How Important is Temptation Spending? Maybe Less than We Thought

Jason T. Kerwin, Qingxiao Li, and Lasse Brune*

October 18, 2019

Click here for the latest version of this paper

Abstract

Temptation plays a key role in theoretical work on spending and saving in developing countries. The limited empirical evidence on its importance, however, suggests that cash transfers do not induce increases in temptation spending. This paper expands the evidence base by studying the effect of randomized exposure to temptation on spending decisions. Consistent with the cash transfer literature, a more tempting environment does not induce significant changes in temptation spending. However, the magnitudes of both temptation spending levels and the treatment effects are somewhat sensitive to the definition of temptation spending used. We discuss the potential factors that may be driving these null results and conclude that future research should not expect a large role for temptation spending.

Keywords: Temptation Spending, Self-Control, Behavioral Economics, Development Economics

JEL Codes: D90, D91, O12

* Kerwin (corresponding author): Department of Applied Economics, University of Minnesota (jkerwin@umn.edu), Department of Applied Economics, University of Minnesota, 1994 Buford Ave, St. Paul, MN 55108; Li: Department of Applied Economics, University of Minnesota (lixx5376@umn.edu); Brune: Global Poverty Research Lab, Buffett Institute for Global Studies, Northwestern University (lasse.brune@northwestern.edu).. We are grateful to Marc Bellemare, Jeff Bloem, Audrey Dorélien, Annet John, Soomin Lee, Joe Ritter, Laura Schechter, and Dan Silverman for helpful comments. We are grateful for research support from the IPA/Yale Savings and Payments Research Fund (funded by the Bill and Melinda Gates Foundation), the University of Michigan Population Studies Center, and the Michigan Institute for Teaching and Research in Economics. Kerwin's work on this study was supported in part by an NIA training grant to the Population Studies Center at the University of Michigan (T32 AG000221), as well as by fellowship funding from the Rackham Graduate School. This study is registered with the AEA RCT Registry under registration number AEARCTR-0000437. All errors and omissions are our own.

Does temptation play a major role in the spending decisions of the poor? Prominent theories suggest that it does (Gul and Pesendorfer 2001; Fudenberg and Levine 2006; Banerjee and Mullainathan 2010; Ozdenoren et al. 2012), and policymakers are often concerned that participants will misspend cash transfers on temptation goods. "Temptation spending" typically refers to money "wasted" by people on things that policymakers would prefer they not buy (e.g. Harvey 2007, Ikiara 2009), or that the people themselves would prefer not to buy when asked at a different time as in O'Donoghue and Rabin (1999). Despite the important role of temptation in both theory and policy, empirical evidence suggests that temptation spending is either unchanged or reduced by cash transfers on temptation goods (Evans and Popova 2016). The disconnect between the theoretical literature and the evidence from cash transfers raises the question of how important temptation spending really is in the financial lives of people living in poor countries.

In this study, we address this puzzle by attempting to experimentally vary workers' exposure to temptation at the time that they receive cash payments, using a field experiment in southern Malawi. We do this by requiring some workers to pick up their pay during the major local market, which takes place weekly on market day and is commonly identified by members of our sample as a highly-tempting environment; the control group picked up their pay on the day before (a non-market day) at the same location. This market day treatment is cross-randomized against a second experiment that varied whether workers received their pay weekly or in a deferred lump sum (Brune and Kerwin, 2019), which allows us also to test whether the effects of the deferred lump sum payment differ by temptation exposure. The experiment holds transaction costs, such as time and transportation costs, constant.

Our results do not provide strong evidence in favor of the typical temptation spending narrative. Direct exposure to a tempting environment induces no appreciable changes in

expenditure, nor in temptation spending in particular; this pattern is similar for both the weekly and lump sum payments. We show that altering the timing of income receipt does not substantially change that finding: the minimum detectable effect sizes on temptation spending for both our lump sum treatment and our market day payment treatment are less than 0.25 SDs, letting us rule out larger effects with reasonable confidence. These findings align with the results of cash transfer studies, and also with previous research in Malawi, which has found that recipients of a large cash windfall spent little on temptation goods (Brune et al., 2017).

This study also contributes evidence on methods for measuring temptation spending. Most previous studies define temptation spending as spending on alcohol and tobacco (Banerjee and Duflo, 2007), and high-calorie savory and sweet foods (Aker, 2013; Dasso and Fernandez, 2013) are sometimes included as well. In our study, we collect rich data on temptation spending — defining it not only using the standard definitions from the literature, but also by allowing the respondents to identify categories of expenditure that they themselves see as problematic, and computing the share of all expenditure that is deemed to be temptation spending. We find that respondents' own designations of temptation goods can differ sharply from those that would be chosen by a researcher.

Based on these expanded definitions of temptation spending, we find that respondents report wasting non-trivial amounts of money — and we estimate that the magnitude of temptation spending is substantially larger (3 to 20 times larger depending on the definitions) using respondents' own self-designations than using the conventional approaches that focus on alcohol and tobacco. These alternate measures do not change our core finding: we cannot reject that spending decisions are unaffected by exposure to temptation. However, the magnitude of the point estimate is quite sensitive to the definition used, ranging from 3% of control-group temptation

levels up to 59%. In addition, we find significant gender heterogeneity, with men spending more when we use the preferred approach of temptation spending (goods respondent says they waste money on or are tempted to buy). Male respondents report MK114 (USD $0.70) more temptation spending than female respondents when they receive wages in a tempting environment.

