energy levels and willingness to accept offered rickshaw journeys to provide suggestive evidence for the impact of caloric consumption on the utility cost of labor supply.
2. "Energy" is the participants' self-reported daily energy levels on a scale of 1 (very low) to 5 (very high). However, there was very low variability in reported energy levels with 3 and 4 accounting for 89 percent of responses. "Number of Trips Accepted" sums the number of cases in which a participant opted to take a trip of the six different types of trips offered during as an experimental task at enrollment and endline. One of the offers was randomly selected and the participants' choices were carried out.
3. Both regressions include controls for the week in the study. In addition, Column (1) controls for a baseline measure of energy, the participants self-assessment of whether low energy had been a problem for them in the week preceding enrollment.
4. Robust standard errors clustered by individual are in brackets.
5. *** p<0.01, ** p<0.05, * p<0.1

1.5.4 High Discount Rates

Even if an individual is does not face liquidity constraints in obtaining sufficient caloric intake, he may choose to purchase better tasting foods that provide fewer calories per rupee or other non-food items if he has a very high discount rate or time inconsistent preferences. Although I cannot rule out these possibilities with the evidence currently available, survey evidence suggests that time inconsistent preferences are unlikely to be a primary reason for the low investment in calories. For example, over 70 percent of survey respondents reported that they did not wish to make any changes, including amount, timing, or type, to their current food consumption. In addition, the discount rates implied by the rate of return to the investment would be extremely high. Finally, it is possible that high discount rates would increase rather than decrease caloric intake given that many foods which are typically considered “tempting” are also high calorie per rupee foods (e.g. oils, sugar) (Banerjee and Mullainthan 2010).

1.5.5 Lack of Knowledge

Knowledge has considerable scope to influence caloric intake. In particular, there are two critical pieces of knowledge that individuals need to make optimal caloric intake decisions: 1) knowledge of the returns to caloric intake, and 2) knowledge of which foods are calorically dense (i.e. provide a large number of calories per Rupee). Despite being important, knowledge in these domains is likely to be difficult to acquire for a number of reasons. First, income streams among the poor are typically very volatile, making it difficult to learn about the returns to an investment. This is particularly true for investments like caloric intake in which, despite high returns, the absolute changes are small relative to the variance in income. In addition, the returns to increased caloric intake grow over time, so relatively long periods of experimentation may be required to learn about the returns available. Finally, as will be discussed in greater detail below, the short run and long-run effects of increased consumption may work in opposite directions; it is common to feel sleepy or tired right after a large meal. But, over a longer time horizon of days or weeks, baseline energy levels increase.

If learning about the returns to higher caloric intake were not challenging enough by itself, learning about the caloric density of food is also quite difficult. While the human body has developed mechanisms to regulate intake, these mechanisms evolved during a time period when food was relatively more expensive such that appetite regulation may not be well-calibrated to the current environment. In addition, the mechanisms to regulate intake are imperfect because feelings of satiety are often not representative of the caloric content of the food. For example, foods which contain large amounts of fiber or that have high water content are quite filling, yet provide less usable energy because fiber cannot be digested and water does not contain calories (Gerstein et al. 2004; Samra and Anderson 2007).

In short, there are many difficulties inherent in this learning process which make knowledge a potentially important factor limiting caloric intake. I explore this possibility via the survey described previously because, to the best of my knowledge, no pre-existing data are available on beliefs about returns to caloric intake or the nutritive properties of foods in a relevant population.

Survey Evidence – Knowledge of Returns to Increased Caloric Intake

To gather data about respondents' beliefs regarding the productivity consequences of increased caloric intake, respondents were asked whether consuming additional food would increase, not change, or decrease their productivity. Participants responded to a multiple variants of this question considering different types of foods (an additional meal or the equivalent energy/calories in snacks) and different outcomes (labor supply or earnings) to explore the consistency of the beliefs. To ensure that questions were considered carefully, participants were paid Rs. 2 for each correct response to these questions.

There were two particularly striking features of these results (Table 1.10). First, nearly one-third of respondents report being uncertain about the returns to increased caloric intake. Second, only a relatively small fraction of individuals, typically about 20 percent, believe that increased caloric consumption leads to increased labor supply or earnings.[30] In addition, among those individuals who responded that increasing food consumption would increase labor supply and earnings, beliefs about the labor supply margins that would change were highly variable and not well aligned with the results of the RCT.[31] For example, despite being asked to indicate all changes they thought would occur, only 7 percent of respondents listed an increase in the number of days worked – the margin with the largest change in the RCT.

Table 1.10: Knowledge of Returns to Higher Caloric Intake

Question type 1: Do you think that if a person similar to you (i.e those in a similar profession, and with the same height and weight) ate [*Food type*] every day for the next [*Timeframe*] they would [*Labor outcome*] less, the same, or more [*Timeframe*] from now?

	Timeframe	Food type	Labor Outcome	Less	Same	More	Unsure
(1)	"consistently"	n/a	work	25.5	35.6	13.4	25.5
(2)	1 month	meal	work	12.8	26.8	27.5	32.9
(3)	1 month	snacks	work	10.7	31.5	23.5	34.2
(4)	1 month	meal	earn	3.4	40.3	20.8	35.6
(5)	1 month	snacks	earn	3.4	43.0	18.1	35.6
		Average		11.1	35.4	20.7	32.8

Question 2: In which, if any, of the following ways would eating more help you or others like you to work more?

		Take fewer or shorter breaks	Work more days	Work longer hours	Work faster/ harder	Look more for better jobs	Unsure or other
(6)	Percent of participants	16.8	7.4	46.3	22.2	0.7	3.3

Notes:
1. Data are from a survey of 149 respondents with a BMI less than 20 who are active in the labor market. Demographic information for respondents is presented in Appendix 1A, Table 1A.7, Column (2).
2. Participants were asked multiple variants of Question type 1 (provided at the top of the table) where items in brackets were completed with options from the lists below:
Food type: "an extra meal," "the equivalent of an extra meal in another form such as snacks like biscuits or nuts"
Time frame: "in one month," "consistently"
Labor outcome: "work," "earn"
3. Participants were paid Rs. 5 for each correct response ("more") according to the RCT results.
4. Only participants who had responded "more" to any of the variants of Question type 1 were asked Question 2. Respondents were paid Rs. 2 each for "work more days" and "work long hours."

[30] Note that although the sample for the survey was demographically diverse, respondents were selected to be similar to the rickshaw drivers in that they were active in the labor market in a job with a substantial physical component, some flexibility in labor supply, and had a BMI under 20.
[31] The list of options for ways in which productivity might change included: fewer/shorter breaks, working more days, working longer hours, working faster/harder, looking for better jobs, and "other."

Although the existence of incorrect beliefs with substantial economic consequences in equilibrium is unusual, it is not without precedent.[32] In addition, the surveys combined with data from the RCT provide preliminary but suggestive evidence for how incorrect beliefs are developed and maintained. First, with 57 percent of survey respondents reporting a decline or no change in energy and/or desire to work immediately after eating, survey evidence suggests that the immediate impact of increased food consumption is neutral or even in direct opposition to the long-run impact on energy levels found among treated participants in the RCT (Table 1.9, Column (1)).[33] These higher energy levels in the long-run are, however, difficult for individuals to link to caloric consumption given the time lag and large number of other intervening factors.

Second, changes in labor supply are difficult to track over long horizons in populations with irregular and flexible work schedules. As evidence of this difficulty even over short time horizons, I compare data on the number of days worked per week collected in two different ways; day by day collection during the RCT and a single response to the question "how many days did you work last week" in the endline survey. Limiting the sample to participants with full attendance in the week leading up to their endline day to ensure complete data, only 40 percent of respondents provide a weekly total that matches the daily responses and the mean difference is 1.4 days. Finally, low caloric intake itself may hamper learning given the observed decrements in cognitive function associated with low caloric intake.

Survey Evidence - Caloric Density of Commonly Consumed Foods

Although the number of calories per unit cost provided by different foods (i.e. "caloric density") is important knowledge to make effective investments in caloric intake, it is a difficult concept. Hence, survey questions were designed specifically to minimize the difficulty of the assessment and avoid confounding the results due to low numeracy in this population. Specifically, participants' knowledge of the caloric density of foods was tested in two different ways. First, participants were provided with a set of ten

[32] For example, Jensen (2010) provides evidence of incorrect beliefs regarding the returns to education. In addition, widespread misperceptions about health ranging from the appropriate care of children with diarrheal disease to the safety of vaccinations have also been documented (Banerjee and Duflo 2011).
[33] Participants were asked to rate their energy on a scale of 1 to 5 during their daily visit to the study office.

photographs of different food items of the same cost (Rs. 10, or roughly $0.25).[34] Food items were specifically chosen to be types of foods which are commonly consumed and adjustments to the tradeoffs were made across survey locations to ensure participants would be familiar with the food items and that costs were consistent given local prices. In addition, food items were chosen to have a wide range in caloric density. For example, in Orissa, foods ranged from 5 calories per Rupee to 134 calories per Rupee.

Each photograph was printed to scale and labeled with a description including the name, any relevant descriptors (e.g. "boiled rice"), and quantity of the food. The written label was also communicated verbally. Participants were then asked to indicate which three photographed foods they believed had the "most energy" and which three had the "least energy," accounting for both the type and amount of food in the photographs.[35] This arrangement allowed participants to assess a relatively simple metric (e.g. "most energy") while still providing information about their knowledge of the more complicated concept of calories per unit cost. Participants were paid a fee for each correct answer. If all responses to this question were correct, the earnings would amount to over 15 percent of the daily income for the average respondent, providing a strong incentive to consider responses carefully.

The ten food items are listed in Table 1.11 with the corresponding fractions of participants indicating the item was among the three highest or lowest caloric density items. This data is also displayed graphically in Figures 1.9a and 1.9b. Participants' beliefs about the caloric density of foods are not well correlated with the true values. In fact, on average, respondents were roughly equally likely to indicate that any given incorrect option was correct as they were to indicate that a correct option was indeed correct. For example, in Tamil Nadu, both the three actual highest energy foods and the remaining seven lower energy foods were identified, on average, by roughly 30 percent of the participants as being "high energy."

[34] To determine the cost of the food, field staff visited local shops and bargained to actually purchase the food items. The costs were averaged across a minimum of five purchases per food item. Caloric information was determined from a publication listing the caloric content of various commonly consumed Indian foods produced by the Indian Government (Gopalan et al. 1989). For the few food items which were not included in this publication, nutrition information was determined from large online calorie databases such as the USDA National Nutrient Database for Standard Reference.

[35] The term "energy" was used rather than calories because respondents are unfamiliar with the concept of calories.

Table 1.11: Knowledge of Caloric Density of Common Food Items, Selection From a List

Food	Calories per Rupee	"Most energy"	"Least energy"
Panel A: Tamil Nadu (N = 147)			
Beef (56g cooked, no sauce)	13	25.5	28.2
Buttermilk (400ml)	14	38.3	12.8
Chicken (76g cooked, no sauce)	17	27.5	34.2
Chips (38g)	21	12.8	47.7
Bananas (2 regular, 250 g)	26	68.5	8.1
Samosa (1 large, 83g)	27	14.1	51.0
Butter biscuits (10, 84g)	37	24.2	26.8
Sunflower oil (100ml)	84	0.0	16.1
Sugar (234g)	93	4.0	17.4
Rice (908g cooked)	121	77.2	3.4
Panel B: Orissa (N = 75)			
Mutton (19g)	5	44.0	4.0
Chicken (32 g)	7	18.7	16.0
Badam biscuits (73 g)	15	14.7	54.7
Chips (43 g)	24	5.3	64.0
Bananas (274 g)	28	53.3	14.7
Milk (400 ml)	34	88.0	1.3
Samosa (134 g)	43	1.3	84.0
Sunflower oil (110ml)	92	2.7	16.0
Sugar (265 g)	105	1.3	40.0
Rice (1005 g cooked)	134	70.7	4.0

Notes:
1. Data are from a survey of 222 respondents with a BMI less than 20 who are active in the labor market. Demographic information for respondents is presented in Appendix 1A, Table 1A.7, Column (1).
2. Participants were presented with a set of 10 photographs of the food items listed above. Each photo was printed to scale and labeled with a description including the name and quantity of the food. The written descriptions were also read to each participant. Each of the photographed items had the same cost (Rs. 10). Participants were asked to indicate which three items they believed had the “most energy” and which three items had the “least energy.” Column (2) reports the percentage of respondents including the food item among the set of three with the "most energy" and Column (3) the percentage of respondents including the food item among the set of three with the "least energy."
3. Participants were paid Rs. 3 for each correct response to these questions.

Figure 1.9a: Participants' Knowledge of Caloric Density (Tamil Nadu)

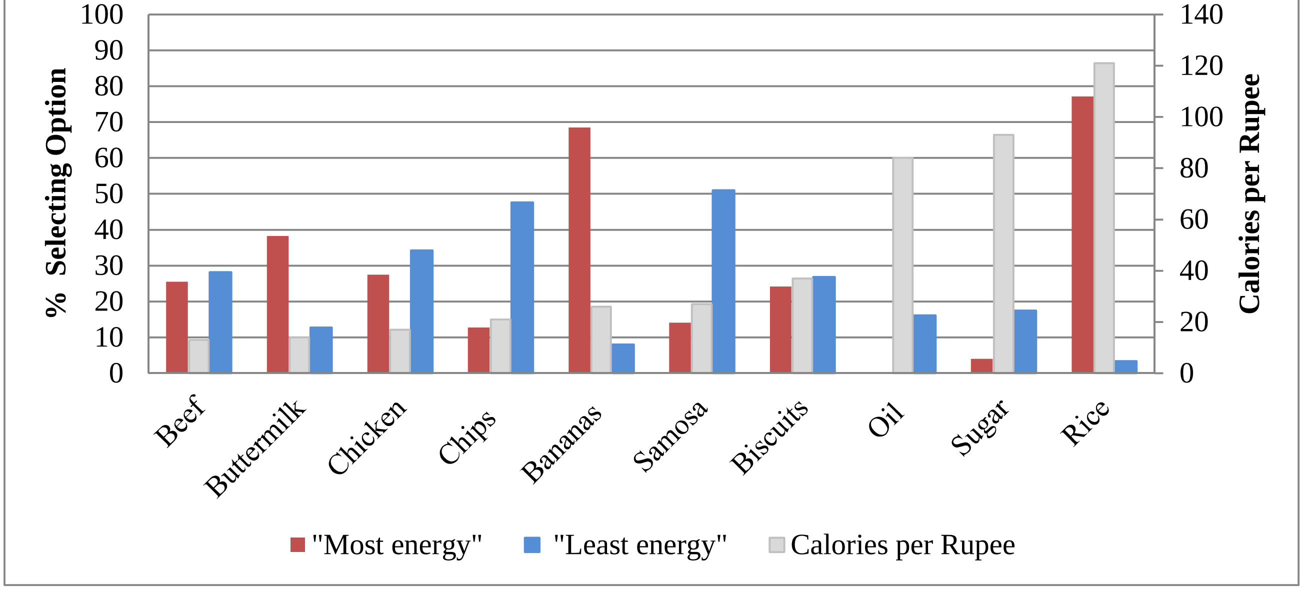
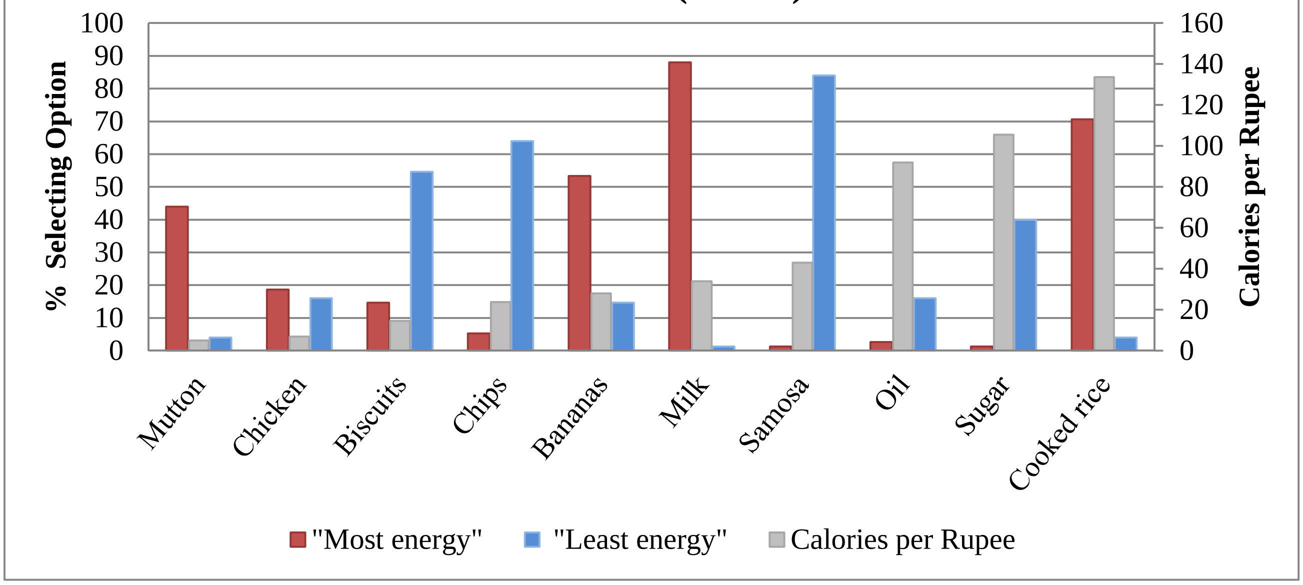


Figure 1.9b: Participants' Knowledge of Caloric Density (Orissa)



Notes: All food photographs used were printed to scale for participants. In addition, the descriptive labels, including amounts which are not displayed here, were read aloud to each participant. Participants were paid a small fee for each correct response. Demographics of survey respondents are included in Table 1A.7, Column (1).

Not only are the responses frequently incorrect, the magnitude of the errors is large. As a metric of the magnitude of the error I calculate the calories in the most (least) calorically dense foods and compare them to the total calories selected by participants in each of these categories. There are actually 3088 calories in the three most calorically dense foods. Yet, participants, on average choose three foods totaling only 1472 calories as those with the most energy; or, less than half of the possible calories. Similarly, the three foods with the fewest calories contain 378 calories, but, on average, participants choose three foods with 1052 calories as those with the least energy.[36]

As a second and even simpler test of knowledge, participants were also presented with ten pairwise tradeoffs, each between two food items with the same cost, and were asked to indicate which of the two foods provided more energy. For example, participants were asked to compare one coconut to six bananas. These tradeoffs were again presented via photographs as well as written and verbal descriptions and participants were compensated according to the accuracy of their responses. Two example tradeoffs are presented in Figure 1.10. As in the first test of knowledge, foods were selected to be commonly consumed and to have large differences in the number (and ratio) of calories in each comparison.

Table 1.12 lists these tradeoffs, the correct responses, and percentage of participants providing the correct response to each tradeoff. On average, participants chose the correct option in 41 percent of the comparisons, significantly fewer than the number expected based on purely random guessing ($p < 0.01$). And, given that the average ratio of calories (difference in calories) in each comparison was 3.7 (283 calories), the magnitude of the errors was, once again, substantial.

[36] These figures were calculated using averages weighted across locations by the number of participants in each location.

Figure 1.10: Examples of Pairwise Comparison to Identify High Caloric Density Foods

A. Raw tomato (200g)

B. Raw onion (100g)

A. 18 Butter biscuits (Medium round, 148g)

B. 1 Plate pongal (256g)

Notes: All food photographs were printed to scale for participants. In addition, the descriptive labels were read aloud to each participant. Participants were paid a small fee for each correct response. Demographics of survey respondents are included in Table 1A.7, Column (1).

Table 1.12: Knowledge of Caloric Density of Common Food Items, Pairwise Comparisons

Food A	Food B	Correct response	Percent of Participants Correct	Difference in Calories	Ratio of Calories (larger/ smaller)
Panel A: Tamil Nadu (N = 149)					
1 banana (large, 168 g)	one bun (medium, 36g)	A	78.5	71	1.7
1 tender coconut	6 bananas (1034g)	B	23.5	923	8.0
butter biscuits (18, 150g)	1 plate pongal (256g)	A	48.3	308	1.9
1 cup of tea (83ml)	1 vada (medium, 42g)	B	36.9	95	4.0
one bun (medium, 36g)	buttermilk (200ml)	A	28.9	28	1.4
tomato (100g)	onion (100g)	B	42.3	35	2.8
gram flour (100g)	2 eggs (boiled, 96g)	A	8.1	224	2.9
chicken (100g, no sauce)	fried peanuts (100g)	B	68.5	368	2.7
curd (100ml)	ghee (30ml)	B	36.9	207	4.3
sambar (230ml)	kurma (165ml)	B	20.8	150	2.2
Average, Tamil Nadu			39.3	241	3.2
Panel B: Orissa (N = 75)					
1 banana (59g)	1 bun (59g)	B	45.3	104	2.7
6 bananas (385g)	1 tender coconut	A	53.3	261	3.0
curd (90g)	2 buns (103g)	B	34.7	233	5.3
1 cup of tea (61ml)	2 vada (medium, 88.4g)	B	66.7	242	11.6
2 buns (103g)	1 glass milk (145 mL)	A	16.0	163	2.3
tomato (200g)	onion (100g)	B	36.0	15	1.4
gram flour (165 g)	2 eggs (boiled, 86g)	A	20.0	457	5.4
chicken (158g cooked)	fried peanuts (340g)	B	48.0	1648	5.8
curd (90g)	ghee (30g)	B	86.7	209	4.9
dali (260ml)	mixed vegetables (303g)	B	42.7	339	5.0
Average, Orissa			44.9	367	4.7
Average, Overall			41.2	283.1	3.7

Notes:
1. Data is from a survey of 222 respondents with a BMI less than 20 who are active in the labor market. Demographic information for respondents is presented in Appendix 1A, Table 1A.7, Column (1). The "Difference in Calories" column is the absolute value of the difference in calories between Food A and Food B. "Ratio of Calories" divides the calories in the higher calorie item by the calories in the lower calorie item.
2. Participants were presented with the ten pairwise tradeoffs listed above and were asked to indicate which of the two foods in each comparison provided more energy. These tradeoffs were presented via life-size photographs as well as written and verbal descriptions. Foods in each comparison have the same cost.
3. Participants were paid Rs. 1 for each correct response.

1.5.6 Summary of Reasons for Low Investment

There are a wide variety of possible reasons for low investment in the face of apparently high returns. However, neither liquidity, structural features of the economy, nor the disutility of labor appear to be constraints limiting caloric investment for many individuals with low caloric intake. Although it is possible that high discount rates play a role in low investment, the implied discount rates would be exceptionally high and the effect of discount rates could easily work in the opposite direction given that food is a consumption good which many people enjoy. The incentivized surveys described above do, however, provide evidence that beliefs and knowledge have the potential to play a role in low caloric intake. Only one-fifth of the respondents have correct beliefs about the returns to caloric intake. In addition, responses from both sets of incentivized survey questions provides strong evidence that respondents' beliefs regarding the caloric density of commonly consumed foods are frequently and substantially incorrect. Given the demographic diversity of the respondents, this evidence also demonstrates that these incorrect beliefs are not limited to a small sub-segment of the population. While this evidence is far from conclusive given that there are other factors (e.g. habit formation) which may limit change even if beliefs are accurate, the presence and extent of the incorrect beliefs does open the possibility that knowledge could play an important role in the low caloric consumption observed in the presence of high returns.

1.6 Conclusion

Over 800 million people still consume fewer calories than are recommended to maintain a healthy weight and most of these individuals are among the worlds' poorest (FAO 2011). While proof for or against the existence of nutrition-based poverty traps is beyond the scope of this paper, the randomized trial presented here suggest that low levels of adult nutrition may play an important role in productivity and economic development. Specifically, treated cycle-rickshaw drivers who were provided with an additional 700 calories per day increased in labor supply and earnings in a relatively linear fashion, leading to an increase in labor supply and earnings of approximately 10 percent in the fifth and final week of the study.

In addition, treated individuals improved performance on both physical and cognitive tasks in the laboratory. Physical performance in the laboratory followed a similar pattern of increase to labor market outcomes with no significant change at enrollment, but an increase of 7 percent among treated individuals relative to control individuals in the final week of the study. In contrast, the gains in cognitive performance were immediate, with a 12 percent gain among treated individuals relative to control individuals at enrollment. This gain was sustained, but not increased, by the fifth week of the study. However, the cognitive improvements observed in the randomized controlled trial suggest that the estimated gains may be a lower bound for the benefits associated with higher caloric consumption and that malnourishment may remain central to productivity despite a shift toward cognitive labor rather than physical labor in many developing countries.

These relatively large changes in productivity driven by changes in caloric intake generate correspondingly high estimated returns on investment. Although there is greater uncertainty in these estimates due to the greater number of assumptions required, the returns appear to be roughly 75 percent over six months. The high returns to caloric intake which accrue over short time horizons in the absence of binding liquidity constraints make the continued low caloric intake in India a puzzle.

This paper provides suggestive evidence generated via an incentivized surveys of hundreds of participants similar to those in the studies of returns that this apparent contradiction may be related to incorrect beliefs which limit investment in nutrition. Specifically, beliefs about the returns to increased caloric intake were highly variable and substantially inaccurate, with only 20 percent of respondents holding the belief that higher caloric intake will increase productivity. In addition, beliefs regarding the caloric densities of foods are often substantially wrong; in pairwise comparisons between commonly consumed foods with compensation for correct responses, respondents were unable to identify which foods contained greater numbers of calories per unit cost even at the rate expected by chance.

If incorrect beliefs about the returns to increased caloric intake and the caloric density of foods are indeed driving low investment, there would be a number of implications for both economics and for food policy. For example, India spends nearly one percent of GDP on food subsidies (Kumar and Soumya 2010).

Although many of these subsidies do target calorically dense foods such as grains, a variety of subsidies also apply to foods which provide few calories per rupee, such as dairy products. With poor knowledge about nutrition, these subsidies may actually negatively impact caloric intake (and productivity) rather than raise it. For example, if the price of milk declines, individuals may substitute dairy products for more calorically dense foods such as grains rather than supplement their previous consumption.

The evidence of incorrect beliefs reported here also suggests a number of new avenues of research exploring the consequences of these beliefs. For example, nutrition-based poverty trap models make strong predictions about who in society will be in high income/food or low income/food equilibria. Specifically, these models predict that the landed (or more broadly those with productive assets) will be hired first and attain the high equilibrium because their additional income allows them to consume more food and provide more effective labor at a lower wage. However, incorrect beliefs would break this linkage; individuals would not consume food to maximize their productivity. Instead, in a world in which individuals have incorrect beliefs about the returns to caloric intake, individual heterogeneity in preferences for food or features of the local diet (e.g. variety, food taboos, the caloric density of the staple grain) will drive which individuals consume sufficient calories and which individuals do not.

References

Banerjee, A.V., and E. Duflo. *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*. New York: Public Affairs, 2011.

Baumeister, R. and K. Vohs. “Self-regulation, Ego Depletion, and Motivation.” *Social and Personality Psychology Compass* 1.1 (2007): 115-128.

Behrman, J.R., and A.B. Deolalikar. "Will Developing Country Nutrition Improve with Income? A Case Study for Rural South India." *The Journal of Political Economy* 95.3 (1987): 492-507.

Bliss, C. and N. Stern. “Productivity, Wages, and Nutrition.” *Journal of Development Economics* 5 (1978): 331-362.

Bouis, H.E., and L.J. Haddad. "Are Estimates of Calorie-Income fx Elasticities Too High?: A Recalibration of the Plausible Range." *Journal of Development Economics* 39.2 (1992): 333-364.

Committee on Military Nutrition, Food and Nutrition Board. Commission on Life Sciences. National Research Council. “Predicting Decrements in Military Performance Due to Inadequate Nutrition.” Washington D.C.: National Academy Press, 1986.

Danziger, S., J. Levav, and L. Avnaim-Pesso. “Extraneous Factors in Judicial Decisions.” *PNAS* 108.7 (Apr. 2011): 6889-6892.

Dasgupta, P. and D. Ray. “Inequality as a Determinant of Malnutrition and Unemployment: Theory.” *The Economic Journal* 96.384 (Dec. 1986): 1011-1034.

Deaton, A., and J. Drèze. "Food and Nutrition in India: Facts and Interpretations." *Economic and Political Weekly* (2009): 42-65.

Deolalikar, A. “Nutrition and Labor Productivity in Agriculture: Estimate for Rural South India.” *The Review of Economics and Statistics* 70.3 (1988): 406-413.

Directorate of Census Operations. “Census of India, 2011.” Available online: http://www.census.tn.nic.in/. Cited Oct. 2013.

Fonseca-Azevedo, K. and S. Herculano-Houzel. “Metabolic Constraint Imposes Tradeoff Between Body Size and Number of Brain Neurons in Human Evolution.” *PNAS* 109.45 (Nov. 2012): 18571–18576.

Food and Agricultural Organization (FAO) of the United Nations. “FAOSTAT.” Available online: http://faostat.fao.org. Cited August 2013.

Food and Agriculture Organization (FAO) of the United Nations. “The State of Food Insecurity in the World.” September, 2010. Available online: http://www.fao.org/publications/sofi/en/. Cited Aug., 2011.

Friedlander, A.L., B. Braun, M. Pollack, J.R. MacDonald, C.S. Fulco, S.R. Muza, P.B. Rock, G.C. Henderson, M.A. Horning, G.A. Brooks, A.R. Hoffman, and A. Cymerman “Three Weeks of

Caloric Restriction Alters Protein Metabolism in Normal-weight, Young Men." *American Journal of Physiology – Endocrinology and Metabolism* 289 (2005): E446–E455.

Gailliot, M.T., R.F. Baumeister, B.J. Schmeichel, C.N. DeWall, J.K. Maner, E.A. Plant, D.M. Tice, and L.E. Brewer. "Self-control Relies on Glucose as a Limited Energy Source: Willpower is More Than a Metaphor." *Journal of Personality and Social Psychology* 92.2 (2007): 325–336.

Gerstein, D.E., G. Woodward-Lopez, A.E. Evans, K. Kelsey, and A. Drewnowski. "Clarifying Concepts about Macronutrients' Effects on Satiation and Satiety." *Journal of the American Dietetic Association* 104.7 (2004): 1151-1153.

Gopalan, C., R. Sastri, and S.C. Balasubramanian. *Nutritive Value of Indian Foods*. Hyderabad: National Institute of Nutrition, Govt. of India, 1989.

Government of the United Kingdom. Cycling Calorie Calculator. Available online: <http://www.nidirect.gov.uk/cycling-calorie-calculator>. Cited September, 2013.

Haddad, L. and H. Bouis. "The Impact of Nutrition Status on Agricultural Productivity: Wage Evidence from the Philippines." *Oxford Bulletin of Economics and Statistics* 53.1 (1991): 45-68.

Immink, M. and F. Viteri. "Energy Intake and Productivity of Guatemalan Sugarcane Cutters: An Empirical Test of the Efficiency Wage Hypothesis, Part I." *Journal of Development Economics* 9 (1981): 251-271.

Jensen, R. "The (Perceived) Returns to Education and the Demand for Schooling." *The Quarterly Journal of Economics* 125.2 (2010): 515-548.

Keys, A., J. Brozek, A. Henschel, O. Mickelsen, and H.L. Taylor. *The Biology of Human Starvation*. Minneapolis: University of Minnesota Press, 1950.

Kraut, H.A., and E.A. Muller. "Calorie Intake and Industrial Output." *Science* 104.2709 (1946): 495-497.

Leibenstein, H. "The Theory of Underemployment in Backward Economies." *Journal of Political Economy* 65.2 (Apr. 1957): 91-103.

Mirrlees, J. "A Pure Theory of Underdeveloped Economies." *Agriculture in Development Theory* 4 (1975): 84-108.

Samra, R.A., and G.H. Anderson. "Insoluble Cereal Fiber Reduces Appetite and Short-term Food Intake and Glycemic Response to Food Consumed 75 min Later by Healthy Men." *The American Journal of Clinical Nutrition* 86.4 (2007): 972-979.

Stiglitz, J.E. "The Efficiency Wage Hypothesis, Surplus Labour, and the Distribution of Income in LDCs." *Oxford Economic Papers* 28.2 (1976): 185-207.

Strauss, J. "Does Better Nutrition Raise Farm Productivity?" *Journal of Political Economy* 94.2 (1986): 297-320.

Subramanian, S. and A. Deaton. "The Demand for Food and Calories." *Journal of Political Economy* 104.1 (Feb. 1996): 133-162.

US Army Research Institute of Environmental Medicine. *Nutritional Status and Physical and Mental Performance of Special Operations Soldiers Consuming the Ration, Lightweight, or the Meal, Ready-to-eat Military Field Ration During a 30-day Field Training Exercise*. Natick, Ma. (1987).

Wolgemuth, J., M. Latham, and A. Chesher. "Worker Productivity and the Nutritional Status of Kenyan Road Construction Laborers." *The American Journal of Clinical Nutrition* 36 (1982): 68-78.

World Bank. "World Bank Development Indicators Database." Available online: http://databank.worldbank.org/data. Cited August 2013.

World Health Organization (WHO). "Global Database on Body Mass Index." Available online: http://apps.who.int/bmi/index.jsp?introPage=intro_3.html. Cited Aug., 2013.

Appendix 1A

Table 1A.1: Experimental Tasks at Enrollment and Endline of RCT

Task Name	Task Type	Task Description	Notes
Exercise cycle, short	Physical	Participants cycled as far as possible in 2 minutes on an exercise cycle in the office. Payment was a function of distance covered.	The proper use of the exercise cycle was demonstrated to each participant before they began.
Exercise cycle, long	Physical	Participants were given 60 minutes to cycle as much or little as they would like. However, payment was a function of distance covered.	The proper use of the exercise cycle was demonstrated to each participant before they began. Participants were provided with clean drinking water and a place to rest when not cycling.
Trip choices	Physical	Participants were first asked to make six choices between different sets of job opportunities using their cycle-rickshaws. A number was then drawn from a bowl to determine which of the choices is actually carried through. If a respondent selected a response that involved taking a load on that day, they completed that task immediately. If he selected a task that involves taking a load (or receiving a payment) in the future, that decision was noted in their record and he took the load (or was paid) on the appropriate day.	Each choice set contained an option to not take a trip. The participant did not receive a payment if he chose this option.
Crossing out symbols	Cognitive	Participants were given a workbook in which they were asked to locate and cross out specific symbols in a large matrix of randomly ordered symbols. This activity lasted 60 minutes. Participants could take breaks during this time but were paid as a positive function of the number of symbols correctly crossed out and a negative function of the number of symbols incorrectly crossed out.	Symbol sheets were printed in large font to mitigate concerns about eyesight impacting performance.
Job planning	Cognitive	Participants were presented with a hypothetical set of six possible jobs of varying durations and monetary values. They were asked to identify which jobs they would select if they had one hour available to complete the jobs and wanted to maximize the amount earned. Payment was a direct function of the value of the jobs selected.	If the sum of the duration of the jobs selected was greater than one hour, only those jobs which fit within an hour were included (and compensated).

Table 1A.1 (Continued): Experimental Tasks at Enrollment and Endline of RCT

Task Name	Task Type	Task Description	Notes
Dietary Recall	Survey	Dietary recall surveys elicited information on the food that the participant ate in a 24 hour period on the previous day. The survey followed a five step procedure: 1) Ask the participant to list of the foods and beverages consumed the previous day. 2) Collect the time at which the foods and beverages were consumed. 3) Probe for additional foods that might have been forgotten (e.g. snacks, side dishes, condiments). 4) Gather a detailed description of the amount and specific types of the foods and beverages consumed. 5) Final review of the information and probe for forgotten foods or details.	Step 4 is the most complex step. Participants were provided with a set of example containers for each type of food to help them estimate the amount consumed. For example, if a participant listed tea he was provided with a set of cups of varying sizes and asked to identify which was the closest to the amount consumed and to identify how full the cup was filled. The data on foods and beverages consumed was entered into a database which was run through a conversion program to translate the foods consumed into a total nutrient intake for the day.
Enrollment / Endline Survey	Survey	This survey elicited basic demographic information (e.g. age, household size, household assets), baseline work habits (e.g. whether the participant had any regular customers), risk preferences, anthropometric measures, general health and wellbeing, and alcohol and tobacco consumption.	
Daily Survey	Survey	The survey elicited information about labor supply (whether the individual worked, and if so the hours worked and the number of trips taken) and earnings the previous day. In addition, participants were asked to rate their energy on a scale of 1 to 5 and to indicate the number of meals eaten the previous day. Because participants were not asked to visit the study office on Sunday, information about both Saturday and Sunday was reported on Monday.	

Table 1A.2: Change in Caloric Intake During RCT

Dependent Variable	(1) Change in Calories (Endline-Enrollment)	(2) Meals	(3) Change in BMI (Endline-Enrollment)
Treated	178.35	0.15***	0.20*
	[235.16]	[0.04]	[0.12]
Mean of dependent variable	158.48	2.73	0.05
Observations	152	5,600	168
R-squared	0.00	0.02	0.02

Notes:
1. This table provides three measures of the effectiveness of the treatment at increasing caloric intake.
2. "Change in Calories" is the number of calories consumed in the 24-hour period before endline, excluding snacks consumed in the office, minus the number of calories consumed in the 24-hour period before enrollment. "Meals" is the number of meals eaten per day as reported daily by participants. "Change in BMI" is defined as a participant's BMI at endline (in the fifth week) minus their BMI at Enrollment.
3. Columns (1) and (3) have one observation per participant while in Column (2) the participant-day is the unit of observation. Standard errors are clustered by participant in Column (2).
4. *** p<0.01, ** p<0.05, * p<0.1

Table 1A.3: Linear Trends in Cycle-rickshaw Driver Labor Supply and Earnings

Dependent variable	(1) Work	(2) Hours if > 0	(3) Total hours	(4) Earn	(5) ln(earn+1)	(6) IHS(earn)
Treatment	0.019	-0.353*	-0.086	-11.722*	0.034	0.047
	[0.021]	[0.206]	[0.244]	[6.282]	[0.110]	[0.124]
Treatment*Day in study	0.003***	0.016*	0.036***	0.611**	0.013**	0.015**
	[0.001]	[0.010]	[0.012]	[0.303]	[0.005]	[0.006]
Mean of dependent variable	0.81	8.88	7.21	163.57	4.21	4.77
Observations	5,675	4,604	5,669	5,675	5,675	5,675
R-squared	0.081	0.144	0.119	0.112	0.086	0.086

Notes:
1. This table provides the results of regressions examining changes in labor supply and earnings among treated participants across time. "Work" is a binary variable indicating whether the participant worked that day. "Hours if >0" is the total number of hours worked, conditional on working. "Total hours" is the number of hours worked per day where hours are zero if the participant did not work. "Earn" is the total daily earnings as a rickshaw driver in Indian Rupees. This variable takes the value zero if the participant did not work. "ln(earn+1)" is the log of earnings + 1 where earnings is defined as previously described. "IHS(earn)" is the inverse hyperbolic sine of earnings. "Day in study" indicates the number of days from enrollment for that participant. The participant-day is the unit of observation.
2. Regressions include stand location, calendar week, and study day fixed effects as well as controls for baseline work habits which are detailed in Appendix 1B. Results are less precisely estimated, but qualitatively similar if these controls are omitted.
3. Robust standard errors clustered by individual are in brackets.
4. *** p<0.01, ** p<0.05, * p<0.1

Table 1A.4: Cycle-rickshaw Driver Labor Supply and Earnings

Dependent Variable	(1) Work	(2) Hours if >0	(3) Total hours	(4) Earn	(5) ln(earn+1)	(6) IHS(earn)
Treatment	0.07***	-0.07	0.57*	-0.81	0.27**	0.32**
	[0.02]	[0.29]	[0.34]	[8.92]	[0.12]	[0.13]
Mean of dependent variable	0.81	8.88	7.21	163.57	4.21	4.77
Observations	5,675	4,604	5,669	5,675	5,675	5,675
R-squared	0.05	0.13	0.09	0.08	0.05	0.05

Notes:
1. This table provides the results of regressions testing for changes in labor supply and earnings among treated participants aggregating across the five weeks of the study. "Work" is a binary variable indicating whether the participant worked that day. "Hours if >0" is the total number of hours worked, conditional on working. "Total hours" is the number of hours worked per day where hours are zero if the participant did not work. "Earn" is the total daily earnings as a rickshaw driver in Indian Rupees. This variable takes the value zero if the participant did not work. "ln(earn+1)" is the log of earnings + 1 where earnings is defined as previously described. "IHS(earn)" is the inverse hyperbolic sine of earnings. The participant-day is the unit of observation.
2. Regressions include stand location, calendar week, and study week fixed effects as well as controls for baseline work habits which are detailed in Appendix 1B. Results are less precisely estimated, but qualitatively similar, if these controls are omitted.
3. Robust standard errors clustered by individual are in brackets.
4. *** p<0.01, ** p<0.05, * p<0.1

Table 1A.5: Cycle-rickshaw Driver Daily Earnings (High Attendance)

	(1)	(2)
Dependent Variable	Work	Earn
Treated (Week 1)	0.02	-0.26
	[0.03]	[11.25]
Treated*Week 2	0.06*	8.16
	[0.03]	[10.26]
Treated*Week 3	0.06	10.06
	[0.04]	[11.42]
Treated*Week 4	0.05	5.06
	[0.03]	[12.00]
Treated*Week 5	0.09*	20.53
	[0.05]	[13.33]
Mean of dependent variable	0.83	170.58
Observations	3,883	3,883
R-squared	0.05	0.11

Notes:
1. "Work" is a binary variable indicating whether the participant worked that day. "Earn" is the total daily earnings as a rickshaw driver in Indian Rupees. This variable takes the value zero if the participant did not work. The participant-day is the unit of observation.
2. The sample is limited to participants who had at least 90 percent attendance.
3. Regressions include stand location, calendar week, and study week fixed effects as well as controls for baseline work habits. Results are less precisely estimated, but qualitatively similar if these controls are omitted.
4. Robust standard errors clustered by individual are in brackets.
5. *** p<0.01, ** p<0.05, * p<0.1

Table 1A.6: Heterogeneous Treatment Effects by BMI

	(1)	(2)
Dependent Variable	Work	Earn
Treatment	0.03	-5.65
	[0.03]	[15.30]
Low BMI	-0.07**	-26.76*
	[0.03]	[15.67]
Low BMI*Treatment	0.07	9.20
	[0.04]	[19.14]
Mean of Dependent Variable	0.81	163.34
Observations	5,648	5,648
R-squared	0.05	0.09

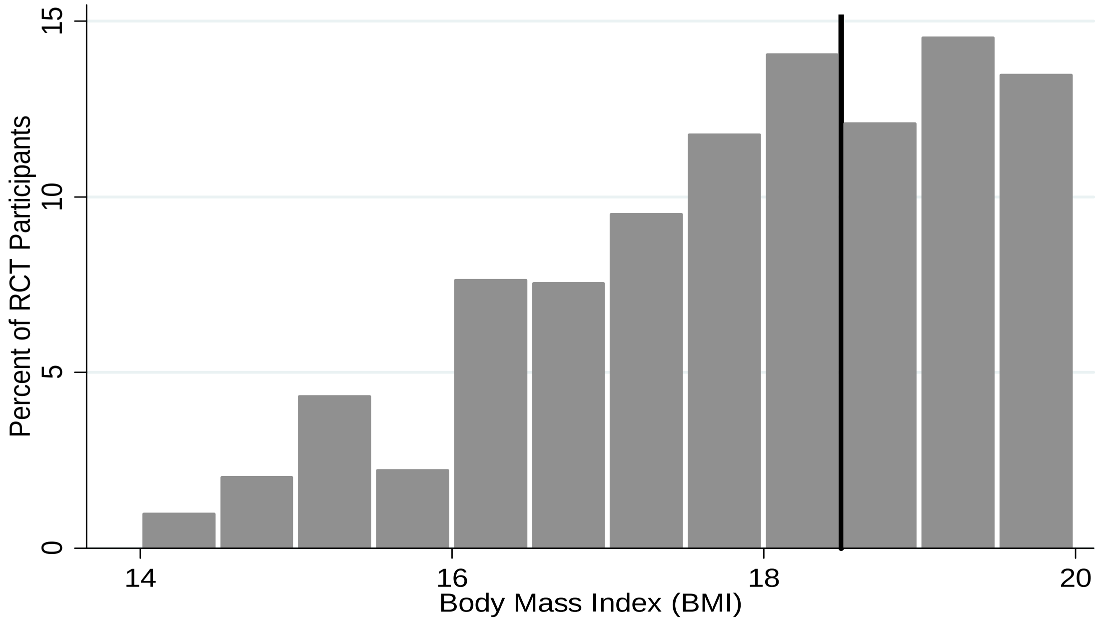
Notes:
1. This table presents results of regressions of participants' labor supply and earnings as a function of treatment status and BMI.
2. "Work" is a binary variable indicating whether the participant worked that day. "Earn" is the total daily earnings as a rickshaw driver in Indian Rupees. This variable takes the value zero if the participant did not work. "Low BMI" is defined as a BMI under 18.5, the World Health Organization cutoff for being "underweight," at enrollment.
3. Regressions include stand location, calendar week, and study week fixed effects as well as controls for baseline work habits. Results are less precisely estimated, but qualitatively similar if these controls are omitted.
4. Robust standard errors clustered by individual are in brackets.
5. *** p<0.01, ** p<0.05, * p<0.1

Table 1A.7: Demographics of Survey Respondents

Question types		(1) Caloric Density, Food Habits and Preferences	(2) Returns
Age		41.37	39.88
		[12.48]	[12.46]
People in household		4.72	4.16
		[2.04]	[1.74]
Monthly household expenditures (Rs)		5430.77	5552.74
		[2808.23]	[3135.65]
Body Mass Index (BMI)		18.13	18.21
		[1.51]	[1.63]
Male		0.68	0.55
		[0.47]	[0.5]
Married		0.89	0.89
		[0.32]	[0.32]
Years of education		4.52	4.72
		[3.96]	[4.16]
Urban		0.44	0.66
		[0.5]	[0.47]
Meals per day		2.78	2.77
		[0.47]	[0.47]
Hungry in typical day			
	"No"	4.02 %	6.04 %
	"< 1 hour/day"	38.84 %	38.93 %
	"1-2 hours/day"	20.09 %	10.74 %
	"2-5 hours/day"	33.93 %	43.62 %
	">5 hours/day"	3.13 %	0.67 %
Religion			
	Hindu	85.71 %	93.29 %
	Muslim	12.05 %	4.03 %
	Other	1.79 %	2.68 %
Years in current profession		15.74	12.59
		[11.92]	[11.01]
N		222	149

Notes:
1. Statistics are presented mean [standard deviation].
2. Common professions of respondents include: agricultural labor, construction, casual laborer, load carrier/porter, auto driver, housekeeper/sweeper, textile workers, food services, and vendors.
3."Urban," “Male,” and "Married" are binary variables.

Figure 1A.1: Daily Attendance



Note: Figure is based on measurements taken at enrollment.

Appendix 1B

Literature – The Impact of Caloric Intake on Economic Productivity

There has been significant interest in the effect of caloric intake on productivity and labor market outcomes for many years. Yet, despite many and varied approaches to studying the changes in productivity associated with caloric intake, the conflicting results and difficulties inherent in isolating the casual effect of increased caloric consumption on production have made it difficult to draw firm conclusions about the existence of this relationship and its strength.

The earliest of these studies, relying on quasi-exogenous changes in caloric intake among men working at various tasks (e.g. coal mining, embankment construction) during war time, found that caloric availability was positively correlated with output (Kraut and Muller 1946). Strauss (1986) improved on these correlational studies by utilizing a vector of food prices, farm assets, and household demographic characteristics to instrument for average caloric intake per adult equivalent to estimate a farm production function in Sierra Leone. The estimated production function showed a strong positive relationship between caloric consumption and farm production, with higher returns at lower levels of consumption. Yet, given the relatively high price of calories and the levels of caloric intake in Sierra Leone, the increased productivity did not account fully for the cost of the additional calories for the median individual. Using panel data from southern India to estimate both wage equations and farm production functions with individual fixed effects, Deolalikar (1988) found no significant impact of caloric intake on wages or farm output once stature (weight for height) was accounted for.

Two studies have also examined the impact of adult caloric supplementation on productivity through experimental methods in an effort to address the endogeniety of caloric consumption. Immink and Viteri (1981) find no significant change in production when treated participants were provided with an additional 350 calories per day. However, due to logistical constraints, this experiment relied on two villages randomized at the level of the village. In addition, laborers in this study worked in groups of four and were paid one quarter of the group’s total production, potentially weakening individual incentives. Wolgemuth et al. (1982) were able to randomize at the level of the individual and provided the treated individuals with an additional 800 calories per day relative to the control group. However, because participants in this study were construction workers, productivity was estimated by visual inspection of the quantity of earth moved and required adjustments of productivity measures to account of differences in output across types of tasks.[37] Although production per day showed a small marginally significant increase in the intention to treat estimate, the number of days worked on the roads project declined significantly among treated individuals making an overall assessment of economic productivity changes difficult.

Literature – The Impact of Caloric Intake on Physical and Cognitive Function

In contrast to the generally inconclusive results regarding the impact of caloric intake on economic productivity, there is substantial evidence from the physiology literature that low caloric intake, even over short durations such as a week, is associated with declines in a variety of indicators of physical performance. Of note, declines appear to become more pronounced as body weight drops further (Bender and Martin 1986; Taylor et al. 1956; Friedlander et al. 2005; Committee on Military Nutrition 1986). For clear reasons, there are a limited number of controlled trials studying low caloric intake over extended time periods.

However, one well designed within-subjects trial, the Minnesota Semi-Starvation Study, does exist. Participants, all young males, were volunteers drawn from a pool of conscientious objectors. The study lasted 56 weeks with a 12 week baseline, a 24 week "starvation" period restricted to 1,560 calories per day, and 20 weeks of varying "recovery" diets. Participants lived at the study site and underwent regular and extensive physiological testing in addition to maintaining a fairly extensive exercise regimen. The study provides strong evidence of substantial declines in physical performance including maximal performance capacity as measured by a treadmill test (50-70 percent decrease in time to exhaustion and overall fitness score) and strength (30-40 percent decrease on each measure) during the starvation period (Keys et al. 1950).[38] Although this study differs from the context of long-run deprivation that many of the worlds'

[37] The specific compensation method for workers is not explicitly stated in the paper. Wages are mentioned, however, the structure of the wages is not. Given that the labor provided was for a public works project, it is likely that there was simply a flat days wage for work on the project such that workers most likely faced relatively weak incentives.

[38] Participants were not monetarily incentivized during these physical tests. However, given that the study population consisted of conscientious objectors willing to live at the study site for an extended period, it is likely that participants were highly internally motivated to provide maximal effort. Study administrators also reported high levels of effort

malnourished suffer from childhood, the continued decline throughout the starvation period suggests that the participants were unable to fully compensate for low caloric intake and maintain performance over fairly long horizons.[39]

In addition to declines in physical performance, participants in the Minnesota study complained of decreased alertness, lack of self-control, and general apathy. A number of other studies in psychology and other disciplines have found similar cognitive changes related to nutrition even over just a few hours (e.g. Baumeister and Vohs 2007; Danziger et al. 2011; Gailliot et al. 2007). Although many studies find that cognitive ability (maximal performance) is maintained, persistence and motivation for cognitive work and thought decline substantially. For example, the US Army found that soldiers consuming 2,000 calories per day maintained their tested cognitive ability but completed just over half as many cognitive exercises as soldiers consuming 2,700 calories/day (US Army 1987). Although perhaps initially surprising, these changes likely reflect the fact that despite accounting for only 2 percent of body weight the brain consumes roughly 20 percent of the energy used by the body; hence, as with any organ requiring substantial energy inputs, limiting the total energy available is likely to constrain the brain's ability to function (Fonsec-Azevedo and Herculano-Houzel 2012).

Characteristics of Cycle-Rickshaw Drivers in the Randomized Controlled Trial in Chennai, India

Chennai is a city of approximately 4.7 million people in the state of Tamil Nadu, located on the eastern coast in the south of India (Census of India 2011). Participants for the RCT were drawn from the population of cycle-rickshaw drivers, an all-male profession. Members of this profession drive large three wheeled vehicles to transport passengers, and occasionally, luggage or goods. The vehicles are large enough to

---

by participants during these tests, reporting that "runs were terminated at a state of actual or near collapse" in the later periods of the study.

[39] Participants did experience a decline in their basal metabolic rate (the energy used in basic bodily processes such as respiration, excluding energy for digestion) from 1570 calories per day to 960 calories per day throughout the starvation period. This increased "efficiency" is one mechanism the body can use to limit weight loss despite low caloric intake. However, this decline had tapered substantially by week 24 due to basic physiological limits (Keys et al. 1950).

comfortably fit two adults in addition to the driver, although can be used for loads up to a few hundred kilos. In Chennai, cycle-rickshaws are primarily used in a 2.5 square kilometer area in the center city due to the narrow streets which make travel by larger vehicles difficult there.

Although Tamil Nadu is a relatively wealthy state, rickshaw drivers are in lower socio-economic brackets. The median participant has 3 years of education and can write numbers but cannot write the letters of the alphabet. In addition, 90 percent of the participants are in historically disadvantaged castes and 70 percent have ration cards which entitle them to subsidized foods and cooking gas.[40] The median participant has a family size of 4, lives in a one room house with electricity but without a bathroom, and has four small appliances (e.g., rice cooker or fan). However, 25 percent of the participants are migrants who typically sleep in their rickshaw while they are working in the city. The average weekly household income of Rs. 1,523 results in a per capita income of $1.10 per day for study participants and their families.[41] Although not much information is available on consumption, of the average daily earnings of Rs. 165 (Rs. 200 conditional on working) rickshaw drivers report spending an average of Rs. 14 per day on tobacco and Rs. 68 per day on alcohol.[42] Among households with similar per capita daily expenditure in the greater Chennai area, 34 percent of overall household expenditures were dedicated to food.[43]

Participants are typically in their 40s and 50s (mean age 46) and have a small physical stature with an average height of 161cm (5 feet, 3 inches) and an average weight of 46 kilos (101 pounds) (the BMI distribution of participants is presented in Figure 1A.1. Based on a random sample of drivers in the city, roughly two-thirds of rickshaw drivers have a BMI under 20). While only 14 percent of participants rate

---

[40] "Historically disadvantaged castes" are defined as belonging to a schedule caste, scheduled tribe, or "other backward class."

[41] The average exchange rate during the study was roughly 45Rs to 1 USD.

[42] Most rickshaw drivers spend money on food, alcohol, and tobacco during the day as they see fit and then give the remainder of their earnings to their wives to allocate to other needs. Because wives are typically responsible for the decisions on how to spend that money, very few men are able to provide information on the allocation of those funds.

[43] This figure was calculated from the 64th round of the NSS (2007-2008). The sample is limited to households in the greater Chennai region to match the RCT population as closely as possible while still maintaining a reasonable sample size. In addition, I restrict the sample to households with a per capita daily expenditure of Rs. 23 to Rs 43 (The average per capita earnings for the RCT participants was Rs 49 in 2012 rupees, which corresponds to Rs. 33 in 2008 rupees. The sample is based on a Rs. 10 per day band around this figure.) Using the NSS sampling weights, the average daily expenditure on food in these households is Rs. 10.7 per person while the average total expenditure is Rs. 31.

their general health as being “somewhat unhealthy” or “unhealthy,” the average participant also reports missing 1 day of work every two weeks due to ill health.

With an average experience of 20 years, participants had generally worked in the profession for most or all of their working lives. A typical work week is 5.7 days per week and 9 hours per day, conditional on working. Three-quarters of the population rents their rickshaw, typically on a weekly or monthly basis for roughly Rs. 190 per week.

Caloric Intake in the Randomized Controlled Trial

A number of steps were taken to maximize the increase in caloric intake among treated individuals. First, participants were asked to visit the office during non-meal hours (i.e. between 10am and 12:30pm and from 3pm to 6:30pm). In addition, participants were provided with “snack” foods (e.g. potato chips, biscuits, fried dough) rather than foods that typically constitute a meal. Finally, participants were provided with an additional 700 calories per day such that even if some crowd out did occur, there was likely to be a net increase. It is also worth noting that if some crowd out did occur, the expected ROI would be higher in that the changes in labor supply and earnings would then have occurred with less than 700 calorie supplementation of the treatment participants’ daily caloric intake.

Three different measures were taken to proxy changes in caloric intake between treated and control individuals during the study: 1) caloric intake as reported in a 24-hour dietary recall survey at enrollment and endline, 2) number of meals consumed reported daily, and 3) BMI (or weight) as measured at enrollment and endline.

All three measures suggest an increase in caloric intake in the treatment group. Although the standard errors are quite large, such that it is difficult to calibrate the exact magnitude of the difference, the estimated coefficient on treatment is positive when comparing changes in caloric intake in the dietary recall surveys despite excluding caloric intake in the lab. In addition, the coefficient on the number of meals consumed each day is positive and significant, suggesting that the treatment did not crowd out consumption of other food.[44] Finally, while control participants do not change BMI over time, treated participants have a marginally statistically significant ($p = 0.09$) increase of 0.2 in BMI, corresponding to a weight gain of roughly 0.5 kg (or approximately 1.1 pounds) (Appendix 1A, Table 1A.2).[45]

Controls Used in Labor Supply and Earnings Regressions

At enrollment, participants reported the number of days worked, the average number of hours per day conditional on working, and earnings conditional on working for the previous week. These variables are included as baseline controls in all regressions of labor supply and earnings.

Bounding Treatment Effects in the Randomized Controlled Trial

The bounding process follows the method proposed and described by Lee (2003), with minor adjustments to account for the structure of the data. This method relies on the trimming of extreme observations in the condition with lower attrition to bound the estimated treatment effect. Specifically, the process begins by calculating the fraction of observations that must be censored from the data by differencing the fraction of non-attriters between the groups and dividing by the fraction of non-attriters in the group with lower attrition.[46] Due to the panel nature of the data in this study, there are many possible measures of attrition. Hence, I generate five different measures of attrition and use the smallest and the largest to create a liberal and a conservative bound on the estimated treatment effect. The measures of attrition used are: 1) never return after enrollment, 2) do not attend endline, and 3) thru 5) attendance less than 25 percent, 50 percent, and 75 percent, respectively. The smallest trimming fraction, 5.3%, is a result attrition measure (1), while the largest, 15.9 percent is a result of measure (5). After calculating the trimming fraction, the individuals

[44] Snacking is relatively uncommon and meals purchased at street stalls are of fairly uniform size, such that the number of meals per day is likely to correlate highly with caloric intake in this population.
[45] One pound of weight corresponds to 3500 calories. Hence, 700 calories per day over five weeks with no changes in behavior of basal metabolic rate would increase weight roughly 5 pounds. However, given that treated participants increase labor supply substantially and may experience some increase in BMR, this smaller gain is not surprising.
[46] Although attrition is generally greater in control groups, the results of Lee (2003) hold regardless of whether attrition is higher in the treatment or control. This analysis calls for trimming from the control group (because attrition is lower in that condition). Hence, I use the fraction of non-attriters in the control as the denominator in calculating the trimming fraction.

with the highest (lowest) values of the outcome, where the outcome is the average value of the dependent variable over all observations, are removed from the dataset and the treatment effect is estimated on the trimmed dataset. The results of this bounding procedure for labor supply and earnings outcomes are reported in Table 1.7a and Figure 1.8.

A similar procedure was used to bound the estimated treatment effect on the enrollment and endline payments. However, in this case, the trimming fraction was calculated using the fraction of individuals in each condition who did not attend endline and payment on the enrollment tasks were used to determine which observations to censor. The results of this bounding procedure for earnings on laboratory tasks are reported in Table 1.7b.

Calibrating the Disutility of Labor

By the end of the five-week RCT, these individuals face a choice between their current activities and a bundle that includes spending an additional Rs. 6 for food, working 12 percent more, and earning Rs. 19 more each day (Table 1.5). Hence, participants must be willing to increase labor supply 12 percent for an additional Rs. 13 per day. The median number of trips per day reported by participants is 6. Hence a 12 percent increase in labor supply would be roughly 0.7 trips per day.[47] Because fractional trips are not possible, I scale both the earnings and trips by 1.4, resulting in a tradeoff of one trip for Rs. 21.

I use two sources data to assess willingness to accept at this rate. The first source is one of the experimental tasks in which RCT participants were offered additional paid fares. Specifically, participants were asked to make a series of choices between taking or not taking paid trips with varied rates, loads, and time horizons and one of their choices was randomly selected and carried out to ensure the choices were incentive compatible. One of the choices offered to participants was between not taking a trip and not earning any additional money, taking a 1 km journey with a 100 kg load (in addition to the cycle and driver)

[47] This approach abstracts from disutility from simply waiting for fares. This is done because it is impractical to offer entire days of work that can be monitored by research staff to ensure compliance. However, given the physical difficulty of the task, waiting is likely to be significantly less taxing than actual labor. In addition, this calibration also ignores the positive utility from additional food consumption which would work in the opposite direction.

for Rs. 20, or taking a 1 km journey with a 150 kg load for Rs. 30.[48] The 1 km journey with a 100 kg load is representative of a fairly typical journey given that most rickshaws will carry 1 to 2 passengers (each weighing roughly 50 kgs) and typically travel short distances.

The second dataset was generated from 150 actual offered rickshaw trips around the city. Research staff were randomly assigned a commonly traveled route (e.g. the bus station to the train station) and a price and were told to offer no more than that price and determine whether the driver would accept. If the driver accepted, the trip was taken and the fare paid. If the driver declined, the staff member moved to the next route and price combination.

In the RCT task, 85 percent of participants opted to take one of the two trips offered. This is consistent with the willingness to accept found in the 150 additional offered rickshaw trips in which 83 percent of participants were willing to accept a trip for Rs. 21.

In short, roughly 85 percent rickshaw drivers are willing to take fares at the rate implied by the labor supply and net earnings changes observed in the RCT. Hence, even if there is no utility provided by the consumption of additional food, there appear to be positive utility returns to increased consumption for the vast majority of rickshaw drivers.

[48] A staff member would accompany the driver on the trip to ensure that the driver completed the full distance required.

2 The Economic Costs of Low Caloric Intake: Evidence from Agricultural Production in India[*]

2.1 Introduction

The randomized trial detailed in Chapter 1 provides internally valid evidence that the labor supply, earnings, and physical and cognitive function of cycle-rickshaw drivers improve with increased caloric intake. Although cycle-rickshaw drivers were selected as a study population in part due to their similarity to many individuals in the informal labor market (e.g. flexible labor supply, low caloric intake), this group is still specific. In order to address potential concerns about the external validity of the RCT as well as to examine the impact of changes in caloric intake at lower levels of intake, I conducted a second analysis. This analysis examines the impact of the reduced caloric intake on agricultural production, an industry which accounts for 52 percent of employment in India (National Academy of Agricultural Sciences 2012; World Bank 2012).

To study the relationship between caloric consumption and economic output in agriculture, I exploit a natural experiment drawing on the quasi-random declines in caloric intake caused by fasting during Ramadan, a month-long Muslim holiday. Following background information about the holiday and the context of the study, the analysis begins by estimating the caloric decline among Muslims during Ramadan in India via the consumer expenditure portion of the National Sample Survey (NSS). The estimated decline in caloric intake is roughly 700 calories per person per day for rural agricultural Muslim households.[1]

[*] I am deeply grateful to Sendhil Mullainthan, Lawrence Katz, Michael Kremer, and David Cutler for feedback and encouragement. For helpful discussions, I thank David Laibson, Rohini Pande, Asim Khwaja, Rick Hornbeck, Ed Glaeser, George Loewenstein, Kevin Volpp, Shawn Cole, Max Bazerman, David Bloom, Supreet Kaur, Frank Schilbach, Dan Bjorkegren, Ian Tomb, Joana Naritomi, Raluca Dragusanu, Laura Trucco, Anjali Adukia and seminar participants at Harvard University. I thank Dave Donaldson for sharing data.

[1] Although this study examines a decline in consumption, the observed levels of consumption are ones which are relevant to consider. Specifically, the declines during Ramadan result in an overall average consumption of roughly 1,500 to 1,600 calories per person per day (See Section 2.3 for details of this analysis). Yet, one-quarter of the Indian population consumes less than 1,625 calories per day on a regular basis (Deaton and Dreze 2009).

I then utilize a differences-in-difference approach to assess the impact of reduced caloric intake on agricultural production at the crop-district-year level. Specifically, to estimate the impact of this decline in calories on agricultural production, I exploit three sources of variation in the overlap between fasting and the labor intensive portions of cropping cycles (sowing and harvesting). The first, generated by the fact that Ramadan cycles throughout the calendar year, is variation for a given crop-district combination over time. The second, variation among crops within a district-year, is the result of natural variation in cropping cycles for different plants. Finally, different climatic patterns across space generate variation in cropping cycles between districts, resulting in spatial variation for a given crop within a year.

The results of this analysis suggest that overlap between Ramadan and sowing/harvesting leads to economically meaningful declines in total agricultural production, both in weight and in value. The estimated decreases in output imply a decline in productivity of approximately 20 to 40 percent per Muslim individual. Building on these baseline results, I also examine heterogeneity in the production declines arising from the uneven spatial distribution of Muslims in India.[2] Reassuringly, in districts with very few Muslims there is no decline in production. However, districts with many Muslims experience substantial production declines when sowing or harvesting overlaps the fasting period. This heterogeneous effect supports the conclusion that Ramadan, rather than other factors, drives the production declines.

Finally, I turn to an analysis of behavioral changes during the holiday to assess whether features of the holiday beyond to the decline in caloric intake underlie the declines in production. Specifically, I discuss and assess three other potential behavioral changes during the holiday as possible channels: changes in labor supply due to religious or social obligations, sleep deprivation, and dehydration. As an omnibus test, I examine whether overlap between sowing/harvesting and the period following Ramadan also generates declines in production. Given the slow recovery from reduced caloric intake but rapid recovery from dehydration and sleep deprivation, the persistence of the effects observed in this test simultaneously provides evidence in favor of nutrition driving the declines and against these three other forces playing key

[2] The fraction of the population which is Muslim varies substantially across districts, ranging from less than 0.1 percent to over 40 percent (Census of India 1961).

roles.[3] Drawing on evidence from relevant literatures and additional direct empirical tests utilizing both agricultural production data and additional data sources, I also assess each of these potential channels individually. These empirical tests draw on both the agricultural data and additional data sources such as the employment portion of the National Sample Survey and the ICRISAT village level studies survey. Each of these sources of evidence and the data analyses are consistent with an effect on production driven primarily by changes in caloric consumption rather than other behavioral shifts. So, while it is difficult to fully rule out all possible alternative channels in any natural experiment, these analyses suggest that these three channels are unlikely to be significant drivers of the changes in production during Ramadan in India.

As in the previous analysis of cycle-rickshaw drivers, the substantial impact of low caloric intake on productivity suggests a high return on investment: the estimated 1-month return to a 700 calorie per day increase in consumption is over 200 percent. These positive returns are robust to a wide variety of different assumptions about the cost of calories and the return on investment.[4]

The remainder of this paper is divided into six parts. Section 2.2 provides background about agriculture and caloric intake in India as well as the Ramadan holiday. Section 2.3 estimates the change in caloric intake for rural Muslim agricultural households during Ramadan. The differences-in-differences approach examining the impact of decreased caloric intake during Ramadan on agricultural production is presented in Section 2.4. Section 2.5 examines three alternative changes in behavior during Ramadan which have the potential to drive production declines and provides evidence suggesting these channels are not primary contributors to the observed declines. Calculations of the returns associated with the estimated productivity changes are detailed in Section 2.6. Section 2.7 concludes.

[3] While recovery from dehydration and sleep deprivation occurs in hours or days, recovery from low caloric intake often takes weeks (Drummond et al. 2006; Dinges et al. 1997; Sawka et al. 2007). Details of this test are provided in Section 2.5.

[4] A wide variety of alternative assumptions about the cost of calories and productivity changes, and the returns implied by these assumptions, are detailed in Section 2.6.

2.2 Background

Ramadan

Ramadan is a month-long Muslim holiday observed primarily through fasting during daylight hours. Fasting includes abstinence from both food and liquids and is obligatory for practicing Muslims with the exception of children, the elderly, individuals who are ill, infirm, or traveling, and women who are pregnant or breast feeding. Muslims are also expected to abstain from smoking, sexual relations, and swearing during daylight hours throughout the holiday. In addition, there is an added emphasis on prayer, reading the Koran, and charity during this time (Blackwell 2009; Ahmad et al. 2012).

The holiday is lunar, shifting by roughly eleven days per year, and cycling through the calendar year approximately once every 30 years. Ramadan is followed by Eid, a holiday marking the end of Ramadan.[5] During Eid, Muslims are not allowed to fast, and typically engage in a special prayer in a communal area and visit family and friends. Eid is a minimum of one day but can last up to three days (Blackwell 2009).

Caloric Intake and Body Mass Index in India

The Indian Planning Commission's recommended caloric intake has typically been 2,100 calories per adult in urban areas and 2,400 calories per adult in rural areas (Sharma 1999). Caloric intake in India has, however, remained significantly below these levels for many years and, of note, has been declining over time despite strong economic growth (Deaton and Dreze 2009). In addition, the distribution of calories in India is quite skewed such that mean caloric intakes are substantially higher than median caloric intakes. The low caloric intakes lead to correspondingly low body mass indices (BMI), with over half of the population below the WHO cutoff for underweight (BMI < 18.5) in 1971, the median year in this study.[6] Roughly one-third of the population remains underweight at present (Deaton and Dreze 2009; WHO 2013).

[5] There are two holidays in the Muslim calendar referred to as Eid: Eid al-Fitr and Eid al-Adha. However, this paper will always reference Eid al-Fitr, which follows Ramadan, when referring to Eid.
[6] As a benchmark, a person who is 5 foot 8 inches tall (1.73 m) would have a BMI of 18.5 at 122 pounds (55.3 kg).

Agriculture in India

Agriculture is a critical sector in the Indian economy. Although the share of GDP generated from agriculture has declined from 43 percent in 1960 to 17 percent at present, agriculture still accounts for 52 percent of employment (National Academy of Agricultural Sciences 2012; World Bank 2012). Driven in part by technology improvements during the Green Revolution, growth in this sector has been relatively rapid with a four-fold increase in real value over the same period (FAO 2013; Evenson and Gollin 2003). India's primary crops are food grains such as rice, wheat, and millet as well as cash crops such as sugarcane, oilseeds, and peanuts. Section 2.4.1 provides additional details about the agricultural data used in this analysis.

2.3 Caloric Consumption During Ramadan

Although Muslims fast during daylight hours during Ramadan, it is possible that there could be substitution of food consumption across time to the evenings and mornings, limiting or eliminating declines in caloric intake. Hence, I directly estimate the change in caloric intake in order to confirm that caloric consumption does indeed decline as well as to estimate the magnitude of the change.

2.3.1 Data

This analysis utilizes the consumer expenditure portion of the 60th, 62nd, and 64th rounds of India's National Sample Survey (NSS, Schedule 1).[7] The dataset provides monthly household level consumption data for nearly 150 food items allowing the calculation of each of the 36,000 households' per capita caloric intake.

[7] Additional information about the NSS can be found on the Indian Ministry of Statistics and Progamme Implementation (MOPSI) website: http://mospi.nic.in. While it would be preferable to include NSS rounds during the period covered by the agricultural production data, earlier rounds of the NSS do not include the survey date and hence do not permit a calculation of the extent of the overlap between the survey period and Ramadan. The 60th round was conducted in 2004, the 62nd in 2005 and 2006, and the 64th in 2007 and 2008.

2.3.2 Results - Changes in Caloric Intake

I estimate caloric declines among Muslim households during Ramadan according to the following equation:

$$(2.1) \quad C_{idt} = \beta_0 + \beta_{1r}\mathbf{R}_i + \beta_2 F_{it} + \beta_{3r}\mathbf{R}_i{*}F_{it} + \beta_2 E_{it} + \alpha_r\mathbf{R}_i{*}E_{it} + \Theta_{dt} + \lambda_m + \varepsilon_{idt}$$

C_{idt} is calories per capita per day in household i, in district d, in year t. $\mathbf{R}_i$ is a vector of binary variables for the household's religion. F_{it} is an indicator variable denoting full overlap of the survey period and Ramadan. E_{it} is an indicator for overlap between the survey period and Eid, and Θ_{dt} and λ_m are district-year and month of survey fixed effects, respectively. The primary sample is limited to rural households whose primary occupation is agricultural work and which have either no overlap or complete (29 day) overlap between the survey period and Ramadan.

The results demonstrate a fairly substantial and significant decline in caloric intake among rural Muslim households during Ramadan (See Table 2.1). Although standard errors are large, the decline appears to be approximately 700 calories per person per day, an estimate consistent with not eating a midday meal.[8] This decline is substantial relative to total caloric consumption, which is approximately 2200 calories per capita per day despite the fact that most individuals in this population engage in heavy physical labor. Results are robust to a variety of methods of topcoding used to address implausibly high consumption levels (See Appendix 2A, Table 2A.1).

[8] The marginal increase in caloric intake during Ramadan for non-Muslims may be due to the fact that during these rounds of the NSS Ramadan fell between September and November, the season when the many of holidays and festivals for other religions occur (e.g. Diwali, Dussehra, Navaratri, Ganesh Chaturhi, Krishna Janmashtami). However, as noted in the previous footnote, it is not possible to use earlier NSS rounds due to omission of the survey dates in those rounds. Of note, the decline in intake among Muslim households does not appear to substantially alter the proportional macronutrient (i.e. carbohydrate, fat, protein) content of the consumption.

Table 2.1: Daily Calories per Capita in Rural Agricultural Households

	(1)	(2)
Dependent variable	Daily Per Capita Calorie Availability	
Topcoding Method	No topcoding	At 99th percentile
Muslim	-107.5***	-92.46***
	[41.54]	[28.86]
Full survey period overlaps Ramadan	455.1*	427.2
	[259.1]	[271.5]
Muslim*Full survey period overlaps Ramandan	-612.6	-725.5**
	[384.8]	[368.9]
Mean of Dependent Variable	2263.49	2200.51
Observations	36,618	36,618
R-squared	0.1	0.21

Notes:
1. This table tests for changes in caloric consumption during Ramadan in Muslim households. The dependent variable is per capita daily caloric availability. Column (1) makes no adjustments to reported values. Column (2) reassigns values above the 99th percentile, conditional on having positive consumption of the food item, to the 99th percentile of consumption conditional on having positive consumption. Topcoding is done by food item before aggregating across the food items.
2. The sample is drawn from the Indian National Sample Survey, Schedule 1 (consumption), rounds 60, 62, and 64 and is limited to households with survey periods with no overlap or full (29 day) overlap between the survey period and Ramadan. These rounds are included because they contain survey dates while earlier rounds do not.
3. All regressions include district-year fixed effects and month of interview fixed effects. In addition, indicator variables for other major religions and their interaction terms with overlap between the survey period and Ramadan as well as overlap between the survey period and Eid (to address the fact that food purchases are lumpy and made in advance) and the interaction between this variable and the religion indicators are included in the regressions but omitted from the table for simplicity.
4. NSS sampling weights are used and are reweighted to weight each round of the survey equally.
5. Results are similar using other methods of topcoding and winsorizing. See Appendix 2A, Table 2A.1 for additional robustness checks.
6. Robust standard errors clustered by district-year are in brackets.
7. *** p<0.01, ** p<0.05, * p<0.1.

2.4 Changes in Agricultural Production

2.4.1 Data and Estimation Strategy

In order to examine the impact of Ramadan fasting on agricultural production this analysis draws on four main sources of data. First, data on agricultural production was obtained from the World Bank India Agriculture and Climate dataset complied by Apurva Sanghi, Kavi Kumar, and James McKinsey. This dataset contains production and price information for 20 crops between 1956 and 1987 in 271 Indian districts covering 85 percent of the land area of India and all of the major agricultural areas with the exception of Kerala and Assam.[9] Second, data on agricultural production cycles was generously provided by Dave Donaldson who compiled it from the 1967 Indian Crop Calendar published by the Indian Directorate of Economics and Statistics (Donaldson 2013). This dataset includes the typical sowing and harvesting periods for 18 of the crops included in the agricultural production data at the district level. Table 2A.2 in Appendix 2A provides additional information about the 18 crops and their cropping cycles. Third, data on the fraction of individuals in each district who are Muslim was gathered from the 1961 Indian census. Finally, rainfall and temperature data were obtained from the University of Delaware monthly rainfall and temperature series.

2.4.2 Overview of Sources of Variation and Estimation Strategy

As demonstrated previously, Ramadan generates a month-long quasi-exogenous shift in caloric intake of roughly 700 calories per day among Muslims. However, in order to assess the impact of this change in calories on economic productivity it is also necessary to measure output as related to "exposure" to fasting that is independent of spatial or temporal confounds. In addition, in order to increase external validity, this variation should occur in an economically important industry in which production is well measured at the same "level" as the variation in the exposure to fasting.

Agricultural production offers a number of advantages in this respect. Although many industries are relatively stable across the year (i.e. manufacturing), agricultural production has significant seasonality, both for given crops and in different locations. Specifically, crops are produced at different times of the year in different locations. For example, rice may be grown in February in one district and in June in another. This variation is significant in India, with an average standard deviation of approximately 1.5 months in the timing of agricultural cycles for any given crop (See Table 2A.2 in Appendix 2A). In addition,

[9] A district in India, similar to a county in the United States, is the administrative unit immediately below the state. The average district in this sample had roughly three million people, of which over 80 percent lived in rural areas in 1961.

different crops are produced at different times in the same location. For example, a district may grow rice in May and wheat in September. A typical district has a standard deviation of approximately 2.5 months and a range of roughly 9 months for the timing of crop cycles within the district (See Table 2A.3 in Appendix 2A). This seasonality provides useful variation in "exposure" to Ramadan fasting both within and between districts. In addition to the variation related to heterogeneous cropping cycles, the fact that Ramadan is a lunar holiday and cycles throughout the calendar year can be exploited to provide temporal variation in exposure to Ramadan fasting for each crop-district combination.

In short, because the same crop is planted following different cycles in different climatic zones, various crops are planted at different times within a district, and the holiday cycles throughout the year, the overlap between the period of low caloric intake and the labor intensive portions of the crop cycle varies across districts, crops, and years. These sources of variation allow a triple-difference approach to identifying the impact of Ramadan at the district (N = 271), crop (N = 18), year (N = 32) level.

2.4.3 Results - Agricultural Production During Ramadan

I begin by estimating Equation (2.2) regressing the production of crop c produced in year t in district d (q_{cdt}), on the fraction of Ramadan covered by each of the labor intensive portions of the agricultural cycle for that crop-district-year (S_{cdt} to indicate overlap with sowing and H_{cdt} to indicate overlap with harvest), as well as district-crop (θ_{cd}), district-year (π_{dt}), and crop-year (α_{ct}) fixed effects and a vector of time varying controls for rainfall and temperature relative to the cropping cycle (X_{cdt}).[10] Details of the calculation of the fraction of Ramadan covered by sowing and harvesting are included in Appendix 2B and example calculations are provided in Figure 2A.1. The extent of the "overlap" between Ramadan and the sowing (harvesting) seasons is measured as a fraction of Ramadan to provide a constant denominator and clearer interpretation of the coefficients of interest.[11] Specifically, β_1 (β_2) multiplied by 100 corresponds to the

[10] See Appendix 2B for a more detailed description of the rainfall and temperature controls.
[11] Results are qualitatively similar, although more difficult to interpret in terms of magnitude, when measuring overlap as a fraction of the season rather than as a fraction of Ramadan. See Appendix 2A, Table 2A.5 for regressions using the fraction of the season rather than the fraction of Ramadan as the measure of overlap.

percentage decline in total production for complete overlap between Ramadan and sowing (harvesting).[12] Table 2A.4 in Appendix 2A displays the distribution of overlap between Ramadan and sowing and harvesting.

$$(2.2) \quad q_{cdt} = \beta_0 + \beta_1 S_{cdt} + \beta_2 H_{cdt} + \theta_{cd} + \pi_{dt} + \alpha_{ct} + \lambda X_{cdt} + \varepsilon_{cdt}$$

As can be seen in Table 2.2, Column (1) total agricultural production in log metric tons declines significantly with increased overlap between Ramadan and sowing or harvesting. Although the estimated coefficients are relatively small, they do correspond to meaningful production declines. Specifically, complete overlap between Ramadan and a sowing (harvest) period would result in an overall decline in production of 1.7 (2.5) percent. However, because only 10 percent of the population is Muslim, these overall declines in output correspond to a decline in productivity of roughly 17 (25) percent per fasting individual when the sowing (harvesting) season fully overlaps the holiday.[13]

In addition to examining the impact of overlap on net production by weight across all crops, Table 2.2 displays estimates of changes in the value of production in Column (2) and the production of rice (the primary staple grain and the crop of greatest total economic value in India) in Column (3) as a function of the fraction of Ramadan covered by sowing and harvest seasons. These estimates are similar although somewhat larger than the percentage declines in production for all crops, indicating that the declines are observed among economically meaningful crops and are not simply driven by marginal ones.

---

[12] This continuous measure of overlap with Ramadan is in contrast to the binary measure used in the analysis of caloric intake. The difference in the choice of right hand side variables is driven primarily by technical details about survey administration in the NSS in which some respondents were surveyed about consumption in the previous month and others about consumption the previous week, which would artificially censor the extent of overlap for these households. In addition, food purchases are likely to be lumpy. Hence, continuous measures of overlap are likely to have significant noise, biasing estimates down. Thus, a discrete measure of overlap is used in the analysis of caloric changes while a continuous measure is used here to improve precision.

[13] The percent of the population that is Muslim is calculated as a population weighted average across the rural portions of the 271 districts included in the agricultural data. Data are drawn from the 1961 Indian Census. It is possible that Muslims are more or less likely than the average rural resident to be involved in agriculture, resulting in biased productivity estimates. However, in the employment/unemployment portion (Schedule 10) of the National Sample Survey, Muslims are marginally less likely to be primarily employed in agriculture than the average rural resident, suggesting that this estimate is likely to be a lower bound.

Although the net declines are relatively small, making it unlikely that farmers would substitute between crops, I confirm that declines are robust to aggregation at the district-year level. These results are presented in Table 2.3. The point estimates for aggregate declines are generally somewhat larger, but statistically indistinguishable from, the declines estimated using the crop level data. These results suggest that substitution between crops is unlikely to be driving these effects and biasing the crop level estimates.

Table 2.2: Effect of Overlap Between Ramadan and Cropping Cycles on Output

	(1)	(2)	(3)
Crop(s)	All	All	Rice only
Dependent Variable	ln(q)	ln(value)	ln(q)
Fraction of Ramadan covered by sowing	-0.017*	-0.028***	-0.058***
	[0.01]	[0.01]	[0.016]
Fraction of Ramadan covered by harvest	-0.025***	-0.044***	-0.041***
	[0.009]	[0.009]	[0.015]
Mean of dependent variable	1.517	1.837	3.91
Observations	103,104	103,088	7,741

Notes:
1. This table tests for changes in agricultural output in each district-crop-year as a function of overlap between Ramadan and the sowing and harvesting season for that crop-district-year. Overlap is defined as the fraction of Ramadan covered by the sowing (harvesting) season such that multiplying coefficients by 100 produces the decline associated with complete overlap between Ramadan and the season. A more detailed description of calculation of the overlap variables is included in Appendix 2B.
2. The dependent variables are: Columns (1) and (3), log production in thousands of tons, and Column (2), log value of production (in 1,000,000 Rs deflated to 1973). Columns (1) and (2) include all 18 crops while Column (3) is limited to rice only. Rice is both the most commonly consumed staple grain and the most economically important crop in India.
3. Columns (1) and (2) include district-crop, district-year, and crop-year fixed effects. Column (3) includes fixed effects for district and year. In addition, all regressions include time varying controls for average rainfall and temperature during the sowing and harvesting seasons, two month leads to each season, and a two month lag following the sowing season are included.
4. The agricultural data are from the India Agriculture and Climate data set. Crop cycles are from Donaldson (2013). Weather data are from the University of Delaware monthly rainfall and temperature series taken for the centroid of each district.
5. Robust standard errors clustered by district-year are in brackets.
6. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.3: Effect of Overlap Between Ramadan and Cropping Cycles on Output, District-Year Aggregation

	(1)	(2)
Dependent Variable	ln(q)	ln(value)
Fraction of Ramadan covered by sowing	-0.060***	-0.035**
	[0.013]	[0.014]
Fraction of Ramadan covered by harvest	-0.039***	-0.066***
	[0.014]	[0.016]
Mean of dependent variable	5.807	5.96
Observations	8,636	8,636

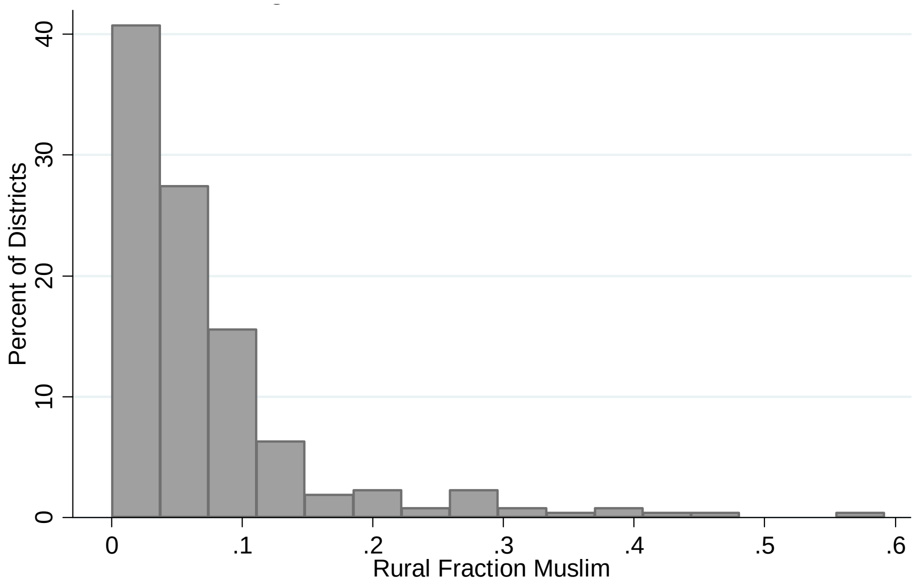
Notes:
1. This table tests for changes in agricultural output in each district-year as a function of overlap between Ramadan and the sowing and harvesting seasons for that district-year. Overlap at the district-year level is defined as weighted sum of the overlap between Ramadan and each of the 18 crops where the weights are the average fraction of production that crop accounts for in all years excluding the current year. A more detailed description of calculation of the overlap variables for each crop is included in Appendix 2B.
2. The dependent variables are log production in thousands of tons in Column (1) and log value of production (in 1,000,000 Rs deflated to 1973) in Column (2).
3. Regressions include district and year fixed effects as well as vector of time-varying controls for monthly rainfall and temperature during both the current and leading agricultural year (to account for long crop cycles).
4. The agricultural data are from the India Agriculture and Climate data set. Crop cycles are from Donaldson (2013). Weather data are from the University of Delaware monthly rainfall and temperature series taken for the centroid of each district.
5. *** p<0.01, ** p<0.05, * p<0.1.

2.4.4 Heterogeneity in Production Declines

Despite the relatively low overall fraction of Muslims in the population, there is substantial heterogeneity in the fraction of Muslims in each district, ranging from less than 0.1 percent Muslim to over 40 percent Muslim (Figures 2.1a and 2.1b and Table 2A.6 provide distributional information on the fraction of the rural population that is Muslim.[14] The sample is restricted to those districts included in the agricultural dataset).

[14] The religious affiliation of individuals in a district is often substantially different between urban and rural areas. Hence, I rely on the fraction of Muslims in rural areas rather than in the overall population of a district.

Figure 2.1a: Geographic Distribution of Muslims in India

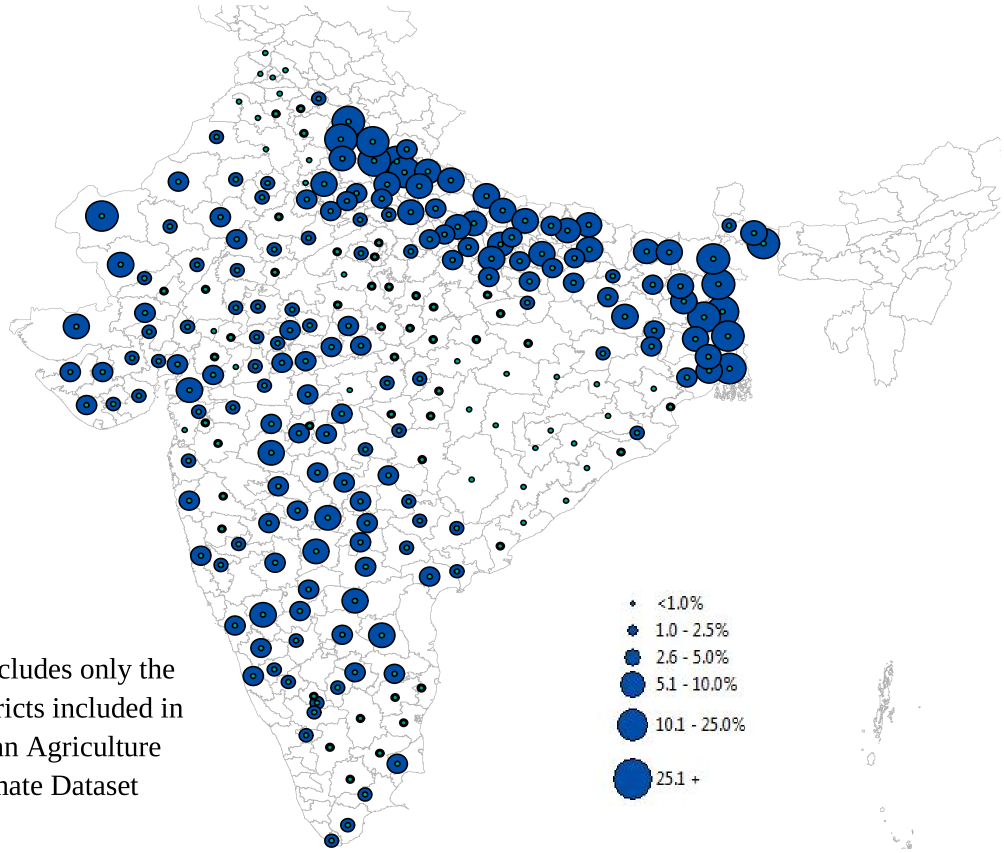


Note: Figure includes only the 270 districts in the Agricultural and Climate dataset

This heterogeneity provides an additional test to confirm that the decline in production is driven by Ramadan rather than other unobserved factors. If the declines are indeed driven by Ramadan, then they should only be observed in areas in which individuals observe the holiday, or in other words, areas in which a reasonably large portion of the population is Muslim. Equation (2.3) tests this prediction by augmenting the previous regression with an interaction terms between the overlap variables and variable(s) for the fraction of Muslims in a district (M_d^n).

$$(2.3) \quad q_{cdt} = \beta_0 + \beta_1 S_{cdt} + \beta_2 H_{cdt} + \beta_n^s S_{cdt} * M_d^n + \beta_n^h H_{cdt} * M_d^n + \theta_{cd} + \pi_{dt} + \alpha_{ct} + \lambda X_{cdt} + \varepsilon_{cdt}$$

Figure 2.1b: Geographic Distribution of Muslims in India



Note: Includes only the 270 districts included in the Indian Agriculture and Climate Dataset

In accordance with this prediction, in below median fraction Muslim districts there is no measurable impact of overlap on production. But, there is a strongly negative and statistically significant impact of overlap in above median fraction Muslim districts (Table 2.4, Columns (1) and (3)). In Columns (2) and (4), I also estimate the net decline in production as a function of the percentage of Muslims in the district using both linear and quadratic interaction terms. These regression results are presented graphically in Figure 2.2.

This heterogeneity provides additional evidence that the observed productivity declines are driven by Ramadan rather than other unobserved causes. In addition, these estimates suggest that these declines may be somewhat larger than those estimated in the non-interacted specification in Table 2.2. Taking a population weighted sum of the average decline at each fraction Muslim and averaging over the four measures (sowing and harvesting for both production by weight and by value), the average decline in productivity per Muslim with full overlap is 39.1 percent. Figure 2.3 plots the average decline in productivity by the fraction of Muslims in the district.

Table 2.4: Heterogeneity in Production Declines by Fraction Muslim

Dependent Variable	(1) ln(q)	(2) ln(q)	(3) ln(value)	(4) ln(value)
Fraction of Ramadan covered by sowing	0.023	0.034**	0.009	0.017
	[0.029]	[0.015]	[0.029]	[0.015]
Fraction of Ramadan covered by sowing*Above 50th percentile Muslim	-0.078**		-0.072**	
	[0.034]		[0.036]	
Fraction of Ramadan covered by sowing*Fraction Muslim		-1.155***		-1.078***
		[0.203]		[0.207]
Fraction of Ramadan covered by sowing*Fraction Muslim squared		2.65***		2.651***
		[0.429]		[0.439]
Fraction of Ramadan covered by harvest	0.002	0.019	-0.012	0.002
	[0.018]	[0.013]	[0.019]	[0.013]
Fraction of Ramadan covered by harvest*Above 50th percentage Muslim	-0.052**		-0.062**	
	[0.025]		[0.026]	
Fraction of Ramadan covered by harvest*Fraction Muslim		-1.082***		-1.121***
		[0.192]		[0.198]
Fraction of Ramadan covered by harvest*Fraction Muslim squared		2.889***		2.893***
		[0.432]		[0.46]
Mean of dependent variable	1.517	1.517	1.837	1.837
Observations	103,104	103,104	103,088	103,088

Notes:
1. This table tests for changes in agricultural production in each district-crop-year as a function of the fraction of Ramadan covered by the sowing (harvesting) season and the interaction of that variable with various indicators for the fraction of Muslims in the district. Columns (1) and (3) interact the overlap variable with an indicator for the district having an above median fraction of Muslims. Columns (2) and (4) interact the overlap variable with a continuous variable for the fraction Muslim as well as the square of this variable. A more detailed description of calculation of the overlap variables is included in Appendix 2B.
2. Quantities (Columns (1) and (2)) are in log thousands of tons and values (Columns (3) and (4)) are in log 1,000,000 Rs deflated to 1973.
3. The net decline in production and in productivity per Muslim at each fraction Muslim based in the estimates in Columns (2) and (4) are plotted in Figures 2.2 and 2.3, respectively.
4. All fraction Muslim variables are based on the fraction of Muslims in rural portions of each district in 1961. The median district is roughly 5 percent Muslim. Further distributional information is provided in Figures 2.1a and 2.1b and Appendix 2A, Table 2A.6.
5. All regressions include district-crop, district-year, and crop-year fixed effects. In addition, time varying controls for average rainfall and temperature during the sowing and harvesting seasons, two month leads to each season, and a two month lag following the sowing season are included.
6. Robust standard errors clustered by district-year are in brackets.
7. *** p<0.01, ** p<0.05, * p<0.1.

Figure 2.2: Production Declines by Fraction Muslim

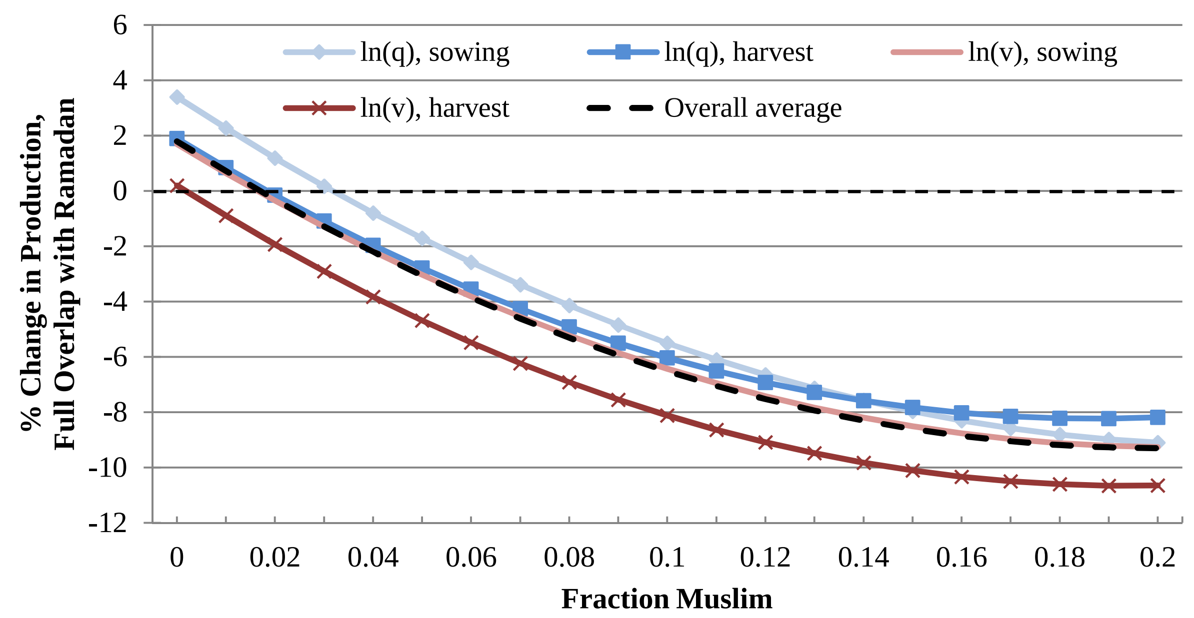
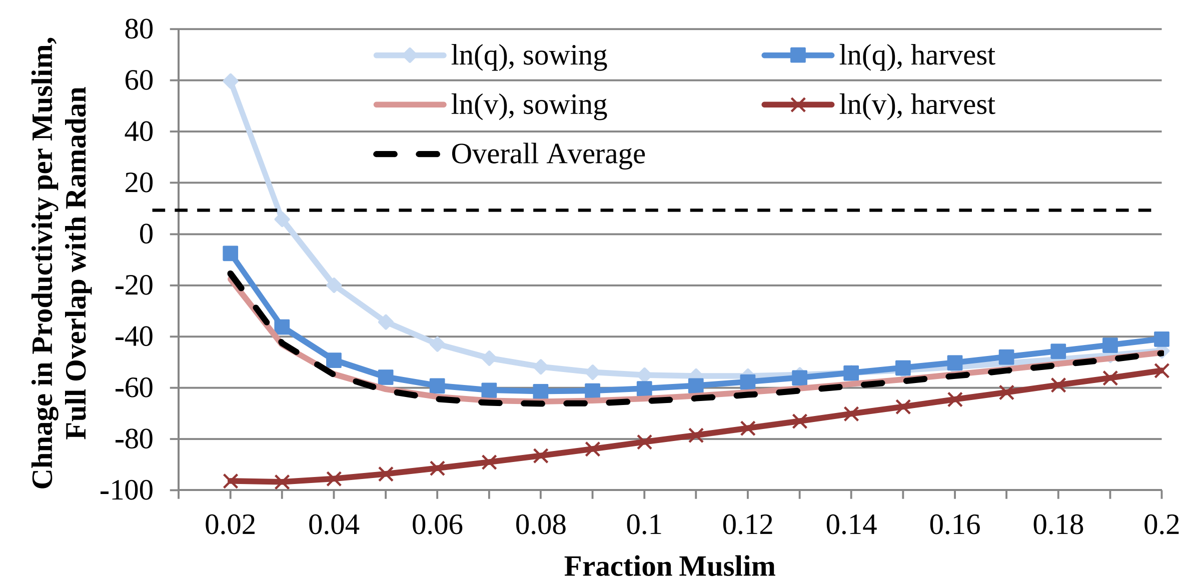


Figure 2.3: Change in Productivity per Muslim Individual



Although this decline is substantial, the magnitude is consistent with a back of the envelope calibration of the "expected decline" given basal metabolic rates (BMR), the estimated declines in caloric intake, and the energy requirements of farm labor (see Appendix 2B for details of this calculation). This calibration of "available labor energy" predicts productivity declines of roughly 50 percent as a result of the reduced in caloric intake if no energy were mobilized from fat stores. This rough calibration does not provide evidence for or against alternative explanations for the decline. However, it does suggest the observed magnitudes are consistent with expectations for productivity changes driven by fasting.

2.4.5 Summary of Results - Production Declines

These results demonstrate that: 1) overlap between Ramadan and sowing/harvesting has a substantial negative impact on total agricultural production both in terms of quantity and value, 2) consistent with Ramadan driving the declines, the impact is larger for districts with a greater percentage of Muslims, and 3) the magnitude of the productivity decline is consistent with the estimated change in caloric intake. However, it is possible that features of the holiday other than the caloric decline associated with fasting are also influencing production. As described previously, there are three main behavioral changes beyond reduced caloric intake which could potentially decrease production: time spent on religious or social activities leading to reduced labor supply, changing sleep patterns or reduced time sleeping, and dehydration. Section 2.5 below examines each of these possibilities, and concludes that they are unlikely to be key channels for the documented productivity declines.

2.5 Evidence Regarding the Forces Driving Production Declines

This section examines whether caloric intake or other behavioral changes are driving the observed declines in productivity during Ramadan. The empirical tests are broken into two portions. First, an omnibus test provides evidence which supports the role of caloric decline in reduced production during Ramadan while simultaneously providing evidence that the declines are unlikely to be driven by the three other notable behavioral changes during the holiday. Following this analysis, I also individually examine whether the declines could be driven by the impact of religious or social obligations on labor supply, by sleep changes, or by dehydration via both relevant literature and additional direct empirical tests.[15]

2.5.1 Omnibus Test - Persistence of Production Declines Following Ramadan

A distinctive feature of low caloric intake is that recovery from this state takes time, often requiring weeks or even months to fully regain physical performance in extreme cases of deprivation (e.g., Keys et al. 1950). This lag between increased caloric intake and improved physical performance suggests that the reduced caloric intake during the holiday would be expected to have an impact beyond Ramadan itself.

In contrast, recovery from sleep deficits and dehydration is rapid. As discussed in greater detail below, sleep deficits primarily influence cognitive rather than physical function.[16] However, a number of researchers have documented that a single night of greater than eight hours of sleep reverses cognitive deficits from extended (24-48 hours) periods of total sleep deprivation (Drummond et al. 2006; Kendall et al. 2006; Brendel et al. 1990). Similarly, Belenky et al. (2003) cannot distinguish measures of attention and reaction time between groups which had either five, seven, or nine hours of sleep per night for one week after a three day recovery period with eight hours of sleep per night. Dinges et al. (1997) also demonstrate that two nights of ten hours of sleep is sufficient to overcome deficits accumulated over seven nights with five or fewer hours of sleep. Similarly, for recovery from dehydration, Sawka et al. (2007) survey a variety of sources and note that normal body water and performance can typically be restored within 8 to 24 hours even after fairly extensive dehydration.

---

[15] One additional concern beyond these behavioral changes is that the timing of caloric intake throughout the day may also influence productivity. For clear reasons, it is not possible to disentangle the effect of timing of intake and of levels of intake in this analysis. However, the RCT presented in Chapter 1 does not have this concern and provides similar results. In addition, given the fixed costs of preparing food, many individuals who consume very few calories will do so by consuming fewer meals, generating a similar temporal pattern of consumption to that observed during Ramadan.

[16] To the best of my knowledge, there are no studies examining physical recovery after extended sleep deprivation. This is likely driven by the fact that except in cases of fairly extreme sleep deprivation, studies do not often find physical declines associated with sleep deprivation. Hence, there is no recovery period to measure.

These rapid recovery periods from sleep deprivation and dehydration suggest that if reduced production during Ramadan were due to sleep deprivation, dehydration, or changes in labor supply related to religious obligations that these deficits should disappear within a few days after the conclusion of the holiday. In contrast, if the reduced production is the result of lower caloric intake during the month, there should continue to be residual effects which reduce production following the holiday. Hence, examining the impact of overlap between the sowing and harvesting seasons and the weeks following Eid (a short holiday immediately following Ramadan) serves as a useful test to distinguish between nutritional deficits and other behavioral changes during the holiday as likely causes of the reduced production.

This test is conducted using an empirical approach similar to the previous analyses, in which the fraction of Ramadan covered by the sowing and harvesting seasons is replaced with the fraction of the weeks following the holiday covered by the sowing and harvesting seasons. However, because the Eid holiday could potentially confound the results and there is a short recovery period for dehydration and sleep deprivation, I calculate overlap with the first and second week following Eid rather than following Ramadan.[17] The estimating equation, Equation (2.4), is displayed below. P_{cdt}^{XN} denotes the overlap between season X (where X = {S for sowing, H for harvesting}) and week N (where N = {1,2}). All other variables are as defined previously.

(2.4) $$q_{cdt} = \beta_0 + \beta_{XN} P_{cdt}^{XN} + \theta_{cd} + \pi_{dt} + \alpha_{ct} + \lambda X_{cdt} + \varepsilon_{cdt}$$

Although the standard errors are larger, the results of overlap for the first week post-Eid are quite similar to the results during the holiday itself for both sowing and harvest (Table 2.5). This effect is relatively robust to longer (4 day) specifications for Eid (Columns (2) and (4)). While the point estimates remain similar for the second week post-Eid for sowing, the effects begin to attenuate by the second week following Eid for harvest, as might be expected given the relatively short period of deprivation. In short,

[17] The holiday is lunar so the exact day of Eid will depend on the sighting of the crescent moon, which varies across locations. I build in a one day buffer to account for the uncertainty in sighting the moon. In addition, while the official holiday is only one day, some individuals continue to celebrate for up to three days. Hence, to span the possible durations of the holiday, I examine periods following both a two-day and a four-day lag for the Eid holiday.

these results are consistent with recovery from a short period of reduced caloric intake but inconsistent with dehydration, sleep deficits, or time spent on religious activities driving the reduced production.

Table 2.5: Persistence of Declines in Output Following Ramadan (and Eid)

	(1)	(2)	(3)	(4)
Eid length	Short (2 day) Eid	Long (4 day) Eid	Short (2 day) Eid	Long (4 day) Eid
Dependent Variable	ln(q)	ln(q)	ln(v)	ln(v)
Fraction 1st week post Eid covered by sowing	-0.016	-0.014	-0.022	-0.023
	[0.017]	[0.017]	[0.017]	[0.017]
Fraction 2nd week post Eid covered by sowing	-0.016	-0.018	-0.026	-0.025
	[0.017]	[0.017]	[0.018]	[0.017]
Fraction 1st week post Eid covered by harvest	-0.023*	-0.004	-0.043***	-0.027**
	[0.014]	[0.013]	[0.014]	[0.014]
Fraction 2nd week post Eid covered by harvest	0.011	-0.006	0.024*	0.011
	[0.014]	[0.013]	[0.014]	[0.013]
Mean of dependent variable	1.517	1.517	1.837	1.837
Observations	103,104	103,104	103,088	103,088

Notes:
1. This table provides a test of whether declines in production during Ramadan are due to low caloric intake, which causes longer lasting decreases in productivity, or other potential confounds including dehydration or sleep deprivation, from which recovery is rapid, or time spent on religious activities during the holiday. To accomplish this, this table examines changes in agricultural production (in log thousands of tons) in each district-crop-year as a function of the overlap between the first and second week following Eid (the holiday following Ramadan) and the sowing (harvesting) season for that crop-district-year.
2. Eid is defined as a two day period in Columns (1) and (3) a four day period in Columns (2) and (4). These formulations are used because the holiday is lunar so the exact day of Eid will depend on the sighting of the crescent moon, which varies across locations. I build in a one day buffer to account for the uncertainty in sighting the moon. In addition, while the official holiday is only one day, some individuals continue to celebrate for up to three days. Hence, to span the possible durations of the holiday, I examine periods following both a two day and a four day lag for the Eid holiday.
3. All regressions include district-crop, district-year, and crop-year fixed effects. In addition, time varying controls for average rainfall and temperature during the sowing and harvesting seasons, two month leads to each season, and a two month lag following the sowing season are included.
4. Robust standard errors clustered by district-year are in brackets.
5. *** p<0.01, ** p<0.05, * p<0.1.

2.5.2 Time Spent on Religious and Social Activities

Literature - Time Spent on Religious and Social Activities

Time spent on religious and social activities during the Ramadan holiday is likely to vary substantially across both individuals and locations, and little information specific to India is available given the relatively low concentration of Muslims in the country. However, the literature does provide general information about the types of behavior changes that Muslim individuals are likely to engage in during the holiday. The primary religious behavioral changes expected during Ramadan are an increase the number of prayers offered, abstinence from smoking, sexual relations, and swearing during the day, and an increase in charity. While a general guideline to increase prayers and reading and recitation of the Koran exists, no specific rules are provided regarding the amount or timing of the prayer that should be completed beyond the standard prayers observed throughout the year five times daily. Similarly, there are no specific guidelines for the amount of charity to be provided during this time.

In addition to these changes in religious practice, Muslims often spend time in social gatherings in the evenings during Ramadan. The most common set of activities following sunset is to break the fast with a small snack, to complete the fourth set of daily prayers, and then to gather with friends or family for a large evening meal (Blackwell 2009; Ahmad et al. 2012). So, although there is an increase in religious and social activity during this time, much of the increase occurs in the evenings when individuals would not typically be working. In addition, because the prescribed changes in behavior are general and flexible, these behavior changes are less likely to interfere with work requirements than observance of many other holidays. Finally, although some heavily Muslim counties see shifts in working hours during Ramadan, the low fraction of Muslims in India make these types of labor market changes unlikely to occur in this context.

Empirical Evidence - No Change in Labor Supply or Earnings During Ramadan

I also examine labor supply changes among Muslims in India during Ramadan directly using two sources of data; the second-generation ICRISAT village level studies survey and the National Sample Survey (NSS).

Data

ICRISAT is a panel dataset containing monthly measures of employment and earnings. The sample is drawn from six villages chosen to be representative of the major agro-climatic zones in the semi-arid tropics of

India.[18] Households were sampled across four categories of landholdings (landless, small farmers, medium farmers, and large farmers).

In addition to the consumer expenditure survey described previously, the National Sample Survey Organization fields a large, nationally representative, repeated cross-sectional survey about employment. The employment/unemployment rounds (Schedule 10) provide information on respondents' labor supply and wages during the seven days preceding the interview as well as their religious affiliation.[19] Households are sampled on a rolling basis such that the sample is temporally spaced within each district.

Empirical Strategy and Results

In the ICRISAT data, labor supply (or earnings) of individual i, in survey round s, and year-month t are regressed on the number of days of overlap between survey period and Ramadan (R_{ist}), an interaction between that variable and an indicator for whether the individual is Muslim (M_i), a variable for the number of days between interviews (D_{ist}), and fixed effects for individuals (θ_i) and the year-month in which the survey occurred (Equation 2.5).[20] "Labor days" is defined as the number of days of labor, including both paid and unpaid labor but excluding domestic work, in the past month. The wages variable is calculated as the sum of cash and in-kind wages during the month. The number of hours worked is only reported for paid labor. If the participant reports more than one paid job during the survey period, the average hours worked per day is calculated as a weighted average across all jobs reported.

(2.5) $$l_{ist} = \beta_0 + \beta_1 R_{ist} + \beta_2 R_{ist} * M_i + \beta_3 D_{ist} + \theta_i + \omega_t + \varepsilon_{ist}$$

[18] The villages are located in Andhra Pradesh and Maharashtra. These states are major agricultural producers, and are included in the agricultural production data used in the primary analysis presented above.
[19] Although it would be ideal to utilize data overlapping the time period covered by the agricultural production data, only rounds 60 (conducted in 2004) onward contain survey dates, permitting calculation of overlap between the survey period and Ramadan. Hence, I rely on rounds 60, 61, 62, 64, and 66.
[20] In order to maximize the sample size given the smaller number of individuals surveyed in the ICRISAT data, the number of days of overlap is used rather than restricting the sample to complete or no overlap and using a binary indicator of overlap, as was done in the analysis of changes in caloric intake.

A similar specification is used in the NSS data. However, I adjust for the fact that the data consists of multiple cross sections rather than a panel. As can be seen in Equation (2.6), labor supply (or earnings) of individual I, in district d, at time t (l_{idt}), is regressed on the days of overlap between the survey period and Ramadan (R_{idt}), whether the individual is Muslim (M_i), the interaction of these two variables, district-year fixed effects (θ_{dt}), and month fixed effects (ω_m). "Labor days" is the combination of days engaged in labor for wages and own-labor (e.g. working on one's own farm) excluding domestic work. Earnings are the total (cash plus in-kind) earnings received during the survey period. Earnings are considered in both levels and logs to account for unemployment and unpaid labor.

(2.6) $l_{idt} = \beta_0 + \beta_1 R_{idt} + \beta_2 M_i + \beta_3 R_{idt} * M_i + \theta_{dt} + \omega_m + \varepsilon_{idt}$

Tables 2.6a and 2.6b provides the results of these regressions. In both datasets, Muslims' labor supply and earnings during the holiday are fairly precisely estimated (a change in labor supply of less than one percent per day of overlap can be detected). Earnings remain unchanged for Muslims during Ramadan. However, interestingly, labor supply appears to either remain constant or increase slightly. The increased labor supply could be caused by an attempt to compensate for reduced productivity and/or to accrue funds for charity, which Muslims are expected to provide during Ramadan.[21]

These results suggest that changes in labor supply driven by religious or social obligations are unlikely to cause the observed declines in production when Ramadan overlaps with important parts of the agricultural cycle and may actually work against such declines.[22]

[21] Although employers are likely to be less willing to hire a less productive employee, many workers do have long-term relationships with employees which may mitigate this effect. Further, as is discussed in greater detail below, many individuals appear to be unaware of the relationship between caloric intake and productivity.

[22] As an additional robustness check, I estimate the same equations in samples limited to the high labor demand seasons and in alternative samples (e.g. rural casual laborers) (See Appendix 2A, Tables 2A.7a and 2A.7b). The smaller sample sizes result in less precise estimates; however, the patterns are very similar and neither wages nor labor supply shows any significant change for Muslims during Ramadan.

Table 2.6a: Labor Supply and Earnings as a Test of Religious Obligations Driving Production Declines

	(1)	(2)	(3)
Sample	ICRISAT Agricultural Laborers		
Dependent variable	Labor days	Average work hours	ln(Total earnings)
Days overlap between survey period and Ramadan	-0.039	-0.002	-0.004***
	[0.035]	[0.001]	[0.001]
Muslim*Days overlap between survey period and Ramadan	0.155*	-0.003	0.002
	[0.089]	[0.011]	[0.006]
Mean of dependent variable	21.752	21.874	17.921
Number of individuals in sample	1,146	910	914
Observations	31,432	19,417	19,287
R-squared	0.699	0.775	0.659

Notes:
1. This table provides the first test of whether time spent on religious or social activities causes reduced production during Ramadan by reducing the labor supply of Muslim individuals.
2. Data is drawn from the second generation ICRISAT village level studies survey. The sample includes only individuals who indicate that agricultural work is their primary occupation.
3. ICRISAT surveys participants approximately once per month and elicits information on labor supply, hours, and wages of the respondent during the month preceding the survey date. The number of days of overlap between Ramadan and the survey period is calculated as the number of days of Ramadan falling within the 30 days preceding the survey date. Labor supply is defined as the number of days of labor including both paid and unpaid labor but excluding domestic work in the past month. Wages are calculated as the sum of cash and in-kind wages during the month. Average hours worked is only reported for paid labor. If the participant reports more than one paid job, average hours worked is calculated as a weighted average across jobs. Religion is captured via the caste variable, which contains a category for Muslim individuals, because it is not directly reported.
4. Regressions include individual fixed effects, a control for the number of days between surveys, and year-month of interview fixed effects. Robust standard errors clustered by individual are in brackets.
5. Results are similar in samples restricted to high labor demand periods, defined as responses for which the interview period had at least 15 days of overlap with a "high labor demand" period defined as the sowing or harvesting seasons for the crop with the greatest acreage by state, and for the full (including non-agricultural worker) sample. Results of these regressions are included in Appendix 2A, Table 2A.7a.
6. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 2.6b: Labor Supply and Earnings as a Test of Religious Obligations Driving Production Declines

	(1)	(2)	(3)
Sample	NSS Agricultural Laborers		
Dependent variable	Labor days	Total Earnings	ln(Total earnings)
Days overlap between survey period and Ramadan	-0.004	-0.689	0.039***
	[0.047]	[2.798]	[0.015]
Muslim	-0.025	-1.775	0.013
	[0.051]	[4.551]	[0.025]
Muslim*Days overlap between survey period and Ramadan	0.077***	-1.775	0.01
	[0.024]	[1.730]	[0.017]
Mean of dependent variable	5.852	103.391	5.521
Observations	121,400	121,408	35,721
R-squared	0.127	0.062	0.314

Notes:
1. This table provides the second test of whether time spent on religious or social activities causes reduced production during Ramadan by reducing the labor supply of Muslim individuals. Given the somewhat coarse nature of the labor supply variable in the NSS, the table also examines wages.
2. The regressions utilize data from the Indian National Sample Survey (NSS), Schedule 10 (Employment), rounds 60, 61, 62, 64, and 66. These rounds are selected because they contain survey dates while earlier rounds do not. The sample includes individuals for whom agricultural work is their primary or secondary occupation.
3. The NSS Schedule 10 provides data on labor supply (to the half day) and wages of the respondent during the week preceding the survey date. Labor supply is calculated as the number of days of labor excluding domestic work in the past week.
4. Regressions include district-year fixed effects and month fixed effects. Robust standard errors clustered by district-year are in brackets.
5. Results are similar in samples restricting the survey periods to high labor demand periods, defined by the sowing or harvesting seasons for the crop with the greatest acreage by state. Results are also similar in samples of rural casual laborers. See Appendix 2A, Table 2A.7b for these additional results.
6. *** p<0.01, ** p<0.05, * p<0.1.

2.5.3 Sleep

Literature -The Impact of Sleep on Cognitive and Physical Function

A limited literature on the impact of Ramadan on sleep habits does exist. However, it focuses on countries such as Saudi Arabia which are nearly exclusively Muslim (Bahamman 2006 provides a review). Because the holiday is celebrated so widely in these countries, the habits of the country as a whole often shift during the month of Ramadan. For example, stores often both open and close later than usual. However, because the median district in this study is only seven percent Muslim, these equilibrium shifts are unlikely to occur in India. In addition, within this literature, changes in sleep patterns appear to vary widely across countries (Bahamman 2006). Hence, the direct literature on changes in sleep patterns during Ramadan is unlikely to aid in understanding the likely impact of possible changes in sleep patterns on agricultural production in India. Instead, I focus on the more extensive and relevant literature surrounding the physiological relationship between chronic partial sleep deprivation, defined as 2 to 7 hours of sleep per night for periods of days to months, and physical and cognitive performance to understand the likely effect of changes in sleep patterns during the holiday.[23]

The effects of sleep deprivation on performance appear to be heavily cognitive and occur primarily when sleep drops below six to seven hours per night (Goel et al. 2009; Durmer and Dinges 2005). When sleep drops below this level, the largest and most consistent effects of sleep deprivation center on changes in mood with increased sleepiness and negative affect (Alhola and Polo-Kantola 2007; Durmer and Dinges 2005). In addition, chronic sleep deprivation of less than 6 hours per night is likely to lead to reduced vigilance, poorer memory consolidation, and slowed reaction time and learning (Blagrove et al. 1995; Rogers et al. 2003; Ferrara and De Gennaro 2001; Van Dongen et al. 2003). These effects on mood and cognitive function typically tend to have a relatively linear relationship with the cumulative deprivation in

[23] The majority of the sleep deprivation literature focuses on total sleep deprivation (i.e. no sleep at all for at least 24 to 72 hours). However, I focus more heavily on the subset of the literature on "partial sleep deprivation" because this pattern of sleep deprivation is more likely to reflect changes that Muslims experience during Ramadan. Results from both literatures show similar patterns, with more extreme outcomes under total sleep deprivation.

the early stages and then often level off over a longer time horizon (multiple weeks) as individuals acclimatize to the new regimen (Ferrara and De Gennaro 2001).

Despite cognitive declines associated with less than six hours of sleep per night, reduced sleep has little if any impact on physical performance until the deprivation becomes relatively extreme (e.g. Sinnerton and Reilly 1992; Guezennec et al. 1994). As VanHelder and Radomski note, “Sleep deprivation of 30 to 72 hours does not affect cardiovascular and respiratory response to exercise of varying intensity, or the aerobic and anaerobic performance capability of individuals” (1989). Some studies do show declines in time to exhaustion after extended periods of sleep deprivation (Reilly and Piercy 1994; VanHelder and Radomski 1989; Vardar et al. 2007). However, results in this domain are mixed and the conditions of deprivation much greater than those likely to be experienced during Ramadan.

In short, the literature relating sleep to cognitive and physical outcomes suggests that changes in sleep patterns during Ramadan are unlikely to be a significant driver of changes in production in this context. Physical capabilities are quite resistant to decreased sleep, even when the physical requirements are substantial as in military combat training. And while extended sleep deprivation is clearly associated with declines in specific areas of cognitive function, the sleep deprivation in these studies tends to be both more extreme than is likely to be experienced during Ramadan and to be limited to specific domains of cognitive function (e.g. memory, reaction time) that are unlikely to be strongly related to agricultural production.

Empirical Evidence – No Change in Hours Worked

In addition to the evidence from the weeks following Eid suggesting that sleep deprivation does not drive productivity declines during Ramadan, it is possible to examine whether changes to sleep patterns crowd out labor supply through naps or returning to sleep after rising early in the morning for prayers. This test is conducted with the ICRISAT data described previously. Respondents who work in paid agricultural labor report the number of hours worked per day for each job. If the participant reports more than one paid job, average hours worked per day is calculated as a weighted average across jobs. As shown in Table 2.6a, average hours worked by Muslims during Ramadan do not change and the lack of change is precisely estimated. While this finding does not rule out changes in the quality of work, it does suggest that naps or other changes to sleep patterns which could influence labor supply on the intensive margin are unlikely to be a significant factor in production declines.

2.5.4 Dehydration

Literature - The Impact of Dehydration on Cognitive and Physical Function

Dehydration typically begins to negatively impact measures of aerobic performance once two percent of body weight has been lost.[24] These declines become more consistent around three to four percent loss in body weight; however, even at these levels results are mixed (Sawka et al. 2007; Casa et al. 2005). There is generally no detectable impact of dehydration on strength and anaerobic performance until at least five percent of body weight is lost from water loss (Sawka et al. 2007; Casa 1999a; Greiwe et al. 1998). There is also more limited evidence that dehydration beyond two percent of body weight may cause declines in cognitive function in areas such as short term memory consolidation (Grandjean and Grandjean 2007).

Given these findings, it is necessary to estimate the expected body weight losses due to dehydration for fasting farmers in India to determine whether dehydration is likely to impact agricultural production during Ramadan. While, to the best of my knowledge, no direct evidence on body water loss rates among farmers in India exists, it is possible to benchmark expected losses relative to sports activities in which sweat rates have been measured based on caloric requirements and environmental factors. This calibration, outlined in Appendix 2B, finds that expected water losses by the end of the day are around the two percent threshold at which aerobic effects become detectable. Given that these losses are cumulative throughout

[24] The technical definitions of aerobic and anaerobic exercise rely on the types of muscular contractions and the energy generating process used. But, a simple delineation based on whether the activity requires breathing hard (i.e. jogging and cycling at a moderate pace are aerobic exercises) or not (e.g. weight lifting and sprinting very short distances are anaerobic exercises) serves as a useful heuristic for differentiating between the types of exercise.

the day, even if this threshold is reached by the end of the day, the majority of the labor supply is likely to be unaffected and the overall impact is likely to be correspondingly small.

Empirical Evidence – No Decrease in Production with Higher Perspiration Rates

I also examine the impact of dehydration empirically via an additional test utilizing data on evaporative potential (PET), a measure of the propensity of water to evaporate into the atmosphere.[25] If dehydration were a significant factor in production, then as evaporative potential increased, body water loss would also increase, and production would decline. However, evaporative potential is also likely to influence agricultural production directly. Hence, this test augments the previous specification not only with a main effect for the average PET over the sowing and harvesting seasons, but also with interactions between PET and overlap between the seasons and Ramadan (Equation 2.7).

$$(2.7) \quad q_{cdt} = \beta_0 + \beta_1 S_{cdt} + \beta_2 PET^S_{cdt} + \beta_3 S_{cdt} * PET^S_{cdt} + \beta_4 H_{cdt} + \beta_5 PET^H_{cdt} + \beta_6 H_{cdt} * PET^H_{cdt} + \theta_{cd} + \pi_{dt} + \alpha_{ct} + \lambda X_{cdt} + \varepsilon_{cdt}$$

The results of these regressions provide further evidence that dehydration is not driving production declines (Table 2.7). In contrast to what would be expected if dehydration were reducing output, the interaction between mean evaporative potential and sowing (harvesting) overlap are positive. Further, the addition of these terms does not eliminate the negative effect of overlap between sowing/harvesting and Ramadan observed in the earlier regressions as would be expected if dehydration were the key factor in production declines. Rather, these terms remain negative and significant.

[25] The measure used is technically referred to as "Potential evapotranspiration" (PET). PET reflects the potential for evaporation from the surface into the atmosphere and is calculated based on three measures of temperature (min, max, mean), vapor pressure, and cloud cover. The calculation to generate the PET from these variables is done by the British Atmospheric Data Center, the source of the data, following a method recommended by the FAO. More information about this measure is available at the BADC website (http://badc.nerc.ac.uk). Because PET data were not available for all districts, missing values were imputed when possible. The imputation procedure is described in Appendix 2B.

Table 2.7: Test of Dehydration Driving Production Declines

	(1)	(2)
Dependent variable	ln(q)	ln(v)
Fraction of Ramadan covered by sowing	-0.257***	-0.216**
	[0.085]	[0.087]
Average evaporative potential during sowing	0.001	0.003
	[0.016]	[0.017]
Fraction of Ramadan covered by sowing*Average evaporative potential during sowing	0.016***	0.013**
	[0.006]	[0.006]
Fraction of Ramadan covered by harvest	-0.242***	-0.203***
	[0.059]	[0.061]
Average evaporative potential during harvest	-0.006	0.002
	[0.013]	[0.013]
Fraction of Ramadan covered by harvest*Average evaporative potential during harvest	0.012***	0.009**
	[0.004]	[0.004]
Mean of dependent variable	1.474	1.786
Observations	58,443	58,443

Notes:
1. This table provides a test of whether declines in production during Ramadan are due dehydration. To accomplish this, this table examines changes in agricultural production (in log thousands of tons) in each district-crop-year as a function of the fraction of Ramadan covered by the sowing (harvesting) season for that crop-district-year fully interacted with a measure of weather features which impact dehydration (evaporative potential).
2. Evaporative potential (technically referred to as "Potential evapotranspiration" or PET) is a measure of the propensity of water to evaporate into the atmosphere. It is calculated based on three measures of temperature (min, max, mean), vapor pressure, and cloud cover following a method recommended by the Food and Agricultural Organization (FAO) of the United Nations (UN). Because PET measures are not available for all districts in all years, PET is imputed for missing observations in states with at least one observation. The imputation procedure is described in Appendix 2B. If dehydration plays an important role in production declines the interaction between mean evaporative potential and overlap is expected to be significantly negative because greater evaporation should increase dehydration and decrease production.
3. The regression includes district-crop, district-year, and crop-year fixed effects. In addition, time varying controls for average rainfall and temperature during the sowing and harvesting seasons, two month leads to each season, and a two month lag following the sowing season are included.
4. Robust standard errors clustered by district-year are in brackets.
5. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

2.6 Returns to Caloric Intake in Agricultural Production

As in the randomized trial presented in Chapter 1, this analysis also suggests that changes in caloric intake cause substantial and widespread changes in productivity. Hence, I next provide the corresponding calculations of the return on investment for higher caloric intake in agricultural production.

Given the estimated changes in caloric consumption in the NSS and in agricultural production for Muslims during Ramadan, it is possible to calculate the estimated returns on investment for caloric intake in the analysis of Ramadan fasting as well. Again, the exact returns depend on a variety of assumptions regarding the foods consumed and the consistency of the returns to higher caloric intake. Yet, all of the estimated returns are positive and even with relatively conservative assumptions, the one month return on investment to greater caloric intake appears likely to be over 200 percent. Details of these calculations are provided in Table 2.8. However, one sample calculation based on data from the Indian Agriculture and Climate data used in the analysis is described below.

One kilo of rice contains approximately 3300 calories (Gopalan et al. 1989; NSS calorie conversion database). The average cost of a kilo of rice in 1971 (the median year in the analysis) was Rs. 1.28 per kilo. Hence, an increase of 700 calories (Table 2.1, Column (2)) would require 210g of rice at a total cost of Rs. 0.27 per day. The average daily agricultural wage in 1971 was Rs. 4.45. Using the estimates of a 20 percent change in productivity per fasting individual during harvest (Table 2.2, Column (1), averaged over sowing and harvesting and scaled by 0.10, the overall fraction of Muslims in the districts in the analysis), the associated increase in earnings would be Rs. 0.89 per day, implying a return over 228 percent in a single month. This high ROI is relatively robust to purchasing foods which are much more expensive per calorie or to high consumption throughout the year even if some periods have much lower returns to caloric intake.

Table 2.8: Return on Investment to Caloric Intake in Agricultural Production

Year	Grain	Cost for 700 calories	Daily wage	Productivity gain	One Month ROI	Minimum ROI
1956	bajra	0.18	4.09	0.2	348	47
1971	bajra	0.17	4.45	0.2	418	70
1986	bajra	0.14	6.03	0.2	761	183
1956	bajra	0.18	4.09	0.4	797	195
1971	bajra	0.17	4.45	0.4	936	241
1986	bajra	0.14	6.03	0.4	1623	466
1956	rice	0.25	4.09	0.2	232	9
1971	rice	0.27	4.45	0.2	228	8
1986	rice	0.18	6.03	0.2	577	123
1956	rice	0.25	4.09	0.4	565	119
1971	rice	0.27	4.45	0.4	556	116
1986	rice	0.18	6.03	0.4	1254	345

Notes:
1) This table calculates the expected return on investment to caloric intake based on the triple-difference estimates of changes in agricultural output.
2) Prices are obtained from the Indian Agricultural and Climate data and are deflated to 1973 Rupees.
3) Estimates of the caloric content of rice and bajra range from 3000 to 3600 calories per kilo depending on the specific variety of the grain and the source used (Gopalan et al 1989; NSS calorie conversion database). An average value of 3300 calories per kilo of grain is used in all calculations.
4) Rice is the commonly consumed grain and also the most expensive. Bajra (pearl millet) is a lower priced grain with similar nutritive properties commonly consumed by lower income individuals.
5) The lower productivity gain, 20 percent, is the expected change in productivity per Muslim individual based on overall changes in log output averaged across the point estimates for sowing and harvesting seasons in Table 2.2. The upper bound productivity gain is based on the specification examining heterogeneity in Table 2.4. Calculations for these changes in productivity are described in the text of the paper.
6) The Minimum ROI is calculated assuming that the individual consumes an additional 700 calories of grain per day throughout the year, but only experiences gains in productivity during the sowing and harvesting seasons. Sowing and harvesting seasons typically last approximately two months each, resulting in a total of 120 days. The "One Month ROI" is calculated assuming that the individual consumes an additional 700 calories per day and is more productive only during the sowing and harvesting seasons (or correspondingly, that he consumes more and earns more throughout the year). No discounting is used given the relatively short time horizons.

2.7 Conclusion

This chapter examines the impact of low caloric intake on agricultural production, a sector which employs over half of the labor force of India. Variation in caloric intake is generated via Ramadan, a month-long Muslim holiday during which individuals in rural agricultural household decrease their caloric intake by roughly 700 calories per day. Using a differences-in-differences approach, I find declines in total agricultural production between 2 and 4 percent, both in weight and in value, when Ramadan and labor intensive parts of the agricultural cycle fully coincide. The extent of the decline shows significant heterogeneity, with larger declines in areas with a higher fraction of Muslims in the population.

Because the proportion of Muslims in the population in India is relatively low overall, these declines in production imply substantially larger declines in productivity per Muslim individual. Specifically, because Muslims account for less than 10 percent of the population of India, these estimates suggest a decline in productivity of approximately 20 to 40 percent per fasting individual.[26]

Although it is also possible that factors other than changes in caloric intake influence production during Ramadan, the evidence from the literature as well as direct empirical tests suggests that three of the leading potential causes of decline -- reduced labor supply due to religious or social obligations, sleep deprivation, and dehydration -- do not appear to play a significant role in productivity in this context. In addition, the omnibus test examining the persistence of declines is consistent with the effect being driven by declines in caloric intake rather than these other potential causes.

It is of note that the magnitude of the productivity declines in agriculture are larger than those observed in the randomized trial presented in Chapter 1. Given that the RCT increases caloric intake, while fasting during Ramadan decreases caloric intake, this may be a function of decreasing marginal returns to caloric intake over this range of intake.[27] However, given the differences in the production functions in

[26] These calculations assume that all Muslims in the labor force observe the holiday. If not all individuals observe the holiday, these estimates would be a lower bound.
[27] A production function with decreasing marginal returns over this range of caloric intake is consistent with estimates by Strauss (1986).

these sectors, it is difficult to determine whether or how much of the difference in productivity is accounted for by the varied levels of caloric intake.

Despite the varied magnitudes of the impact of increased caloric consumption between these two studies, in both studies higher caloric consumption results in greater economic production. The consistency of these results across studies designed to be complementary suggests that the observed relationship is causal, economically meaningful, and widespread and that low levels of adult nutrition may play an important role in productivity and economic development.

References

Ahmad, S., K. Goel, K.A. Maroof, P. Goel, M. Arif, M. Amir, and M. Abid. "Psycho-Social Behaviour and Health Benefits of Islamic Fasting During the Month of Ramadan." *Journal of Community Medicine and Health Education* 2.9 (2012): 178-181.

Alhola, P. and P. Polo-Kantola. "Sleep Deprivation: Impact on Cognitive Performance." *Neuropsychiatric Disease and Treatment* 3.5 (Oct. 2007): 553–567.

Bahammam, A. "Does Ramadan Fasting Affect Sleep?" *International Journal of Clinical Practice* 12 (Dec 2006): 1631-1637.

Barr, S.I. "Effects of Dehydration on Exercise Performance." *Journal of Applied Physiology* 24 (1999):164 -172.

Belenky G., N.J. Wesensten, D.R. Thorne, M.L. Thomas, H.C. Sing, D.P. Redmond, M.B. Russo, and T.J. Balkin. "Patterns of Performance Degradation and Restoration During Sleep Restriction and Subsequent Recovery: A Sleep Dose-response Study." *Journal of Sleep Research* 12.1 (Mar. 2003): 1-12.

Blagrove, M., C. Alexander, and J.A. Horne. "The Effects of Chronic Sleep Reduction on the Performance of Cognitive Tasks Sensitive to Sleep Deprivation." *Applied Cognitive Psychology* 9.1 (Feb. 1995): 21–40.

Blackwell, A.H. *Ramadan*. New York: Chelsea House Publishers, 2009.

Brendel, D.H., C.F. Reynolds, J.R. Jennings, C.C. Hoch, T.H. Monk, S.R. Berman, F.T. Hall, D.J. Buysse, and D.J. Kupfer. "Sleep Stage Physiology, Mood, and Vigilance Responses to Total Sleep Deprivation in Healthy 80-year-olds and 20-year-olds."*Psychophysiology* 27.6 (1990): 677-685.

Candas, V., J. P. Libert, and J. J. Vogt. "Human Skin Wettedness and Evaporative Efficiency of Sweating." *Journal of Applied Physiology* 46.3 (1979): 522-528.

Casa, D.J. "Exercise in the Heat. I. Fundamentals of Thermal Physiology, Performance Implications, and Dehydration." *Journal of Athletic Training* 34.3 (1999a): 246-252.

Casa, D.J. "Exercise in the Heat. II. Critical Concepts in Rehydration, Exertional Heat Illnesses, and Maximizing Athletic Performance." *Journal of Athletic Training* 34.3 (1999b): 253-262.

Casa, D.J., P.M. Clarkson, and W.O. Roberts. "American College of Sports Medicine Roundtable on Hydration and Physical Activity: Consensus Statements." *Current Sports Medicine Reports* 4.3 (2005): 115-127.

Deaton, A., and J. Drèze. "Food and Nutrition in India: Facts and Interpretations." *Economic and Political Weekly* (2009): 42-65.

Directorate of Census Operations. "Census of India, 2011." Available online: http://www.census.tn.nic.in/. Cited Oct. 2013.

Donaldson, D. “Railroads of the Raj: Estimating the Impact of Transportation Infrastructure.” *AER* forthcoming (2013).

Drummond, S.P., M.P. Paulus, and S.F. Tapert. “Effects of Two Nights Sleep Deprivation and Two Nights Recovery Sleep on Response Inhibition.” *Journal of Sleep Research* 15.3 (Sep. 2006): 261-5.

Durmer, J.S., and D.F. Dinges. "Neurocognitive Consequences of Sleep Deprivation." *Seminars in Neurology* 25.1 (2005): 117- 129.

Ferrara, M., and L. De Gennaro. "How Much Sleep Do We Need?" *Sleep Medicine Reviews* 5.2 (2001): 155-179.

Ferro-Luzzi, A., C. Petracchi, R. Kurivan, A.V. Kurpad. “Basal Metabolism of Weight-stable Chronically Undernourished Men and Women: Lack of Metabolic Adaptation and Ethnic Differences.” *American Journal of Clinical Nutrition* 66.5 (1997): 1086-1093.

Fluck, R.C. “Energy of Human Labor.” *Energy In World Agriculture*. Ed. B.A. Stout. Amsterdam: Elsevier Science Publishers, 1992. 31-37.

Food and Agricultural Organization (FAO) of the United Nations. “FAOSTAT.” Available online: http://faostat.fao.org. Cited August 2013.

Food and Agriculture Organization (FAO) of the United Nations. “The State of Food Insecurity in the World.” September, 2010. Available online: http://www.fao.org/publications/sofi/en/. Cited Aug., 2011.

Goel, N., H. Rao, J.S. Durmer, and D.F. Dinges. "Neurocognitive Consequences of Sleep Deprivation." *Seminars in Neurology*. 29.4. NIH Public Access. (2009).

Gopalan, C., R. Sastri, and S.C. Balasubramanian. *Nutritive Value of Indian Foods*. Hyderabad: National Institute of Nutrition, Govt. of India, 1989.

Government of India, Department of Agriculture and Cooperation. *Comprehensive Scheme for Studying the Cost of Cultivation of Principal Crops in India*. New Delhi: Government of India, 1991.

Government of India, Department of Agriculture and Cooperation. *Cost of Cultivation of Principal Crops in India*. New Delhi: Government of India, 2000.

Grandjean, A.C. and N.R. Grandjean. “Dehydration and Cognitive Performance.” *Journal of the American College of Nutrition* 26.5 (2007): S549–554.

Greiwe, J.S., K.S. Staffey, D.R. Melrose, M.D. Narve, and R.G. Knowlton. "Effects of Dehydration on Isometric Muscular Strength and Endurance." *Medicine and Science in Sports and Exercise* 30.2 (1998): 284-288.

Guezennec, C.Y., P. Satabin, H. Legrand, A.X. Bigard. “Physical Performance and Metabolic Changes Induced by Combined Prolonged Exercise and Different Energy Intakes in Humans.” *European Journal of Applied Physiology* 68 (1994): 525-530.

Havenith, G. “Individualized Model of Human Thermoregulation for the Simulation of Heat Stress Response.” *Journal of Applied Physiology* 90 (2001): 1943-1954.

Havenith, G., V.G.M. Luttikholt, and T.G.M. Vrijkotte. "The Relative Influence of Body Characteristics on Humid Heat Stress Response." *European Journal of Applied Physiology and Occupational Physiology* 70.3 (1995): 270-279.

Kendall A.P., M.A. Kautz, M.B. Russo, and W.D. Killgore. “Effects of Sleep Deprivation on Lateral Visual Attention.” *International Journal of Neuroscience* 116.10 (Oct. 2006): 1125-38.

Keys, A., J. Brozek, A. Henschel, O. Mickelsen, and H.L. Taylor. *The Biology of Human Starvation*. Minneapolis: University of Minnesota Press, 1950.

Nag, P.K., N.C. Sebastian, and M.G. Mavlankar. “Occupational Workload of Indian Agricultural Workers.” *Ergonomics* 23.2 (1980): 91-102.

National Academy of Agricultural Sciences. “State of Indian Agriculture 2011-2012.” Available online: http://www.naasindia.org/Announcements/SIA.htm. Cited August 2013.

Rehrer, N.J., and L.M. Burke. "Sweat Losses During Various Sports." *Australian Journal of Nutrition and Dietetics* 53.4 (1996): S13-S16.

Reilly, T. and T. Deykin. “Effects of Partial Sleep Loss on Subjective States, Psychomotor and Physical Performance Tests.” *Journal of Human Movement Study* 9 (1983): 157-170.

Reilly T. and M. Piercy. “The Effect of Partial Sleep Deprivation on Weight-lifting Performance.” *Ergonomics* 37 (1994): 107-115.

Rogers, N.L., J. Dorrian, and D.F. Dinges. “Sleep, Waking, and Nuerobehavioral Performance.” *Frontiers in Bioscience* 8 (Sept. 2003): s1056-s1067.

Sawka, M.N., L.M. Burke, E.R. Eichner, R.J. Maughan, S.J. Montain, and N.S. Stachenfeld. "American College of Sports Medicine Position Stand. Exercise and Fluid Replacement." *Medicine and Science in Sports and Exercise* 39.2 (2007): 377-390.

Shapiro, Y., K.B. Pandolf, and R.F. Goldman. "Predicting Sweat Loss Response to Exercise, Environment and Clothing." *European Journal of Applied Physiology and Occupational Physiology* 48.1 (1982): 83-96.

Sharma, S. *Development and Democracy in India*. Boulder Co.: Lynne Rienner Publishers, 1999.

Sinnerton, S. and T. Reilly. “Effects of Sleep Loss and Time of Day in Swimmers.” *Biomechanics and Medicine in Swimming: Swimming Science VI*. Eds. MacLaren, D., T. Reilly, and A. Lees. London: Chapman and Hall, 1992.

Strauss, J. “Does Better Nutrition Raise Farm Productivity?” *Journal of Political Economy* 94.2 (1986): 297-320.

Van Dongen, H.P.A., G. Maislin, J.M. Mullington, and D.F. Dinges. "The Cumulative Cost of Additional Wakefulness: Dose-response Effects on Neurobehavioral Functions and Sleep Physiology from Chronic Sleep Restriction and Total Sleep Deprivation." *Sleep* 26.2 (2003): 117-129.

Van Helder, T., and M.W. Radomski. "Sleep Deprivation and the Effect on Exercise Performance." *Sports Medicine* 7.4 (1989): 235-247.

Vardar, S.A., L. Ozturk, C. Kurt, E. Bulut, N. Sut, and E. Vardar. "Sleep Deprivation Induced Anxiety and Anaerobic Performance." *Journal of Science and Medicine in Sport* 6 (2007): 532-537.

World Bank. “World Bank Development Indicators Database.” Available online: http://databank.worldbank.org/data. Cited August 2013.

Appendix 2A

Table 2A.1: Changes in Caloric Intake During Ramadan – Robustness Checks

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	Daily Per Capita Calorie Availability					
Topcoding Method	No topcoding	Above p99 to mean	Above p99 to p99	Above p99 to mean, conditional on positive consumption	Above p99 to p99, conditional on positive consumption	Winsorized at p1 and p99
Muslim	-107.5***	-96.84***	-89.49***	-106.7***	-92.46***	-82.57***
	[41.54]	[28.13]	[27.49]	[28.88]	[28.86]	[27.34]
Full survey period overlaps Ramadan	455.1*	279.8	395.8	405.6	427.2	432.07*
	[259.1]	[274.7]	[274.7]	[303.0]	[271.5]	[256.77]
Muslim*Full survey period overlaps Ramadan	-612.6	-601.3*	-620.2*	-880.2**	-725.5**	-545.35
	[384.8]	[340.3]	[336.9]	[400.6]	[368.9]	[356.5]
Mean of dependent variable	2263.49	2026.52	2138.46	2148.34	2200.51	2203.35
Observations	36,618	36,618	36,618	36,618	36,618	36,618
R-squared	0.095	0.188	0.213	0.174	0.211	0.236

Notes:
1. This table tests for changes in caloric consumption during Ramadan in Muslim households. The dependent variables are per capita daily caloric intake with six different methods of topcoding. All topcoding is done by food item before aggregating across the food items. Column (1) makes no adjustments to reported values. Column (2) reassigns values above the 99th percentile to the mean. Column (3) reassigns values above the 99th percentile to the 99th percentile. Column (4) reassigns values above the 99th percentile, conditional on having positive consumption of the food item, to the mean consumption conditional on having positive consumption. Column (5) is the same as Column (4) except the reassignment is to the 99th percentile, conditional on positive consumption. Column (6) is winsorized at the 1st and 99th percentiles.
2. The sample is drawn from the Indian National Sample Survey, Schedule 1 (consumption), rounds 60, 62, and 64 and is limited to households with survey periods with no overlap or full (29 day) overlap between the survey period and Ramadan. These rounds are included because they contain survey dates while earlier rounds do not.
3. All regressions include district-year fixed effects and month of interview fixed effects. In addition, indicator variables for other major religions and their interaction terms with overlap between the survey period and Ramadan as well as overlap between the survey period and Eid and the interaction between this variable and the religion indicators are included in the regressions but omitted from the table for simplicity.
4. NSS sampling weights are used and are reweighted to weight each round of the survey equally.
5. Robust standard errors clustered by district-year are in brackets.
6. *** p<0.01, ** p<0.05, * p<0.1.

Table 2A.2: Summary Statistics for Crops in the Agricultural Analysis

Crop Name	(1) Mean production by district	(2) Mean value by district	(3) Percent of districts with positive production	(4) Sowing start, SD	(5) Sowing end, SD	(6) Harvest start, SD	(7) Harvest end, SD
Bajra (pearl millet)	17.29	15.37	79.70	24.26	35.34	28.47	30.76
Barley	9.02	8.00	69.74	15.27	19.26	18.83	17.57
Cotton	3.74	9.38	73.06	41.95	45.38	57.05	59.64
Groundnut (peanut)	20.06	36.86	92.99	32.54	38.20	31.64	35.49
Gram (lentils, chickpeas)	18.19	23.62	99.26	13.73	18.93	22.44	21.54
Jowar (sorghum)	35.15	32.05	87.45	50.28	57.05	52.80	54.62
Jute	2.93	5.06	32.47	60.58	49.09	50.49	43.00
Maize	17.80	15.04	98.52	31.46	33.32	34.96	33.70
Other pulses	8.40	12.40	78.97	41.56	48.28	45.57	53.12
Potato	20.14	11.33	85.98	62.43	62.22	58.49	58.56
Ragi (finger millet)	0.88	0.78	11.07	41.87	50.75	40.67	56.26
Rice	134.61	162.93	97.79	99.86	88.17	81.54	82.70
Rapeseed and Mustard (oils)	2.72	6.04	87.08	17.80	19.73	24.71	23.47
Sesamum (sesame)	1.21	3.33	93.73	50.95	54.10	46.92	52.71
Sugar	46.87	62.98	98.52	88.36	102.80	65.04	67.13
Tobacco	1.31	6.16	90.77	60.04	76.01	60.00	69.12
Tur (pigeon pea)	8.68	12.75	89.30	27.29	25.45	48.01	39.47
Wheat	82.94	84.93	95.94	10.57	16.05	19.95	21.34
Average	24.00	28.28	81.24	42.82	46.67	43.75	45.57

Notes:
1) Production, Column (1), is in metric tons. Value, Column (2), is in 1,000,000 Rs deflated to 1973.
2) Columns (4) thru (7) provide the standard deviation of sowing and harvesting start and end dates to demonstrate the variation in the timing of cropping seasons.

Table 2A.3: Within District Variation in Cropping Cycles

	(1) Sowing start	(2) Sowing end	(3) Harvest start	(4) Harvest end
Average Standard Deviation	74	73	74	75
Average Range	261	258	233	245

Notes:
1. This table displays measures of the variation of timing of cropping cycles for different crops within a district.
2. The "Average Standard Deviation" is calculated as the standard deviation of the element of the cropping cycle listed at the top of the column within each district averaged across all districts. Similarly, the "Average Range" is the range in the timing of that element of the cropping cycle within a district, averaged across districts.
3. Roughly 7 percent of crop-district combinations have more than one cropping cycle per year. All cycles are included in these figures (as well as all analyses).

Table 2A.4: Distribution of Fraction of Ramadan Covered by Sowing and Harvesting Seasons

Percentile	Fraction of Ramadan covered by Sowing	Fraction of Ramadan covered by Harvest
50	0	0
75	0	0.03
90	0.86	0.93
95	1	1
99	1	1
mean	0.17	0.17
N	103,104	103,104

Notes:

1) This table provides distributional information for the primary independent variable of interest; the fraction of Ramadan covered by the labor intensive portions of the cropping season, sowing and harvesting.

2) The sample utilized for this table matches that of the primary specification listed in Table 2.2, Column (1).

Table 2A.5: Effect of Overlap Between Ramadan and Cropping Cycles (As a Fraction of the Season) on Output

	(1)	(2)	(3)	(4)
Dependent Variable	ln(q)	ln(value)	ln(q)	ln(value)
Fraction of sowing covered by Ramadan	0.000	-0.024	0.061***	0.034
	[0.019]	[0.02]	[0.023]	[0.024]
Fraction of sowing covered by Ramadan*Above median fraction Muslim			-0.13***	-0.125***
			[0.029]	[0.03]
Fraction of harvest covered by Ramadan	-0.041**	-0.069***	-0.002	-0.016
	[0.016]	[0.017]	[0.02]	[0.02]
Fraction of harvest covered by Ramadan*Above median fraction Muslim			-0.083***	-0.113***
			[0.026]	[0.026]
Mean of dependent variable	1.517	1.837	1.517	1.837
Observations	103,104	103,088	103,104	103,088

Notes:
1. This table tests for changes in agricultural output in each district-crop-year as a function of overlap between Ramadan and the sowing and harvesting seasons for that district-crop-year. Overlap is measured as the fraction of the total season(s) covered by Ramadan rather than the fraction of Ramadan covered by the season as in the primary specification. The dependent variables are: Columns (1) and (3), log production in thousands of tons; and Columns (2) and (4), log value of production (in 1,000,000 Rs deflated to 1973).
2. Regressions include district-crop, district-year, and crop-year fixed effects. In addition, time varying controls for average rainfall and temperature during the sowing and harvesting seasons, two month leads to each season, and a two month lag following the sowing season are included.
3. The agricultural data are from the India Agriculture and Climate data set. Crop cycles are from Donaldson (2013). Weather data are from the University of Delaware monthly rainfall and temperature series taken for the centroid of each district.
4. Robust standard errors clustered by district-year are in brackets.
5. *** p<0.01, ** p<0.05, * p<0.1.

Table 2A.6: Distribution of Muslims in Rural Areas

Percentile	Percent of the population that is Muslim
1	0.1
5	0.2
10	0.5
25	1.8
50	4.8
75	9.9
90	18.6
95	28.2
99	42.1
mean (district weighted)	0.076
mean (population weighted)	0.099

Notes:
1) Data in the table are drawn from the 1961 Indian Census and based on the 270 districts in the Indian Agricultural and Climate Dataset.
2) The census disaggregates the district into rural and urban areas. Given that the analysis examines agricultural production, the data here are drawn from rural areas only.

Table 2A.7a: Labor Supply and Earnings as a Test of Religious Obligations Driving Production Declines (ICRISAT)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent Variable	Labor Days	Average work hours	ln(Total earnings)	Labor Days	Average work hours	ln(total earnings)	Labor Days	Average work hours	ln(Total earnings)
Sample	Agricultural Laborers			Agricultural Laborers, High Season			Full		
Days overlap between survey period and Ramadan	-0.039	-0.002	-0.004***	-0.024	-0.002	-0.006***	-0.016	0.001	-0.002**
	[0.035]	[0.001]	[0.001]	[0.038]	[0.002]	[0.001]	[0.022]	[0.001]	[0.001]
Muslim*Days overlap between survey period and Ramadan	0.155*	-0.003	0.002	0.142	-0.008	-0.007	0.180***	-0.001	0.002
	[0.089]	[0.011]	[0.006]	[0.142]	[0.019]	[0.008]	[0.068]	[0.006]	[0.003]
Mean of dependent variable	21.752	7.131	6.816	21.874	7.212	6.778	17.921	7.182	7.014
Number of individuals	1,146	910	914	1,139	890	891	2,953	1,753	1,747
Observations	31,432	19,417	19,287	19,481	12,641	12,520	72,048	34,267	33,977
R-squared	0.699	0.775	0.659	0.754	0.820	0.670	0.631	0.780	0.760

Notes:
1. This table provides a test of whether time spent on religious activities causes reduced production during Ramadan by reducing the labor supply of Muslims during that period.
2. Samples are drawn from the second generation ICRISAT village level studies survey. "Agricultural Laborers" samples include only individuals who indicate that agricultural work is their primary occupation. "Agricultural Laborers, High Season" samples include observations from individuals in the "Agricultural Laborers" sample for which the interview period had at least 15 days of overlap with a "high labor demand" period defined as the sowing or harvesting seasons for the crop with the greatest acreage by state. The "Full" sample includes all individuals in the sample.
3. ICRISAT surveys participants approximately once per month and elicits information on labor supply, hours, and wages of the respondent during the month preceding the survey date. The number of days of overlap between Ramadan and the survey period is calculated as the number of days of Ramadan falling within the 30 days preceding the survey date. Labor supply is defined as the number of days of labor including both paid and unpaid labor but excluding domestic work in the past month. The wages variable is calculated as the sum of cash and in-kind wages during the month. Average hours worked is only reported for paid labor. If the participant reports more than one paid job, average hours worked is calculated as a weighted average across jobs. Religion is captured via the caste variable, which contains a category for Muslim individuals, because it is not directly reported.
4. All regressions include individual fixed effects, a control for the number of days between surveys, and year-month of interview fixed effects.
5. Robust standard errors clustered by individual are in brackets.
6. *** p<0.01, ** p<0.05, * p<0.1.

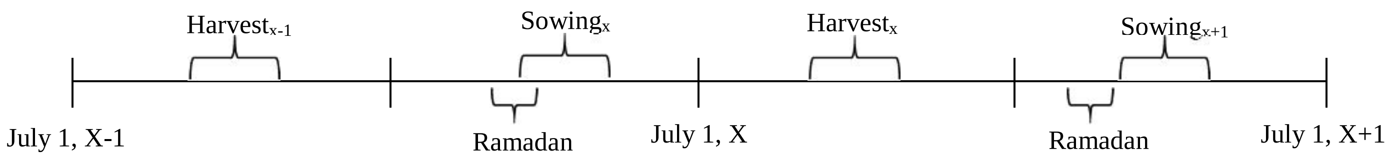
Table 2A.7b: Labor Supply and Earnings as a Test of Religious Obligations Driving Production Declines (NSS)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Dependent variable	Labor days	Total earnings	Ln(Total Earnings)	Labor days	Total earnings	Ln(Total Earnings)	Labor days	Total earnings	ln(Total earnings)	Labor days	Total earnings	ln(Total earnings)
Sample	Agricultural Laborers			Agricultural Laborers			Rural Casual Laborers			Rural Casual Laborers		
Timing	Full Year			High Labor Demand Seasons			Full Year			High Labor Demand Seasons		
Days overlap between survey period and Ramadan	-0.004	-0.689	0.039***	-0.035	4.168	0.040*	-0.022	-1.175	-0.003	-0.006	0.293	0.005
	[0.047]	[2.798]	[0.015]	[0.043]	[3.751]	[0.022]	[0.015]	[1.267]	[0.004]	[0.026]	[2.075]	[0.007]
Muslim	-0.025	-1.775	0.013	0.044	2.143	0.005	-0.039	2.436	0.032*	0.095	-6.211	0.018
	[0.051]	[4.551]	[0.025]	[0.08]	[8.176]	[0.045]	[0.067]	[6.845]	[0.018]	[0.101]	[16.354]	[0.027]
Muslim*Overlap between survey period and Ramadan	0.077***	-1.775	0.01	0.05	-4.375	-0.006	0.021	-0.597	0.002	-0.018	-1.726	-0.007
	[0.024]	[1.730]	[0.017]	[0.04]	[2.787]	[0.021]	[0.038]	[2.965]	[0.009]	[0.032]	[3.958]	[0.010]
Mean of dependent variable	5.852	103.391	5.521	5.808	105.968	5.397	4.811	278.428	5.657	4.895	266.677	5.581
Observations	121400	121408	35721	37882	37885	13825	105397	105397	83930	41593	41593	34625
R-squared	0.127	0.062	0.314	0.141	0.094	0.297	0.171	0.145	0.363	0.182	0.174	0.324

Notes:
1. This table provides a test of whether time spent on religious or social activities causes reduced production during Ramadan by reducing the labor supply of Muslims during this time. Given the somewhat coarse nature of the labor supply variable, the table also examines wages.
2. Samples are drawn from the Indian National Sample Survey, Schedule 10 (Employment), Rounds 60, 61, 62, 64, and 66. These rounds are selected because they contain survey dates while earlier rounds do not. "Agricultural laborer" samples include only individuals who indicate that agricultural work is their primary or secondary occupation. The "Rural casual laborers" sample includes individuals living in rural areas who indicate that they participate in the casual labor market. "Full year" samples include the full survey periods. "High Labor Demand" additionally restricts the samples to surveys conducted during the sowing or harvesting seasons for the crop with the greatest acreage by state.
3. The NSS Schedule 10 provides data on labor supply (to the half day) and wages of the respondent during the week preceding the survey date. Labor supply is calculated as the number of days of labor excluding domestic work in the past week. The wages variable is calculated as the sum of cash and in-kind wages during the week.
4. All regressions include district-year fixed effects and month fixed effects.
5. Robust standard errors clustered by district-year are in brackets.
6. *** p<0.01, ** p<0.05, * p<0.1.

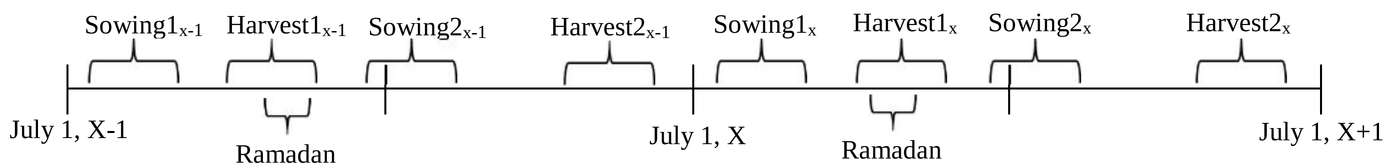
Figure 2A.1: Overlap Between Ramadan and Cropping Cycles, Calculation Examples

Example 1



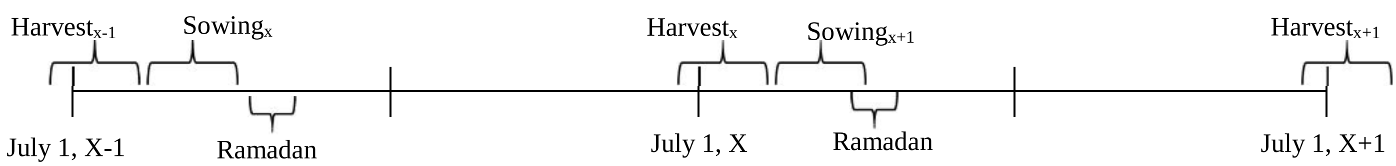
Overlap fraction Sowing, Year X = 0.4
Overlap fraction Harvest, Year X = 0

Example 2



Overlap fraction Sowing, Year X = 0
Overlap fraction Harvest, Year X = 1

Example 3



Overlap fraction Sowing, Year X = 0
Overlap fraction Harvest, Year X = 0

Notes: Agricultural years run from July 1 to June 30. A description of the overlap calculation procedure is provided in Appendix 2B.

Appendix 2B

Calculating the Fraction of Ramadan Covered by Agricultural Seasons

Data in the India Agricultural and Climate dataset is organized by the agricultural year running from July 1 of year X to June 30 of year X+1. To calculate the overlap with between the sowing and harvesting seasons and Ramadan, I first organize the crop calendar relative to the agricultural year such that an agricultural cycle for agricultural year X is defined by the seasons leading up to a harvest which occurs in agricultural year X.[28] This calendar is then overlaid with the Ramadan dates during the agricultural year of interest as well as the preceding year to account for extended agricultural cycles and harvests which occur early in the agricultural year. For example, the cropping cycle for sugar cane is typically between one and two years such that for a harvest occurring in agricultural year X, the sowing season typically occurs in year X-1. Similarly, if a harvest were to occur in July of agricultural year X, the sowing season would typically occur in year X-1. However, the overlap between these sowing seasons and the previous year's Ramadan would be assigned to the year in which the harvest occurred to coincide with the production data in the India Agricultural and Climate data.

In addition, because certain crops (e.g. rice, potatoes) frequently have more than one agricultural cycle per year, I calculate the days of overlap between each season (i.e. sowing, harvesting) and Ramadan for each cycle and then sum across the cycles ending in the same agricultural year. Finally, this total number of days of overlap is divided by 29, the number of days in Ramadan.[29] Figure 2A.1 provides example overlap calculations and Table 2A.4 provides the distribution of overlap between Ramadan and each of the agricultural seasons.

[28] For harvest seasons which overlap two agricultural years I assign the crop cycle to the year in which the majority of the harvest falls. Because some crops such as sugar cane have very long growing cycles, production accrued in year X can begin with a sowing season up to roughly 1.5 years earlier.

[29] The exact dates of Ramadan depend on the sighting of the crescent moon. However, data is not available on when the moon is sighted in each district. Hence, I use a consistent start date for the holiday across all of India for each year and limit the duration of the holiday to 29 days (rather than the possible 30) to generate conservative estimates and ensure that I do not measure overlap with Eid, the holiday following Ramadan.

Calculating Rainfall and Average Temperature Controls

Although the Indian Agricultural and Climate dataset contains information on rainfall and temperature, the values are averages across all years. Hence, I use the University of Delaware data to provide time varying controls for rainfall and temperature. Because each crop is likely to be impacted differently by rainfall and temperature in a given month, instead of using weather controls by calendar month I create rainfall and temperature relative to the sowing and harvesting seasons for each crop. Specifically, the regressions control for the total rainfall and average temperature over the two months preceding each season, during the season itself, and for the two months following the sowing season at the centroid of each district. Because the rainfall and temperature data is provided as a monthly series, these variable are calculated as weighted averages created by summing of the number of days in each month for the relevant period (e.g. sowing season) multiplied by the monthly value and dividing the total by the total number of days. Missing values are replaced with imputed values generated from a regression of existing values on district-crop and year fixed effects.

Calibrating Declines in Agricultural Productivity: "Free Energy" Available for Work

Adjustments to basal metabolic rates and energy use outside of work limit the precision of "free energy" calibrations. However, it is possible to calculate a measure of expected "possible" production based on caloric availability in order to determine whether the observed declines in productivity are roughly congruent with expectations based on energy availability. Rural residents in India were consuming roughly 2240 calories per day in 1983 (Deaton and Dreze 2009). Researchers have measured the basal metabolic rates (BMR) of rural populations in southern India and have determined that typical basal metabolic rates for low BMI individuals require approximately 1,100 and 1,400 calories per day depending on size and gender (Ferro-Luzzi et al. 1997).

Taking the lowest BMR estimates, the strong assumption that all calories not used in basal metabolism are used productively, and the estimated decline in caloric intake for Muslims during

Ramadan, individuals have 1140 calories available for work when not fasting and 440 available calories for work when fasting.

Passmore and Durnin (1955) review a variety of sources across five countries detailing energy expenditures in different agricultural activities and find that sowing season activities such as clearing brush, digging ridges, and ploughing typically require 4 to 9 calories per minute while harvesting activities such as bundling and threshing typically require 3 to 7 calories per minute. Hence, farm work is likely to burn roughly 180 to 540 calories per hour. Taking an estimate from the middle of this range of roughly 350 calories per hours, a typical farmer would be able to complete roughly 4 hours of active labor when not fasting and 2 hours of active labor when fasting. Accounting for the fact that some energy can be mobilized from fat reserves, these estimates suggest the observed 20 to 40 percent decline in productivity per Muslim individual is consistent with the free energy available.

Calibrating Expected Water Losses During Ramadan

Perspiration rates depend on a number of factors, key among which are the intensity of the physical activity and factors which influence evaporation rates (i.e., temperature, vapor pressure). The intensity of physical activity can be approximated via the rate at which calories are burned in the activity and the weather variables can be measured directly.

A 50 kg individual would be expected to burn approximately 1200 calories in a day of hard farm work during the active labor seasons (Fluck 1992; Nag et al. 1980). In order to benchmark the expected perspiration rates, I compare this energy usage to the energy consumed by running, one of the most commonly studied activities in which perspiration rates are measured, in similar weather conditions. Rehrer (1996) reviews the literature and finds that accounting for water intake, running a marathon typically results in a loss of approximately three percent of body weight. Although predicted rather than measured, Sawka et al. (2007) estimate that in temperatures similar to those likely to be experienced by farmers in India running a marathon would result in a loss of four to five percent of body weight.[30] However, running a marathon requires roughly 2500 calories for a 50 kg individual, twice the energy required for a day of intensive farm labor. Hence, one would expect roughly half the water losses for farming as well, suggesting an overall loss of roughly 1.5 to 2.5 percent in body weight.

However, two additional factors are likely to reduce sweat rates relative to this benchmark. First, the lower rate of caloric burn for farming will reduce sweat production due to greater passive cooling and less "wasted" sweat lost to dripping (Candas 1979; Shapiro et al. 1982). Second, a higher surface area to mass ratio (negatively correlated with BMI), lower overall weight, low body fat percentage, and acclimatization to heat all improve core body temperature management and substantially reduce sweat rates in hot humid climates (Havenith, Luttikholot, and Vrijkotte 1995; Havenith 2001; Casa 1999b). This suggests that Indian farmers are likely to have relatively low sweat rates relative to many individuals in the studies cited previously due to their small size, very low body fat percentage, and acclimatization to hot temperatures. Hence, it is unlikely that the two percent of body weight threshold at which decrements in performance begin to be observed is likely to be surpassed for a significant number of farmers during Ramadan.

Evaporative Potential (PET) and Maximum Temperature Calculations

To calculate the average evaporative potential and maximum temperature during the sowing and harvesting seasons I draw on data from the Climatic Unit Research Database at the British Atmospheric Data Center (BADC). BADC calculates a daily PET based on three measures of temperature (min, max, mean), vapor pressure, and cloud cover following a method recommended by the FAO. More information about the details of this measure are available on the BADC website: http://badc.nerc.ac.uk.

[30] This prediction is based on an ambient temperature of 28C, or approximately 82F. The overall mean temperature in the districts included in this study during sowing and harvesting periods was 27 and 24 degrees Celsius respectively (University of Delaware Air Temperature and Precipitation data, authors calculations). Although based on a more limited sample, the mean daily maximum temperature over the sowing and harvesting periods was 31 and 29 degrees C, respectively (Climatic Unit Research Database at the British Atmospheric Data Center, author's calculations).

Although the data are detailed, there is a significant amount of missing data. Hence, in order to calculate the average weather measure for a season, I begin by averaging the weather measure for all points within a district-month. That average is then assigned to the centroid so that, contingent on any readings being taken within a district-month, each district has one value. The average value at the centroid for each district month is regressed on district-month fixed effects and year fixed effects and missing values are assigned the predicted values from this regression. However, because some states do not have any data over the relevant time period and fitted values are unlikely to be representative in those areas, I limit the sample to states in which some data are available. This limitation drives the lower sample size in these regressions. Average evaporative potential over a season is calculated from the district-month values as the sum of the number of days in each month within the relevant season multiplied by the average evaporative potential in that month divided by the total length of the season.

3 Comparing the Effectiveness of Individualistic, Altruistic, and Competitive Incentives in Motivating Completion of Mental Exercises*

3.1 Introduction

Faced with a variety of challenges to public health arising from unhealthy behaviors, such as smoking, poor diet, sedentary lifestyles, and low rates of medication adherence, employers, health insurers, and government agencies have been dramatically expanding the use of monetary incentives to motivate healthy behavior change. Most of the programs that have been implemented in the field as well as most of the incentives tested by researchers (e.g., Cawley and Price 2009; Charness and Gneezy 2009; Perez et al. 2009) employ individual incentives. There are many reasons, however, to suspect that socially oriented incentives that play on motives such as competition and reciprocity, might have beneficial effects.

First, social motives can sometimes provide motivation that is disproportionate to the underlying magnitude of objective incentives. People will, for example, often reciprocate small gifts, such as the address labels provided by charities, or the flowers handed out by Harre Krishnas, with much larger return favors (Cialdini 2006; Fehr and Gächter 2000). Likewise, even in the absence of differential material incentives, pure completion and the feelings of "winning" or "losing" can substantially alter behavior and generate significant levels of effort (Delgado et al 2008). The social forces generated by teams or groups can also significantly increase effort and reduce the cost to produce a given amount of output (Nalbantian and Schotter 1997; Babcock et al 2012). By playing on such non-pecuniary motives, social incentives have, at least in theory, the potential to produce more substantial behavioral changes at lower cost than individualistic incentives.

* This chapter is coauthored with George Loewenstein, Jessica Kopsic, and Kevin G. Volpp. We gratefully acknowledge financial support from the Pennsylvania Department of Health's Commonwealth Universal Research Enhancement Program and the NIA-funded joint Penn/CMU Roybal P30 Center on Behavioral Economics and Health. We also thank Lumosity and Neurotrax for the generous use of their software, Matt Thompson for his outstanding website development, and Ben Cowan and ASHEcon conference participants for helpful comments. All errors are our own.

Second, by associating the desired behaviors with social cues rather than purely financial or economic cues, socially oriented incentives are less likely to be subject to the crowding out of intrinsic motivation that are a potential concern with the provision of financial incentives (Gneezy and Rustichini 2000; Benabou and Tirole 2003; Heyman and Ariely, 2004). Further, despite the predominant focus on individual-level incentives, social incentive schemes may be of particular interest in this domain because many health behaviors (e.g. eating, and some forms of exercise) have strong social elements, suggesting that the use of social forces may substantially augment the effectiveness of such programs both directly and indirectly.

Finally, the importance of peer effects and social motivations in altering behavior has been documented in a number of domains such as labor supply and financial decision making (Kaur, Kremer, and Mullainathan 2010, 2011; Bandiera, Barankay, and Rasul 2010; Duflo and Saez 2003). Although evidence within the health domain is limited, two recent studies coauthored by two of the authors of this paper found beneficial effects of social incentive programs. In the first study, veterans with poorly managed diabetes were either paid direct incentives for controlling their diabetes, or were paired with a peer-mentor whose diabetes had been, but no longer was, poorly controlled. Although both interventions led to improvement, the peer mentoring program was significantly more successful at lower cost (Long et al 2012). In the second study, employees were paid either individualistic rewards or organized into small groups in which joint rewards were allocated to group-members who lost weight. While the group incentive scheme was significantly more effective in motivating weight loss, it also provided higher rewards *ex post*, so it failed to provide a clean comparison of social and non-social incentives of similar value (Kullgren et al 2012). No studies that we are aware of, including the two just noted, have systematically compared the impact of social and non-social incentives of similar magnitude on desired health behaviors.

This paper provides a test of the comparative effectiveness of individualistic and socially oriented monetary incentives in motivating health-promoting behaviors via a randomized controlled trial among 312 elderly individuals who receive no monetary incentives, individualistic incentives, altruistic incentives, or team-based cooperative/competitive incentives for daily completion of cognitive training exercises. In all conditions, including the Control, individuals were randomly paired and provided with daily information about the number of exercises completed (and, if relevant, the earnings) by themselves and by their partner. While participants in the Control condition did not receive monetary incentives for completing cognitive training exercises, the magnitude of the incentives provided in the three treatment conditions were designed to be as similar to one-another as possible, so as to provide a clean test of their relative effectiveness in motivating engagement.

The use of online cognitive training exercises as an outcome serves a variety of purposes. First, given the rapidly aging United States population, cognitive decline is a substantial concern both in terms of population health and healthcare costs. In recent years Alzheimer's has become the sixth most prevalent cause of death in the United States, and accounts for an estimated direct costs of care of approximately $150 billion per year (Mebane-Sims 2009). Further, much of the cost of overall age-related cognitive decline is a result of milder forms of decline. By the sixth decade of life, losses in domains including reaction time, working memory, and attention are widespread (Bäckman, Small, and Wahlin 2006; Park and Payer 2006; Rogers and Fisk 2006). These declines are associated with decrements in functional performance on instrumental activities of daily living, such as problem solving and financial management (Marsiske and Margrett 2006; Finucane et al. 2005; Owsley et al. 2002).

Research examining the effectiveness of cognitive exercises in producing functional improvement on daily tasks or capabilities beyond performance on the exercises has generally been discouraging (Jaeggi et al. 2008). However, disappointing effects may stem in part from low rates of adherence to training programs and the lack of cost-effective approaches to improving adherence to these regimens. For example, the most comprehensive test of cognitive exercises, the ACTIVE study, had an overall budget of $15 million for 2,802 enrolled participants, or approximately $5,000 per participant over 24 months (Ball et al. 2002).

In addition to examining an important domain of health, use of a web-based cognitive training task also facilitates accurate, high frequency data collection as well as high frequency feedback and incentive provision, features which are often difficult to achieve in research of this type. For example, studies examining the impact of incentives on gym usage have generally focused on attendance, measured by sign-ins; it is much more difficult to monitor how much exercise participants complete after signing in. The use of an on-line platform also provides the potential for scalability, since replicating a web-based intervention is much easier than one which involves physical facilities and/or personnel. These technologies have the potential to reach individuals and promote healthy habits on a daily basis at low cost and in an automated fashion.

Finally, the online platform also provides an opportunity to examine the long run impacts of the provision of incentives with minimal experimental demand effects. In this study, the active study period in which incentives are provided lasted six weeks. However participants were given one year of continued access to the cognitive training exercises following the completion of the study, which made it possible to track continued engagement with the exercises after the removal of incentives.

During the six-week active study period, individuals in all experimental groups, including the Control, were paired with another participant in the study and received free access to the training software and information about their own and their partner's completion of exercises. Participants in the incentive groups, but not the Control group, were additionally eligible to receive financial compensation during this period. In the individual incentives condition, referred to as the Atomistic treatment, participants were provided with a flat payment of approximately $0.17 per exercise up to 30 exercises per day. In the Altruistic treatment, the level of compensation was the same; however individuals were paid as a function of the number of exercises completed by their partner rather than as a function of their own exercise completion. In the Cooperative/Competitive treatment, teams, each consisting of a pair of individuals, were randomly paired to form quads, and each of the teams was compensated as a positive function of the fraction of the exercises completed by that team and negative function of the total exercises completed by the opposing team. Further details regarding the exact payment structures are given below.

We find that the use of any monetary incentives, whether direct or socially motivated, approximately doubled engagement with the cognitive training exercises. Surprisingly, Altruistic treatment and Cooperative/Competitive treatments (both of which had much lower average marginal benefit per exercise to the individual engaging in the exercise) generated gains in the number of exercises completed that are statistically indistinguishable from those in the atomistic condition. Despite similar gains in exercises completion across incentivized treatments, we did observe very different patterns of engagement in pairs of participants across the experimental treatments. We also found that while utilization of the software led to substantial improvements on scores in the majority of the incentivized exercises, the gains did not typically generalize to improvements in measures cognitive function more broadly. Finally, examining utilization of the software following the completion of the experimental period, despite dramatic declines across all experimental groups, there were significant differences in the rate of decline across conditions. During the five month follow up period, roughly twice as many exercises were completed by participants in the socially oriented treatments than by participants in the Atomistic and Control conditions.

3.2 Experimental Design

3.2.1 Participants

Three hundred and twelve participants between the ages of fifty-five and eighty were recruited from adult education classes, churches, prior unrelated studies, Craig’s List, and community centers in Pittsburgh PA. All participants were screened either in person or by phone prior to entering the study. Individuals were excluded from the study if they had a history of stroke, dementia, Parkinson’s or Huntington’s Disease, Multiple Sclerosis, major psychiatric disorders, or were using medications to enhance cognitive ability. To participate in the study, individuals had to score at least 26 on the Telephone Interview for Cognitive Status-40 (TICS-40) (roughly equivalent to scoring 27 or above on the Mini-Mental State Exam (MMSE) (Fong et al. 2009)), to have fluent written and spoken English, proficiency with a computer, internet access, and ability to attend a training session and testing sessions at the beginning and end of the active experimental period in the office in Pittsburgh, PA.

Table 3.1: Baseline Participant Characteristics

	All	Control	Atomistic	Altruistic	Cooperative/ Competitive
Age	64.76	65.11	64.85	64.88	64.48
	[6.40]	[6.74]	[6.98]	[6.16]	[6.10]
Female	0.70	0.68	0.73	0.75	0.68
	[0.46]	[0.47]	[0.45]	[0.44]	[0.47]
Married	0.61	0.61	0.56	0.52	0.69
	[0.49]	[0.49]	[0.50]	[0.50]	[0.47]
Left handed	0.12	0.06	0.13	0.20**	0.10
	[0.33]	[0.25]	[0.34]	[0.41]	[0.31]
Not born in US	0.05	0.05	0.03	0.03	0.06
	[0.21]	[0.22]	[0.18]	[0.18]	[0.25]
Retired	0.61	0.61	0.63	0.66	0.58
	[0.49]	[0.49]	[0.49]	[0.48]	[0.50]
White/Caucasian	0.93	0.89	0.92	0.92	0.95
	[0.26]	[0.32]	[0.27]	[0.27]	[0.22]
Family member has/had dementia	0.62	0.58	0.65	0.67	0.60
	[0.70]	[0.74]	[0.70]	[0.71]	[0.67]
Normalized cognitive test score at enrollment	107.32	106.81	107.49	107.47	107.41
	[7.03]	[7.45]	[6.30]	[7.48]	[7.00]
Education					
Less than BA	0.14	0.15	0.19	0.17	0.10
BA	0.32	0.35	0.26	0.30	0.35
More than BA	0.48	0.48	0.47	0.45	0.51
Other	0.05	0.02	0.08	0.08	0.05
Median Household Income Range (USD)	50,000-74,999	50,000-74,999	35,000-49,999	50,000-74,999	50,000-74,999
Observations	312	62	62	64	124

Notes:
1. This table contains the mean and standard deviation of participant characteristics as reported at enrollment.
2. *** Significant at the 1 percent level, ** Significant at the 5 percent level, * Significant at the 10 percent level.

Baseline participant characteristics are presented in Table 3.1. With fewer than one in twenty significant differences between conditions, this table suggests that the randomization was successful in producing comparable samples in each of the four conditions.

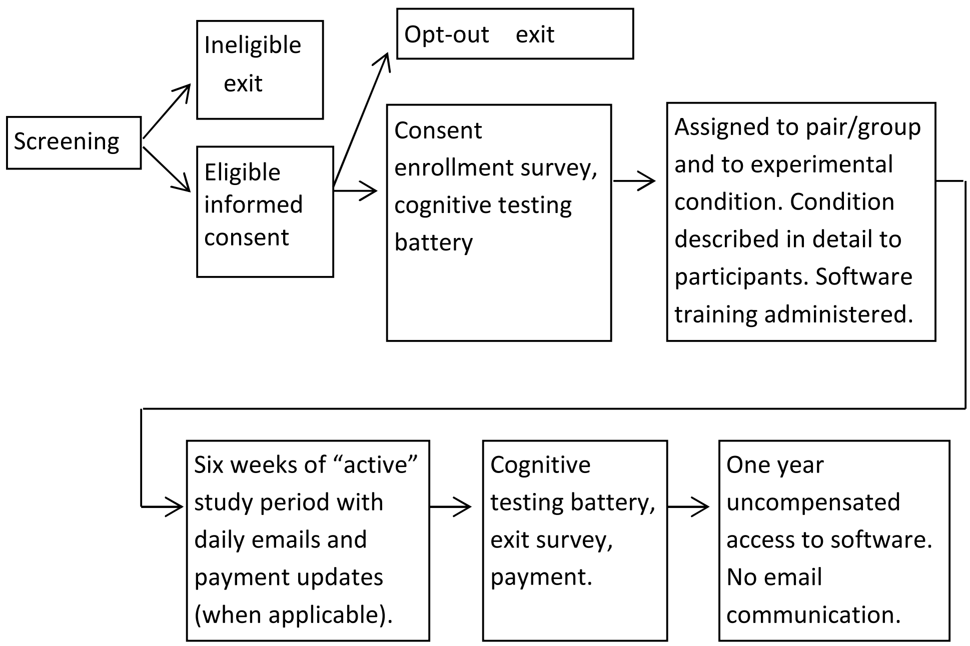
3.2.2 Experimental Timeline and Online Platform

After being screened and completing the informed consent process, participants visited the lab and completed an enrollment survey and a 45 minute baseline battery of computer based cognitive tests utilizing the Mindstreams software developed by Neurotrax.[1] Following the cognitive testing, participants were randomly paired and/or grouped and assigned to an experimental condition. Participants also completed an in-person training session to familiarize them with the cognitive exercise software, including the exercises themselves, the website's messaging features which allowed them to communicate directly with their partner(s), and the information available through the website (their performance, their partner's performance, and, when relevant, each person's earnings). Participants in the treatment groups were also given extensive instruction, in verbal, mathematical, and graphical form, about the monetary incentive structure to which they were randomized. During the six-week study period all participants received free access to the cognitive training software and daily emails regarding their own and their partner's/group's engagement (and earnings, when relevant). In addition, all participants, including those in the Control group, could access, via the website at any time, information regarding their own use of the software as well as their partner's use of the software. This information was updated in real time. At the completion of the study period, participants completed an alternate version of the cognitive testing battery they had completed at intake, and an exit survey.[2] After the active study period all participants were given continued free access to the cognitive training software, and usage was monitored; however no further emails were sent, no information about the partner's utilization of the software was available, and no further payments were made for use of the software. Figure 3.1 details the participant timeline.

[1] The testing battery has also been used in over 50 published peer reviewed studies and allows for measurement of cognitive function in a variety of domains such as reaction time, attention, and memory. Over 17,000 patients have completed this testing battery. More information is available at http://www.neurotrax.com.

[2] The cognitive tests taken at enrollment and the completion of the active experimental period are identical in nature and design, however the stimuli vary between versions to minimize test-retest effects. Correlation in performance across versions is very high (see the Neurotrax website for more details).

Figure 3.1
Participant Timeline



3.2.3 Cognitive Training Software

The cognitive training software consisted of eleven exercises targeting five cognitive domains, including spatial orientation, problem solving, memory, executive function, and reaction time. The exercises used in the training software were provided by Lumosity,[3] a firm that provides online cognitive training exercises. The average exercise took approximately two to three minutes to complete, however the range in duration was approximately one to ten minutes depending on the exercise and the individual's skill level. To ensure that participants were exposed to the full range of exercises, the eleven exercises were presented in a quasi-random order which was changed daily.

[3] Lumosity is a commercial developer of cognitive training exercises. The firm collaborates with cognitive science researchers from a variety of institutions such as Stanford and UCSF in developing their training exercises. More information is available at www.lumosity.com.

3.2.4 Experimental Treatments

Payment formulas for all experimental conditions are presented in Table 3.2. As noted, individuals in all conditions, including the Control, were paired and provided with access to information about their partner's participation (and payments, if relevant), and could also communicate with their partner via a messaging feature of the website.

Table 3.2: Experimental Conditions and Payments

Experimental Condition	Description	Payment formula
Control	No payment	$P_1 = 0$ $P_2 = 0$
Atomistic	Flat rate of \$0.17 per exercise completed	$P_1 = \frac{E_1}{6}$ $P_2 = \frac{E_2}{6}$
Altruistic	Flat rate of \$0.17 paid to partner for each exercise completed	$P_1 = \frac{E_2}{6}$ $P_2 = \frac{E_1}{6}$
Cooperative/ Competitive	Marginal payments vary as a function of exercises by both teams. Team members earn the same amount.	$P_1 = P_2 = \frac{Max[(E_1 + E_2),(E_3 + E_4)]}{6} \; \frac{(E_1+E_2)}{(E_1 + E_2 + E_3 + E_4)}$ $P_3 = P_4 = \frac{Max[(E_1 + E_2),(E_3 + E_4)]}{6} \; \frac{(E_3 + E_4)}{(E_1 + E_2 + E_3 + E_4)}$

Notes:
1. E_x = Exercises completed by partner x, P_x = Payment to partner x.

Individuals assigned to the Control group were provided with free access to the cognitive training software, messaging service, and emails, but were not given any monetary incentives to utilize the software. Participants in the Atomistic treatment were provided with a flat rate monetary incentive of approximately $0.17 per exercise up to 30 exercises for their own participation, resulting in maximum daily earnings of $5. Participants in the Altruistic treatment were compensated at the same rate of $0.17 per exercise. However, their compensation depended on the number of exercises completed by their partner rather than of their own level of participation. Hence, while participants in this treatment could potentially improve their cognitive health via the training, they received no direct financial benefit from completing additional exercises.

Finally, participants in the Cooperative/Competitive treatment were paired with a partner to form a team, and two pairs/teams were matched to form a group of four. The incentives in this treatment were designed to encourage cooperation between members of the teams and competition between the teams. To accomplish this, individuals in this treatment were compensated as a function of both the relative level of participation between the two teams and the total number of exercises completed by the team with the highest level of participation. Specifically, the total amount of money available to be distributed among the group of four was the maximum number of exercises completed by either team multiplied by $0.34. The money was then allocated between the two teams in direct proportion to the number of exercises completed by each team. Each member of a team/pair received the same compensation for a given day. This design provides a strictly positive marginal payment for the individual completing the activity and also for their partner (up to the 30 exercise per participant limit, consistent with the other treatments). However, the marginal payment for one exercise by one member of the team varies significantly and ranges from less than $0.01 to $0.17 per partner[4] based on the performance of both teams. Due to the fact that the payment from each exercise is split between members of the team, individuals in the Cooperative/Competitive treatment receive a weakly lower payment per exercise for themselves than

[4] Note that a marginal payment of $0.34 for the team as a whole is a marginal payment of $0.17 for each member of the team. Marginal payments are high when one team has not completed any exercises but the other team has not yet reached the 30 exercise per participant limit so the total amount available is growing but is only allocated to one team. On the other hand, marginal payments are low when one team has completed the maximum incentivized number of exercises and the other has completed very few because the total amount to be distributed does not grow when the low playing team engages, but the fraction reallocated towards the low engagement team is small. Despite the variability in the marginal payments, the median payment per exercise in this condition was $0.17 and the average payment per exercise was $0.23.

individuals in the Atomistic treatment. But, to keep the total possible payments the same, this difference is compensated for by the fact that when an individual's partner completes an exercise, that individual receives a payment without having completed any exercises. This structure encouraged cooperation among team members (each team member's work benefits the other; both had to participate to get the maximum possible earnings), but competition between the two teams (once the maximum number of exercises was reached by either team the payments became zero-sum across the group).

3.3 Results

3.3.1 Completion of Exercises

There is a large main effect of treatment on engagement with the cognitive exercises. Individuals in the no payment Control completed an average of 11.7 exercises per day (roughly 30 minutes of daily engagement with the software). Individuals in each of the treatment groups completed approximately twice that number, a large and statistically significant increase (See Table 3.3 and Figure 3.2)[5]. The increase in engagement in the treatment groups is statistically indistinguishable across the three treatment arms, with an average of 22.4, 23.1, and 25.5 exercises per day for the Altruistic, Atomistic, and Cooperative/Competitive groups respectively. This result is particularly striking given that participants in the Altruistic condition receive no direct monetary benefit from completing exercises and that the marginal payment for completing an exercise in the Atomistic treatment weakly dominates the payment for completing an exercise in the Cooperative/Competitive condition. Further, while the marginal payment in the Cooperative/Competitive condition was variable, and depended on the level of utilization of both teams, the mean payment per exercise ($0.23) was quite similar and the median payment per exercise ($0.17) was nearly identical to the other compensated treatments.

[5] As expected given the well balanced randomization, results were qualitatively similar with and without controlling for baseline characteristics. Hence, additional covariates are omitted to simplify regression results. All regressions were clustered at the level of the pair for the Control, Atomistic, and Altruistic conditions and at the level of the group (two teams each consisting of a pair) for the Competitive/Cooperative condition.

Table 3.3: Exercises Per Day

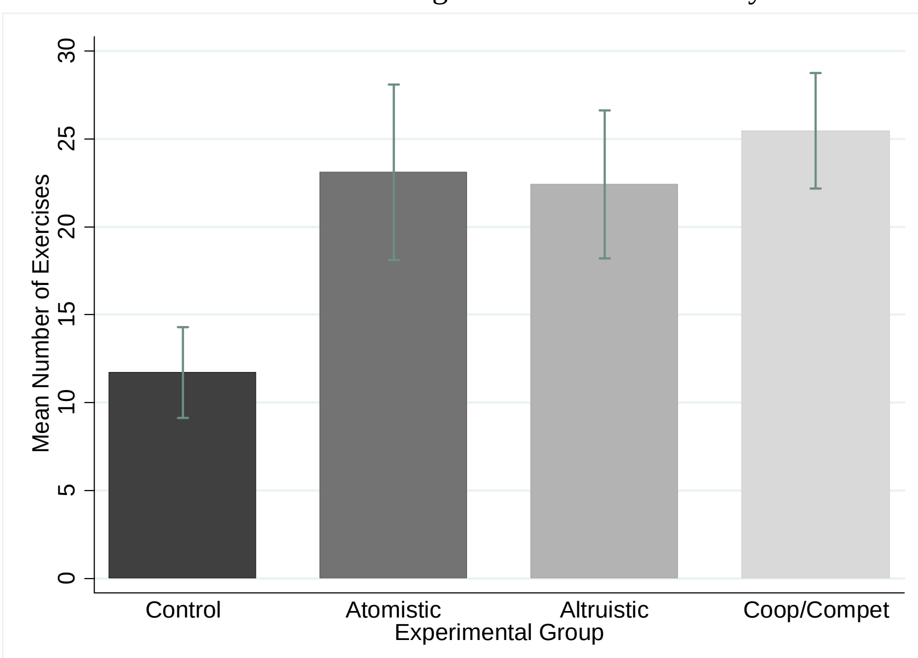
Atomistic	11.39***
	[2.15]
Altruistic	10.70***
	[2.55]
Cooperative/Competitive	13.76***
	[1.68]
Constant	11.72***
	[1.32]
Observations	13,104
R-squared	0.11

Notes:

1. This table reports the OLS regression of the number of exercises completed on indicator variables for each experimental condition.
2. The unit of observation is the participant-day.
3. Standard errors clustered at the level of the pair for all experimental groups except Cooperative/Competitive which is clustered at the level of the group.

4.*** Significant at the 1 percent level, ** Significant at the 5 percent level, * Significant at the 10 percent level.

Figure 3.2
Mean Number of Cognitive Exercises Per Day



This dramatic increase in engagement with the exercises in the treatment groups is the combined effect of both extensive margin changes (i.e. more regular use of the software) and intensive margin changes (i.e. greater participation conditional on logging into the website) (distributional information for the number of exercises completed is presented in Table 3.4). Individuals in the control group logged in 65.8 percent of the days while participants in the Altruistic, Atomistic, and Cooperative/Competitive groups logged on 81.0 percent, 80.3 percent, and 87.5 percent of the days, respectively. Conditional on logging in and completing any exercises, the mean number of exercises completed in each group was 17.8 (Control), 27.7 (Altruistic), 28.8 (Atomistic), and 29.1 (Cooperative/Competitive). Hence, in addition to the large impact on daily use of the software, the treatments dramatically increased the number of exercises completed once logged in.

Table 3.4: Summary Statistics Of Daily Completion of Exercises

Percentile	Control	Atomistic	Altruistic	Cooperative/ Competitive
10^{th}	0	0	0	0
25^{th}	0	10	10	21
50^{th}	9	30	30	30
75^{th}	20	30	30	31
90^{th}	30	32	32	33
95^{th}	35	35	37	37
99^{th}	50	56	94	60
Mean	11.72	23.12	22.43	25.48
SD	12.67	15.03	17.65	12.86
Correlation with partner	0.15	0.12	0.36	0.22
Mean percent of days logging on	65.78	80.30	80.92	87.5
Mean exercises if exercises > 0	17.82	28.79	27.72	29.14

As shown in Figures 3.3a-3.3d and Figure 3.4, the higher average completion of exercises in the treatment groups is driven in large part by the substantial fraction of individuals completing exactly the maximum number of incentivized exercises, 30. While the most immediately striking feature of these figures is the large mass of individuals completing exactly 30 exercises in the treatment groups, the treatments also lead significantly more individuals to complete more than the monetarily incentivized number of activities. Specifically, averaging across participant-days, 19.2 percent, 21.4 percent, and 28.7 percent of the Altruistic, Atomistic, and Cooperative/Competitive groups engaged in more than 30 activities per day.

Figure 3.3
Cognitive Exercises Per Day By Experimental Condition

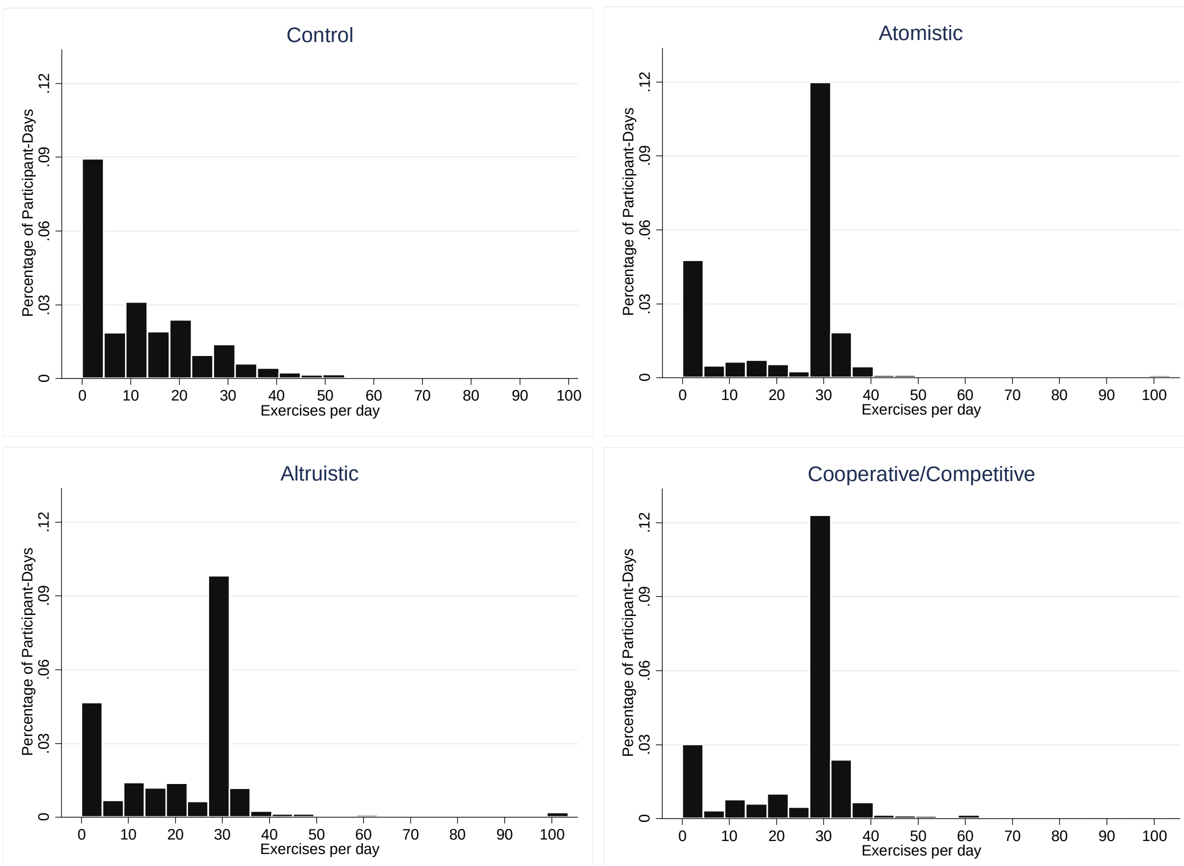
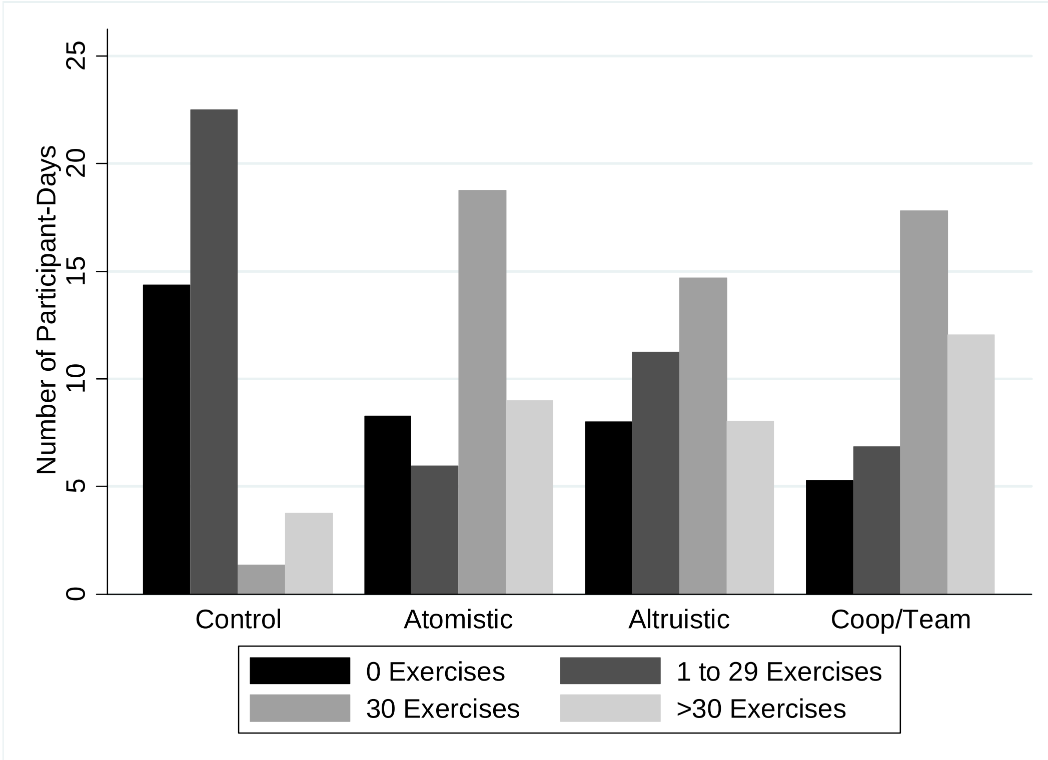


Figure 3.4
Mean Number of Days Completing N Exercises



Although the incentives offered in the three treatment groups had similarly large main effects on the average number of exercises completed per day, treatment assignments had differential impacts on the within-pair patterns of engagement with the exercises. Individuals assigned to the Control group have no financial interaction or interdependency; however play between partners is still correlated ($r = 0.15$, $p = 0.09$), providing evidence for peer effects resulting purely from the daily emails regarding how many exercises that individual and their partner completed. In the Atomistic treatment in which financial rewards are again unrelated to the partner's engagement, the correlation between partner's daily use of the software is very similar ($r = 0.12$, $p = 0.25$). In the Altruistic and Cooperative/Competitive treatments, in which financial rewards are contingent on one's partner's play, the correlation between partners increases to 0.36 ($p<0.01$) and 0.22 ($p=0.01$), respectively.

The simple correlations, while informative, hide other differences in patterns of concordance across treatments. Column (1) of Table 3.5 displays results from a linear probability regression examining the probability that an individual completes zero exercises as a function of their treatment group, binary variables indicating whether their partner completed zero exercises that day or the previous day, and treatment interacted with the binary variables. Column (2) presents similar results with binary variables for completing at least 30 exercises. Both the current day and lagged interaction terms are strongly positive and significant for the Altruistic condition indicating that individuals in this condition were more likely than individuals in the Control group to complete zero (30 or more exercises) if their partner did the same on either the current day or the previous day. Of further interest is the fact that, for the altruistic treatment, the point estimate for positive reciprocity is substantially larger, although not statistically distinguishable from, negative reciprocity. Negative reciprocity may be mitigated in these circumstances by the fact that the exercises are intended to promote health, so even if the absence of financial remuneration for engagement individuals are likely to engage for the health benefits.

To examine this potential avenue of influence between partners and the evolution of the spillovers over time we regress the number of exercises completed by the individual on their partner's exercise completion that day and the previous three days and treatment assignment in a fully interacted model (see Appendix 3A for Table 3A.1. Figure 3.5 summarizes the results from this regression). From the figure it can be seen that point estimates of all contemporaneous and lagged effects are positive and most are significantly different from zero. Initially (contemporaneous effects and one lag), reciprocity effects are greatest in the altruistic condition and second greatest in the cooperative/competitive condition. By two periods (days) back, however, the effects, while positive, are small and indistinguishable across conditions.

Table 3.5: Probability Of Completing Zero/More Than Thirty Exercises

	Column 1 0 exercises		Column 2 ≥30 exercises
Atomistic	-0.17*	Atomistic	0.46***
	[0.07]		[0.11]
Altruistic	-0.24***	Altruistic	0.04
	[0.07]		[0.05]
Cooperative/ Competitive	-0.26*** [0.06]	Cooperative/ Competitive	0.34*** [0.09]
$Partner_t = 0$ (binary)	-0.06	$Partner_t \geq 30$ (binary)	0.08
	[0.05]		[0.04]
Atomistic*$Partner_t = 0$	0.05	Atomistic*$Partner_t \geq 30$	0.00
	[0.09]		[0.08]
Altruistic*$Partner_t = 0$	0.28***	Altruistic*$Partner_t \geq 30$	0.30***
	[0.08]		[0.06]
Cooperative/ Competitive *$Partner_t = 0$	0.18* [0.07]	Cooperative/ Competitive *$Partner_t \geq 30$	0.11 [0.07]
$Partner_{t-1} = 0$ (binary)	-0.00	$Partner_{t-1} \geq 30$ (binary)	0.03
	[0.05]		[0.06]
Atomistic*$Partner_{t-1} = 0$	0.05	Atomistic*$Partner_{t-1} \geq 30$	0.03
	[0.10]		[0.10]
Altruistic*$Partner_{t-1} = 0$	0.17*	Altruistic*$Partner_{t-1} \geq 30$	0.32***
	[0.08]		[0.07]
Cooperative/ Competitive *$Partner_{t-1} = 0$	0.08 [0.07]	Cooperative/ Competitive *$Partner_{t-1} \geq 30$	0.15 [0.08]
Constant (Control)	0.36*** [0.06]	Constant (Control)	0.11*** [0.03]
Observations	13,104	Observations	13,104
R-squared	0.07	R-squared	0.33

Notes:

1. This table examines positive and negative reciprocity between partners [pairs in the cooperative/competitive condition] in each of the experimental conditions. Column [1] reports the results of a linear probability model regressing an indicator for whether an individual completes zero exercises on indicators for experimental condition, an indicator for whether their partner completed zero exercises that day and whether their partner completed zero exercises the previous day, and those indicators interacted with each experimental treatment. Column [2] has the same general design but studies the probability of completing at least 30 exercises.
2. The unit of observation is the participant-day.
3. Standard errors clustered at the level of the pair for all experimental groups except Cooperative/Competitive which is clustered at the level of the group.
4. *** Significant at the 1 percent level, ** Significant at the 5 percent level, * Significant at the 10 percent level

Figure 3.5
Conditional Correlations Between Partners

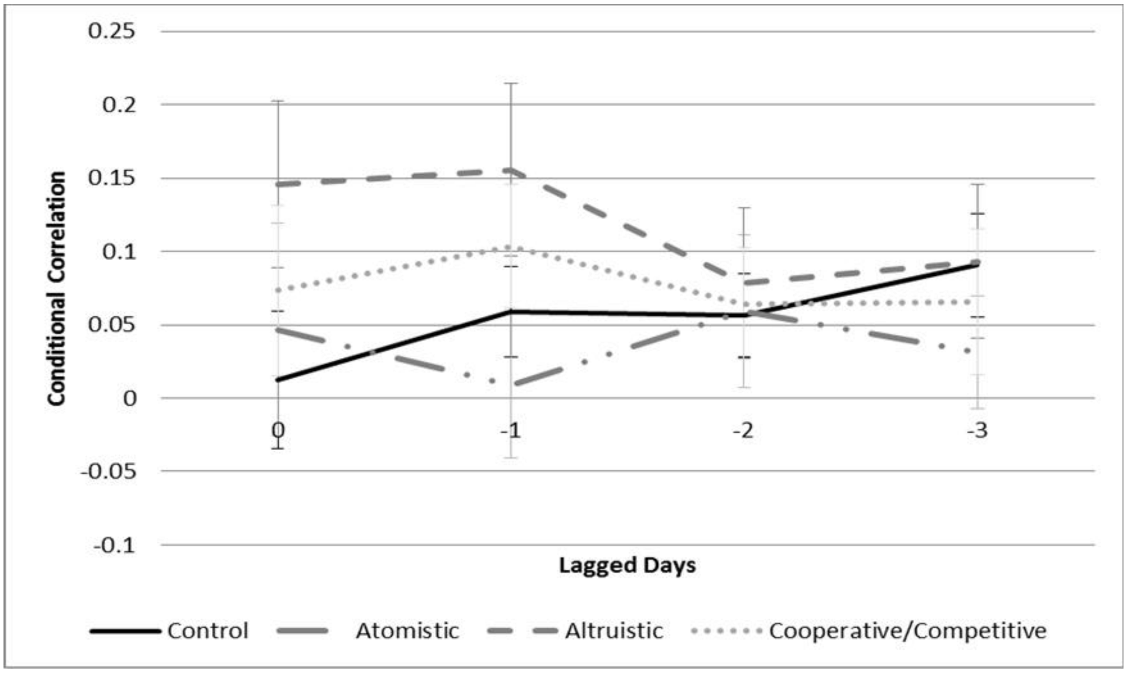


Figure 3.6
Between Partner Correlation in Exercises Completed

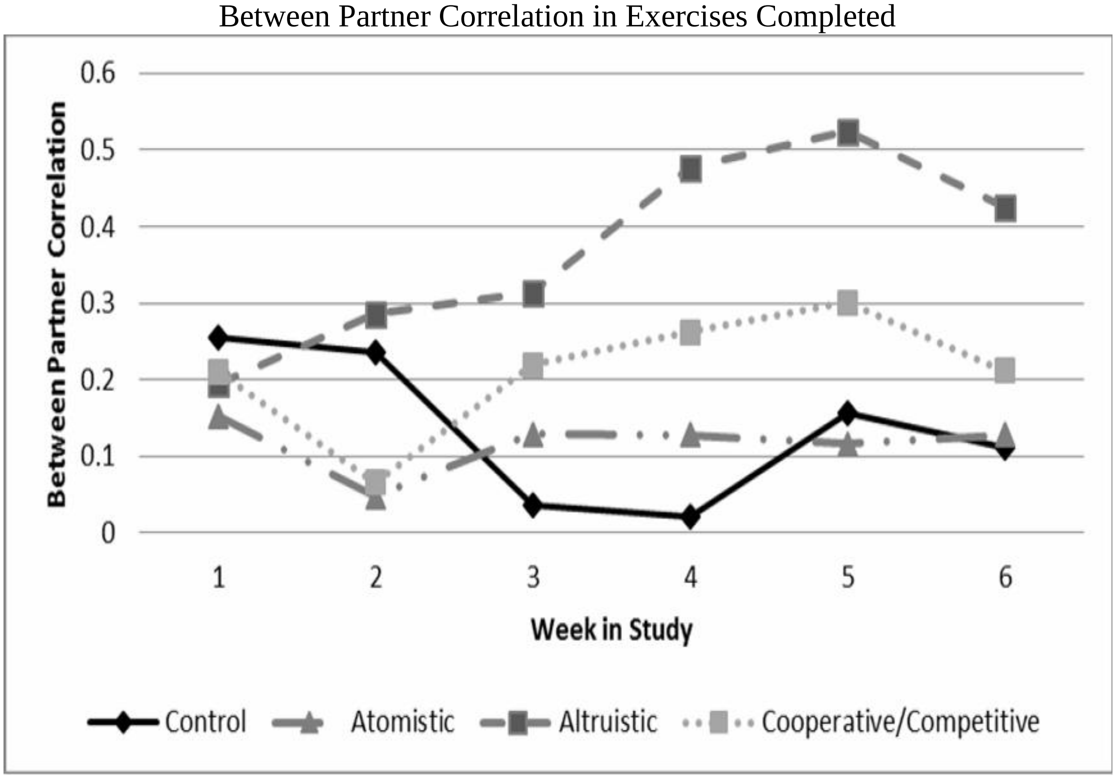


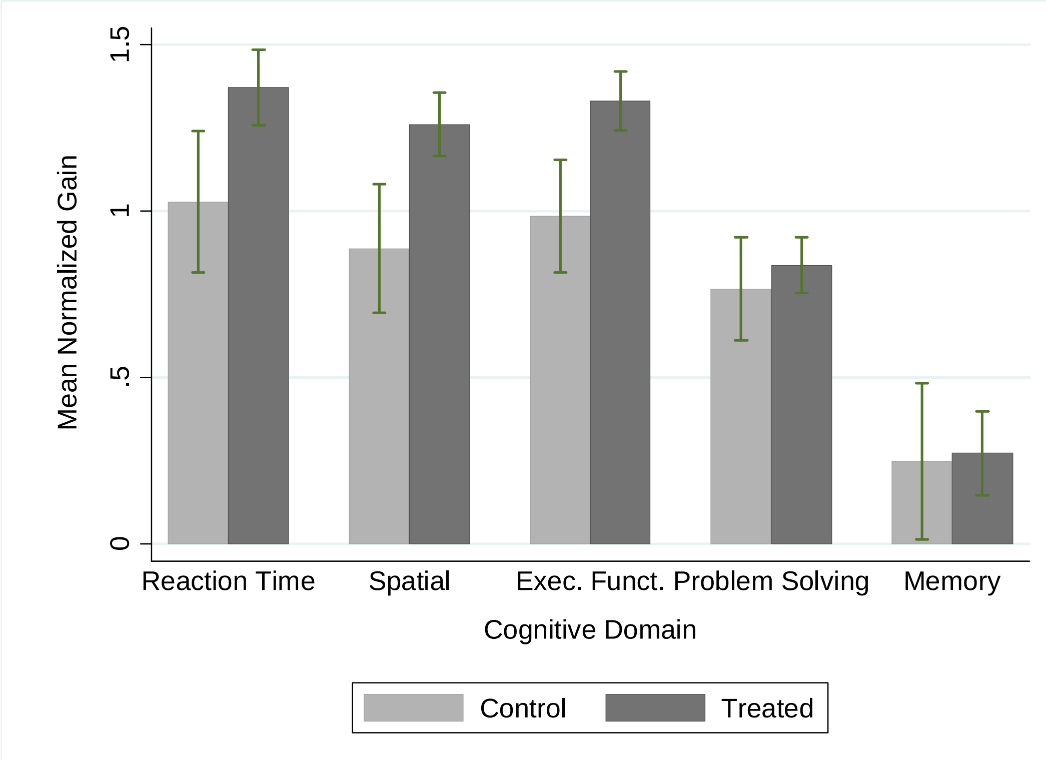
Figure 3.6, which displays the between-partner correlation in daily exercises by study week shows that these reciprocity effects grew stronger over the course of the six week intervention period in the altruistic and cooperative/competitive conditions in which payoffs were interdependent. In contrast, the correlation between partners' play declined over time in the control condition and remained fairly stable, but low, in the atomistic treatment condition.

3.3.2 Performance on Cognitive Exercises

Participants in all experimental treatments significantly improved their performance on all but one of the eleven cognitive exercises included in the software. These improvements were substantial in magnitude, typically between ¾ and 1 standard deviation. The doubling of exercises completed by individuals in the treatment groups also led to differential improvement in those groups relative to the Control group on approximately half of the 11 exercises (See Table 3.6 Panel A for changes in performance on each of the 11 exercises and Figure 3.7 for changes in performance when grouping exercises by cognitive domain. Table 3.6 also indicates which exercises, numbered from 1 to 11, target each of the cognitive domains. Exercise 2 is co-categorized in both Reaction Time and Spatial Reasoning). In particular, the differential gains were largest in exercises focusing on executive function, speed/reaction time, and spatial orientation, on which participants in the incentive treatment groups improved approximately ¼ to ½ of a standard deviation beyond the improvements in the control group on average. Although these estimates suggest decreasing marginal returns to additional exercises, the differential gains still represent substantial improvements on these exercises.

The improvements in scores on the exercises appear to be mediated by the increase in the number of exercises completed (See Table 3.6, Panel B). Each additional 100 exercises is associated with a gain of approximately 0.05 to 0.2 standard deviations. However, congruent with the results presented in Panel A, the negative coefficients on the squared terms indicated diminishing marginal returns.

Figure 3.7
Mean Normalized Gains In Exercise Scores By Cognitive Domain



Given that the number of exercises (i.e. quantity) is incentivized rather than the scores on the exercises (i.e. quality), it is possible the design of the incentives could crowd out the "quality" of the engagement and encourage participants to hurry through with little thought. Table 3.6 Panel B allows us to examine whether this occurred by testing whether receiving financial incentives impacts scores on the exercises, conditional on the number of exercises completed. To be explicit, because improvements in scores are a function of both practice (the number of exercises completed) and "quality" or concentration per exercise, if treated individuals exerted less cognitive effort per exercise we would expect treated individuals to obtain lower scores conditional on the amount of practice (number of exercises). Hence, if motivational crowd out in the quality of cognitive effort occurs we would expect that the regression coefficients on the Treatment indicator to be negative.

Table 3.6: Changes In Normalized Scores On Cognitive Exercises

	Cognitive Domain	Reaction Time	Spatial		Executive Function			Memory		Problem Solving		
	Exercise Number	1	2	3	4	5	6	7	8	9	10	11
Panel A	Treated	0.26*	0.45***	0.31**	0.39***	0.37***	0.30***	0.07	-0.02	-0.04	-0.02	0.28*
		[0.14]	[0.12]	[0.13]	[0.13]	[0.14]	[0.10]	[0.14]	[0.16]	[0.09]	[0.12]	[0.14]
	Constant (Control)	1.08***	0.96***	0.81***	1.06***	1.13***	0.74***	0.07	0.42***	0.75***	0.75***	0.79***
		[0.12]	[0.10]	[0.11]	[0.12]	[0.12]	[0.07]	[0.11]	[0.14]	[0.07]	[0.11]	[0.12]
	Observations	294	293	294	294	293	294	293	293	293	294	292
	R-squared	0.01	0.04	0.02	0.03	0.02	0.02	0	0	0	0	0.01
Panel B	Treated	-0.02	0.06	-0.08	-0.03	-0.12	0.00	-0.05	-0.13	-0.21**	-0.26**	-0.11
		[0.15]	[0.12]	[0.13]	[0.13]	[0.14]	[0.10]	[0.17]	[0.19]	[0.09]	[0.13]	[0.14]
	Total Exercises (00s)	0.09***	0.12***	0.12***	0.13***	0.17***	0.10***	0.06**	0.08***	0.05**	0.06***	0.11***
		[0.02]	[0.02]	[0.02]	[0.02]	[0.02]	[0.02]	[0.02]	[0.03]	[0.02]	[0.02]	[0.03]
	Total Exercises (00s) Squared	-0.002***	-0.003***	-0.003**	-0.003***	-0.004***	-0.002***	-0.002***	-0.003***	0.001	-0.001**	0.002
		[0.001]	[0.001]	[.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
	Constant (Control)	0.71***	0.44***	0.29***	0.50***	0.44***	0.34***	-0.15	0.16	0.54***	0.46***	0.28***
		[0.15]	[0.12]	[0.11]	[0.12]	[0.12]	[0.10]	[0.14]	[0.17]	[0.10]	[0.11]	[0.14]
	Observations	294	293	294	294	293	294	293	293	293	294	292
	R-squared	0.06	0.15	0.13	0.16	0.21	0.11	0.02	0.02	0.03	0.04	0.12

Notes:

1. Panel A reports results of OLS regressions of changes in scores on each cognitive exercise, as defined below, on an indicator for "Treatment" which includes all individuals in the Atomistic, Altruistic, and Cooperative/Competitive conditions. Panel B contains the results of OLS regressions of the same dependent variable on an indicator for treatment, the total exercises completed by each participant over the experimental period, and the square of that total.
2. Scores for each exercise are normalized to a mean of zero and standard deviation of one. Changes are defined as the last score – first score, conditional on having completed an exercise at least twice during the experimental period. Results are qualitative similar using averages of the last three scores – first three scores. Results are also similar examining indicators for each treatment rather than grouping all treatments together.
3. Exercises are categorized into cognitive domains as indicated by Lumosity, the company providing the software. Exercise number 2 has both spatial reasoning and reaction time components and hence is included in both of those categories.
4. The unit of observation is the participant-day. Standard errors clustered at the level of the pair for all experimental groups except Cooperative/Competitive which is clustered at the level of the group. *** Significant at the 1 percent level, ** Significant at the 5 percent level, * Significant at the 10 percent level.

However, as can be seen in Table 3.6 Panel B, nine of the eleven coefficients on the Treated variable are insignificant. Interestingly, the two coefficients (columns (9) and (10)) that are statistically significant are in exercises in which the duration of the exercise increases significantly with improved performance with a range of approximately three to ten minutes. Hence, in circumstances in which the exercises become much more taxing as participants improve, there is some slight evidence of "quality" crowdout, however, quality crowdout appears to have been minimal overall.

3.3.3 Performance on Cognitive Testing Battery

Although individuals in the treatment groups typically had greater improvement on scores on the training exercises, individuals in the incentive conditions did not show greater improvement on scores on the cognitive testing battery over the course of the six-week study as compared with the control group (See Table 3.7 Panel A).

There are universal improvements, defined below as the difference between exit score and baseline score, across all experimental groups in the cognitive testing battery.[6] However, these improvements may be at least partially due to a test-retest effect. As can be seen in Table 3.7 Panel B, while Processing Speed is significantly correlated with the number of exercises completed even after a Bonferroni correction, there is no significant relationship between the number of exercises completed and improvements on the cognitive testing battery for the overall cognitive score or three of the four cognitive domains[7].

Alternatively, it is possible that the large positive constants in these regressions suggest that even the lower levels of training done by the control group can be highly efficacious in increasing scores on this testing battery. However, the improvement of those individuals in the bottom decile of exercises per

[6] The Neurotrax Mindstreams software generates normalized scores for each of four domains as well as an overall "global" score based on a series of underlying tests such as Stroop, Go/No-go, and delayed recall tests.
[7] Results of two stage least squares regression using treatment as an instrument for the number of exercises completed provide similar results.

day (approximately 4 or fewer exercises per day, or less than one exercise per day in each domain) is statistically indistinguishable from that of individuals in the top decile of exercises per day (more than 31 exercises per day). Hence, the improvements from the training would need to be highly non-linear (i.e. all of the benefits accrue from the completing the first exercise or two) for this explanation to hold, suggesting that a test-retest effect is the more likely explanation of the majority of the results in this table.

Table 3.7: Changes in Mindstreams Cognitive Testing Scores By Domain

	Cognitive Domain	Global	Memory	Executive Function	Attention	Processing Speed
Panel A	Treated	0.03	-0.38	0.13	0.02	0.31
		[0.76]	[0.95]	[1.38]	[0.99]	[1.35]
	Constant	4.29***	2.38***	2.80**	2.73***	9.22***
	(Control)	[0.70]	[0.83]	[1.24]	[0.91]	[1.22]
	Observations	310	310	310	310	308
	R-squared	0.00	0.00	0.00	0.00	0.00
Panel B	Total Exercises	0.06	-0.01	0.01	-0.01	0.25***
	(00s)	[0.05]	[0.08]	[0.09]	[0.08]	[0.10]
	Constant	3.81***	2.16***	2.78***	2.81***	7.16***
		[0.62]	[0.81]	[1.02]	[0.82]	[1.01]
	Observations	310	310	310	310	308
	R-squared	0.00	0.00	0.00	0.00	0.02

Notes:

1. This table reports on changes in the cognitive testing scores from enrollment to the end of the active experimental period. Each column in Panel A is an OLS regression of the change in score within the cognitive domain indicated at the top of each column on an indicator [Treated] for belonging to Atomistic, Altruistic, or Competitive/Cooperative condition. The change in score is defined as the normalized score at the end of the active experimental period score minus the normalized enrollment score. Panel B includes OLS regressions of the change in score on the total number of exercises completed by the participant during the active experimental period.
2. The aggregate domain scores are calculated by the Mindstreams test from underlying tests (e.g. Stroop, Go/No-go) in each area of cognitive function. The "Global" score aggregates across all domains tested. More information on the tests used and the scores is available on the Mindstreams website: http://www.neurotrax.com.
3. Standard errors clustered at the level of the pair for all experimental groups except Cooperative/Competitive which is clustered at the level of the group.
4. *** Significant at the 1 percent level, ** Significant at the 5 percent level, * Significant at the 10 percent level.

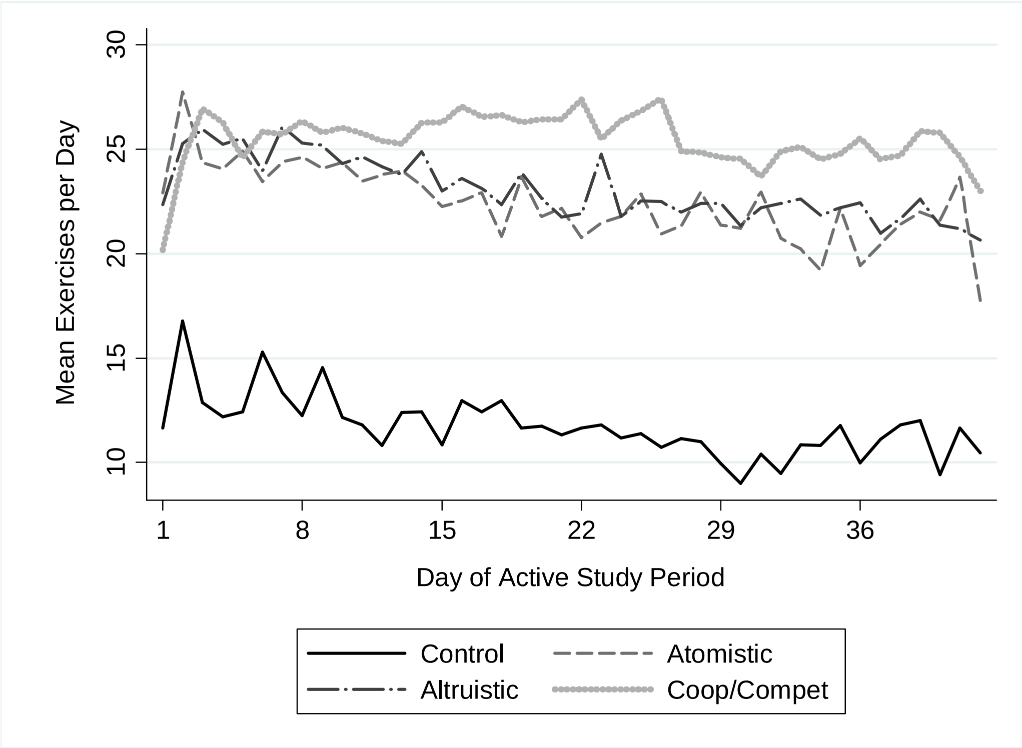
While these results are consistent with a large number of other studies which point to limited or no improvement on general cognitive tasks which are not specifically trained (Jaeggi et al. 2008), it is also possible that this particular cognitive assessment failed to capture changes generated by the training or that additional training is necessary to detect effects on this test.

3.3.4 Time Trends

One of the most significant challenges in changing health related behaviors is to maintain the behavior changes over time. During the six weeks of the active study period there was a small but statistically significant decline in the number of exercises completed in Control, Atomistic, and Altruistic conditions. The effect amounted to a decline of approximately 3.3 exercises per day over the course of six weeks, in a fairly linear trend of approximately 0.5 exercise per week. There was no similar decline in engagement in the Cooperative/Competitive treatment (See Figure 3.8).

At the conclusion of the six-week experimental period, participants were given continued access to the software; however the monetary rewards and daily information about their own and their partner's engagement with the software ceased. In contrast to the moderate decline in engagement with the software during the six-week experimental period (except in the Cooperative/Competitive treatment) there was a large and immediate decline in all experimental conditions at the conclusion of the study. In fact, the average total number of exercises completed per participant in five months following the completion of the study was only 84, or approximately ½ exercise per day. This low level of engagement is in sharp contrast to the previous overall average of 21.6 exercises per day during the active experimental period.

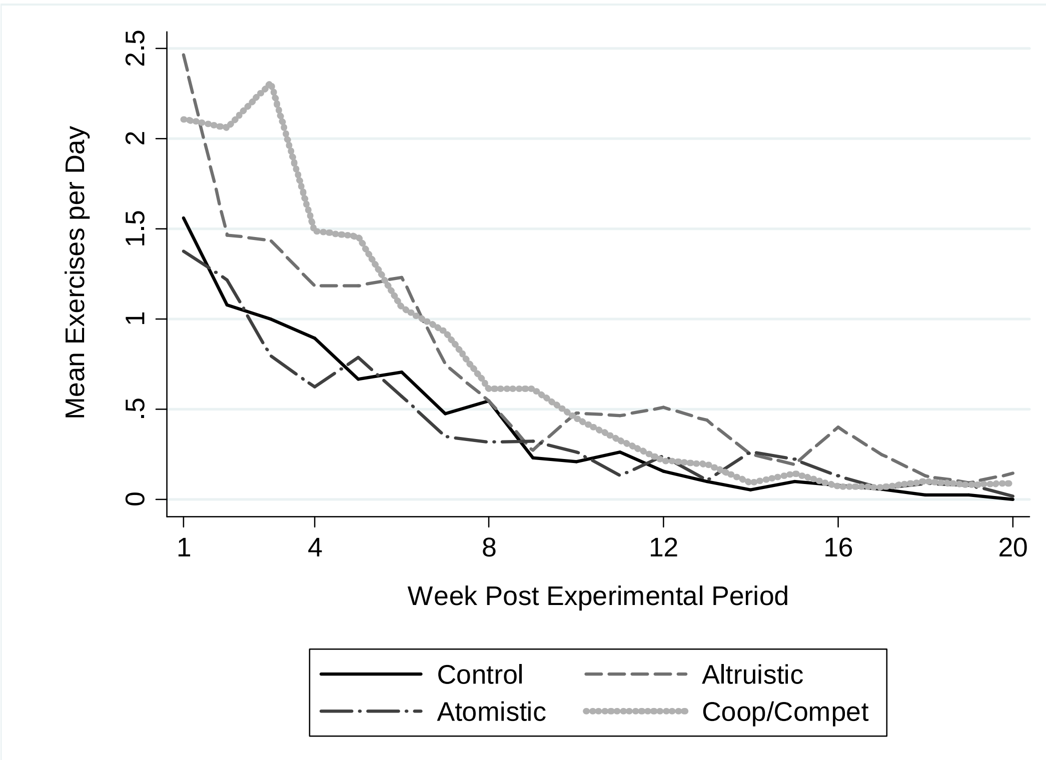
Figure 3.8
Mean Exercises Per Day By Treatment Group



However, a post-hoc analysis comparing the average number of exercises between socially oriented and individually oriented treatments reveals that individuals in the socially oriented incentive conditions completed nearly twice as many exercises (103 in Cooperative/Competitive and 98 in Altruistic) as individuals in the individually oriented conditions (58 in Control and 57 in Atomistic). The difference between the Cooperative/Competitive and Altruistic treatments and the Atomistic and Control treatments is marginally significant ($p = 0.06$) during the first month, but becomes insignificant as the treatments converge over time (See Figure 3.9).

The large standard errors on these estimates are due to substantial variation in the level of software utilization during this period. The fraction of individuals who never log onto the software again after the end of the experimental period is relatively constant across experimental conditions, ranging from 39 percent to 42 percent. However, approximately 7 percent of individuals continue to engage at meaningful levels (>5 exercises per day on average) for at least a month, and 77 percent of these individuals are in the Cooperative/Competitive and Altruistic treatments, a pattern of difference that persists, albeit more weakly, after the first month. Hence, these results suggest that the more socially oriented treatments, as compared with more individualistically oriented treatments, enhance intrinsic motivation, at least for some subset of the population.

Figure 3.9
Post Experimental Period Exercises



3.4 Discussion

In this experiment, all three types of monetary incentives, whether direct or socially motivated, approximately doubled engagement with the cognitive training exercises. Strikingly, the altruistically motivated incentives and the cooperative/competitive incentives (both of which had much lower average marginal benefit per exercise to the individual engaging in the exercise) generated gains in the number of exercises completed that are statistically indistinguishable from the direct monetary incentives in the Atomistic condition. The dramatic increase in the average number of exercises completed each day was the result of gains on both the extensive and intensive margins, with individuals in the treatments logging in on a larger fraction of the days and completing more exercises conditional on logging in.

Despite the fact that the gains in utilization of the software were statistically indistinguishable across the incentivized treatments, the patterns of engagement with the software among paired participants were strikingly different across the experimental treatments. While pairs of participants in the Control and Atomistic treatments exhibited modest correlations in exercises completed each day, suggesting the existence of spillovers purely from the information provided about the partner's use of the software, the correlation between partners in the Altruistic and Cooperative/Competitive conditions was both much higher and increasing over time.

Utilization of the software led to substantial improvements on the majority of the incentivized exercises; these gains were typically 0.75 to 1 standard deviation in the Control group and 1 to 1.5 standard deviations in the Treatment groups. Although there were also substantial gains on a cognitive testing battery administered at enrollment and again at week six, the gains appear likely to be driven primarily by a test-retest effect and did not differ between control and treatment groups, despite substantial differences in the numbers of exercises completed. This finding, which is suggestive of limited generalizability of cognitive changes, is consistent with a wide range of previous studies examining the impact of 'brain exercises' on generalized cognitive function (Jaeggi et al. 2008).

Following the conclusion of the experimental period, utilization of the software declined dramatically across all experimental groups. However, the decline was attenuated in the Altruistic and Cooperative/Competitive conditions. Individuals in these groups completed nearly twice as many exercises in the first month following the cessation of the intervention as individuals in the Control or Atomistic treatments, pointing to the possibility that the social forces generated by those treatments led to less crowding out, or more crowding in, of intrinsic motivation. These differences between conditions in post-incentive engagement, however, disappeared by the end of the second month following the removal of incentives.

The high levels of utilization of the cognitive training software during this study were striking. While the population was likely to be particularly motivated, a fact demonstrated by the substantial utilization even among the control group, financial and social incentives still resulted in large increases in the number of exercises completed. Although the high initial motivation of the participants may reduce generalizability of the magnitude of the effects, the fact that incentives improved engagement from already high levels suggests that these types of socially oriented monetary incentives could potentially have broad applicability.

The scalability of the online platform complements the scope of the socially oriented interventions, both in terms of facilitating further research and in terms of possible use in wellness programs or other contexts in which healthy behavior changes are promoted. From the perspective of study participants or individuals considering whether or not to join a wellness program, web-based platforms have the potential to greatly reduce costs.

Further, in terms of future research, the online platform, and in particular the cognitive training exercises, offer a unique opportunity to gather accurate high frequency data with minimal experimental demand. This feature is important because, although prior research examining the impact of monetary incentives on other health-related behaviors has yielded a number of interesting findings, this research has often been stymied by poor measures of incentivized behaviors. For example, in studies examining monetary incentives for gym attendance, attendance has been measured by card-swipes (e.g. Acland and Levy 2010; Charness and Gneezy 2009). However, it is unclear whether the individual actually completed any exercise or simply swiped the card to receive the promised rewards. A variety of other health behaviors such as medication adherence face similar challenges. Some studies have addressed behaviors with more directly verifiable outcomes such as weight loss or smoking cessation, but it is usually difficult to measure these behaviors with high frequency and accuracy in many settings, both limitations that likely diminish the effectiveness of incentives. Online cognitive training address these concerns by accurately capturing exactly how much exercise was completed and providing high frequency data that can be used to provide rapid accurate feedback and incentives.

Although the platform offers the benefits of scalability in a wide variety of domains, the differential effects on patterns of engagement suggest that the various incentive designs may be more or less appropriate for different health related behaviors. For example, in activities where individuals can "fall off the bandwagon" easily, altruistic designs may provide discouraging results because when one team member fails and is unable to get back on track there are likely to be spillovers to the other team member. The same is true of the competitive/cooperative condition. The higher correlation in behaviors between pairs in the two conditions involving social incentives is, thus, a kind of double-edged sword. On the one hand, each of these conditions may have been successful in channeling powerful social motives to the goal of motivating people to engage in cognitive exercises. On the other hand, the same connectedness between the players also introduces hazards in terms of likely non-engagement if one of the players drops out. This could happen for reasons that have nothing to do with lack of motivation, such as vacations, work, or lack of internet access but nevertheless effectively demotivate the other member of the pair. These are important factors to take into account when deciding what types of incentives to introduce in a particular setting.

The title of a recent important paper on motivational crowding out was "pay enough or don't pay at all." While not contradicting the findings of that paper, since our study focused on different issues and involved fairly substantial payments, the equivalent title for this paper could have been "pay enough, and it doesn't matter who or how."

References

Acland, Dan, and Matthew Levy, "Habit Formation and Naivete in Gym Attendance: Evidence from a Field Experiment," Working paper, 2010.

Babcock, Philip, Kelly Bedard, Gary Charness, John Hartman, and Heather Royer, "Letting Down the Team? Evidence of Social Effects of Team Incentives," NBER Working Paper No. 16687, 2011.

Bäckman Lars, Brent J. Small, and Åke Wahlin, "Aging and Memory: Cognitive and Biological Perspectives," in *Handbook of the Psychology of Aging, 6th ed.,* James E. Birren and K. Warner Schaie, eds. (New York, NY: Academic Press, 2006).

Ball, Karlene, Daniel B. Berch, Karin F. Helmers, Jared B. Jobe, Mary D. Leveck, Michael Marsiske, John N. Morris, George W. Rebok, David M. Smith, Sharon L. Tennstedt, Frederick W. Unverzagt, Sherry L. Willis, "Effects of Cognitive Training Interventions with Older Adults: A Randomized Controlled Trial," *Journal of the American Medical Association*, 288.18 (2002), 2271-2281.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul, "Social incentives in the Workplace," *Review of Economic Studies*, 77.2 (2010), 417-458.

Benabou, Roland, and Jean Tirole, "Intrinsic and Extrinsic Motivation," *Review of Economic Studies,* 70.3 (2003), 489-520.

Cawley, John, and Joshua A. Price, "Outcomes in a Program that Offers Financial Rewards for Weight Loss," NBER Working Paper No. 14987, 2009.

Charness, Gary, and Uri Gneezy, "Incentives to Exercise," *Econometrica,* 77.3 (2009), 909-931.

Cialdini, Robert, *Influence – The Psychology of Persuasion*, (New York, NY: Harper Business, 2006).

Delgado, Mauricio R., Schotter, A., Erkut Y. Ozbay, and Elizabeth A. Phelps, "Understanding Overbidding: Using the Neural Circuitry of Reward to Design Economic Auctions," *Science,* 321.5897 (2008), 1849-1852.

Duflo, Esther, and Emmanuel Saez, "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment," *Quarterly Journal of Economics,* 118.3 (2003), 815-842.

Fehr, Ernst and Simon Gächter, "Fairness and Retaliation: The Economics of Reciprocity," *The Journal of Economic Perspectives,* 14.3 (2000), 159-181.

Finucane Melissa L., C.K. Mertz, Paul Slovic, and Elizabeth Scholze Schmidt, "Task Complexity and Older Adults' Decision-Making Competence," *Psychology and Aging*, 20.1 (2005), 71-84.

Fong, Tamara G., Michael A. Fearing, Richard N. Jones, Peilin Shi, Edward R. Marcantonio, James L. Rudolph, Frances M. Yang, Dan K. Kiely, and Sharon K. Inouye, "Telephone Interview for Cognitive Status: Creating a Crosswalk with the Mini-Mental State Examination," *Alzheimer's & Dementia,* 5 (2009), 492-497.

Gneezy, Uri and Aldo Rustichini, "Pay enough or Don't Pay at All," *Quarterly Journal of Economics,* 115.3 (2000), 791-810.

Heyman, James and Dan Ariely, "Effort for payment: A tale of two markets," *Psychological Science*, 15.11 (2004), 787-793.

Jaeggi, Susanne M., Martin Buschkuehl, John Jonides, and Walter J. Perrig, "Improving Fluid Intelligence with Training on Working Memory," *Proceedings of the National Academy of Sciences*, 105.19 (2008), 6829-6833.

Kaur, Supreet, Michael Kremer, and Sendhil Mullainathan, "Self-Control and the Development of Work Arrangements," *American Economic Review Papers and Proceedings,* 100.2 (2010), 624-628.

Kaur, Supreet, Michael Kremer, and Sendhil Mullainathan, "Works Well with Others: Peer Effects and Social Motivation," Working paper, 2011.

Kullgren, Jeffrey T., Andrea B. Troxel, George Loewenstein, David A. Asch, Laurie A. Norton, Lisa Wesby, Yuanyuan Tao, Jingsan Zhu and Kevin G. Volpp, "Individual vs. Group-Based Incentives for Weight Loss: A Randomized, Controlled Trial," *Annals of Internal Medicine*, 158.7 (2013), 505-514.

Long, Judith A, Erica C. Jahnle, Diane M. Richardson, George Loewenstein, and Kevin G. Volpp, "Peer Mentoring and Financial Incentives to Improve Glucose Control in African American Veterans," *Annals of Internal Medicine,* 156.6 (2012), 416-424.

Marsiske, Michael, and Jennifer A. Margrett, "Everyday Problem Solving and Decision Making," In *Handbook of the Psychology of Aging, 6th ed.*, James E. Birren and K. Warner Schaie eds. (New York, NY: Academic Press, 2006).

Mebane-Sims, Irma, "Alzheimer's Disease Facts and Figures," *Alzheimer's & Dementia*, 5.3 (2009), 234-270.

Nalbantian, Haig R. and Andrew Schotter. "Productivity Under Group Incentives: An Experimental Study." *American Economic Review*, 87.3 (1997), 314-341.

Owsley, Cynthia, Michael Sloan, Gerald McGwin, Jr., and Karlene Ball, "Timed Instrumental Activities of Daily Living Tasks: Relationship to Cognitive Function and Everyday Performance Assessments in Older Adults," *Gerontology*, 48.4 (2002), 254-265.

Park, Denise C., and Doris Payer, "Working Memory Across the Adult Lifespan," In *Lifespan Cognition: Mechanisms of Change*, Ellen Bialystok and Fergus I.M. Craik eds. (New York, NY: Oxford University Press, 2006).

Perez, Amanda Philyaw, Martha M. Phillips, Carol E. Cornell, Glen Mays, and Becky Adams, "Promoting Dietary Change Among State Health Employees in Arkansas Through a Worksite Wellness Program: The Healthy Employee Lifestyle Program (HELP)," *Preventing Chronic Disease*, 6.4 (2009), A123.

Rogers, Wendy A., and Arthur D. Fisk, "Understanding the Role of Attention in Cognitive Aging Research," In *Handbook of the Psychology of Aging, 6th ed.*, James E. Birren, and K. Warner Schaie eds. (New York, NY: Academic Press, 2006).

Appendix 3A

Table 3A.1: Mean Daily Exercises As a Function Of Partner's Lagged Exercises

Atomistic	10.62**
	[3.82]
Altruistic	2.64
	[4.38]
Cooperative/Competitive	8.73*
	[3.56]
Partner's exercises today (p_t)	0.01
	[0.05]
Atomistic* p_t	0.03
	[0.07]
Altruistic* p_t	0.13*
	[0.06]
Cooperative/Competitive* p_t	0.06
	[0.06]
Partner's exercises t-1 (p_{t-1})	0.06
	[0.03]
Atomistic*p_{t-1}	-0.05
	[0.05]
Altruistic*p_{t-1}	0.10
	[0.06]
Cooperative/Competitive* p_{t-1}	0.05
	[0.04]
Partner's exercises t-2 (p_{t-2})	0.06*
	[0.03]
Atomistic* p_{t-2}	0.003
	[0.05]
Altruistic* p_{t-2}	0.02
	[0.05]
Cooperative/Competitive* p_{t-2}	0.008
	[0.04]
Partner's exercises t-3 (p_{t-3})	0.091*
	[0.04]
Atomistic* p_{t-3}	-0.06
	[0.05]
Altruistic* p_{t-3}	0.00
	[0.05]
Cooperative/Competitive* p_{t-3}	-0.03
	[0.05]
Constant	9.01***
	[1.68]

Table 3A.1 (Continued): Mean Daily Exercises As a Function Of Partner's Lagged Exercises

Observations	12,168
R-squared	0.19

Notes:

1. This table examines how an individual's exercise completion relates to the current and previous exercise completion of their partner (other pair in the cooperative/competitive treatment). The table reports results of an OLS regression of exercises completed on indicators for experimental condition, the number of exercises completed by the individual's partner on the current day and three previous days, and interactions between the treatments and the lagged exercise completion.
2. The unit of observation is the participant-day.
3. Standard errors clustered at the level of the pair for all experimental groups except Cooperative/Competitive which is clustered at the level of the group.
4. *** Significant at the 1 percent level, ** Significant at the 5 percent level, * Significant at the 10 percent level.