We discuss seven potential reasons why the market day treatment may not produce substantial changes in spending and savings behaviors. The first is statistical power: the effects could be small enough that we do not have a sufficiently-large sample to detect them. The second is that the treatment may simply have been too weak, i.e. the market day environment was not actually tempting for the workers. Third, our experiment may have suffered from substitution bias: since there are other opportunities available for workers in our study to purchase temptation goods, the effects from our market day treatment may be limited. Fourth, under-reporting of temptation spending could attenuate any effects of the treatment. Fifth, workers may have successfully pre-committed to spending plans that prevent them from spending money on temptation goods. A sixth potential factor is peer effects: since workers show up together in paydays and interact with each other, their choices could mirror one another's leading to similar temptation spending decisions across study arms. The seventh reason is costly self-control: workers are able to resist the temptations posed by the market, but at cost in terms of utility or willpower. These models imply a diminishing ability to resist as temptations increase.

Out of the seven explanations mentioned above, the two most compelling are the role of pre-committed spending plans and substitution bias. Both of these mechanisms suggest that temptation has very limited practical impact on spending decisions. If prior commitments to spending plans reduce people's susceptibility to temptation, and many people have such plans, then the overall average effect of temptation will be small. Sufficiently-strong substitution bias, in

contrast, implies that tempting environments might not matter at all: people will seek out the same level of impulse purchases irrespective of their environmental exposure to temptation.

This paper contributes to the empirical literature on temptation spending by measuring the effects of a natural temptation exposure on overall temptation expenditures in a real-world setting. Previous research has shown strong effects of temptation in lab settings (Toussaert 2018). Sadoff et al. (2019) find large dynamic inconsistency effects and a strong demand for commitment, but focus on food choices alone, and examine choices out of a restricted set of foods. There is also evidence that paying people in cash (as opposed to a bank account), leads to large changes in consumption, but not to increases in temptation spending (Somville and Vandewalle 2018). While this previous work implies that temptation exposure is very important in economic decision-making, our findings suggest that it is not.

These results provide important insights into the potential role of temptation spending in the economic lives of the poor. Given our findings, we argue that researchers should expect that temptation exposure will have at most a weak effect on temptation spending. Men appear to be more responsive to tempting environments for certain definitions of temptation spending, so the gender targeting of cash transfers or other income may affect how much money is spent on temptation goods. Moreover, the measure of temptation spending used may mask effects in temptation spending studies. Alcohol and tobacco are the most used definition for temptation spending; however, not everyone drinks or smokes. Moreover, just because policymakers or consumers themselves want to reduce spending on a good like alcohol doesn't mean the good qualifies as a temptation good in the theoretical sense. In our sample, only 16% of people consume alcohol or tobacco, but more than 34% of people have spent on goods they consider a waste of money or are tempted to buy; the most-common such goods are gifts for their children, clothes,

and savory foods. Estimates of temptation effects that focus on alcohol and tobacco alone may be downward biased. Relying on individuals' own determination of which purchases are temptation spending may generate more-useful measures and while also giving people more agency over how their choices are evaluated and how policies are designed.

The remainder of the paper proceeds as follows. Section 1 describes the data we use and the design of the study, and Section 2 lays out our empirical strategy. In Section 3 we present the effects of exposure to a tempting environment on expenditure decisions. We discuss the mechanisms for our findings in Section 4. Section 5 concludes.

1 Data and Experimental Design

The data we use in this study comes from a field experiment that randomly assigned workers to receive their wages in environments with varying levels of temptation, as well as either in a smooth stream or a lump sum (Brune and Kerwin 2019). The wages were paid through an income support program organized by Mulanje Mountain Conservation Trust (MMCT), a local NGO in the Mulanje District of Malawi's Southern Region, which provides temporary informal employment opportunities during the agricultural offseason. While the workers in our sample have other sources of income, the wages received from this program are an important supplement to their livelihoods.

Two rounds of the experiment happened over a period of three months from November 2013 to January 2014. There were initially 350 workers from seven villages recruited into the study for round one, and an additional 15 workers were added for round two to replace the workers who dropped out after round one. Workers were selected for participation by their respective village development committees, which chose people largely on the basis of perceived disadvantage; thus

the sample is predominantly female and poorer than average for the region. Each worker worked for two weeks during each round of the program, and for about four days per week. The daily wage rate was MK400 (USD $2.50), which was at the national minimum wage level, and is approximately 160% of average daily spending for the workers in our sample. Workers were assigned to work on conservation-oriented activities that promoted the sustainable use of natural resources.

In our study, workers received their wages after the work was completed and were randomly assigned, independently by round, to receive the wages either on the market day (Saturday) or a non-market day (Friday). The total nominal income received by all workers was identical, and workers were informed about when they would be receiving their pay at the beginning of each round. Workers' pay schedules were fixed for each round, the procedure was explained verbally and workers were also provided with a simple handout explaining their schedule. To ensure transaction costs, such as transit and time costs, were held constant across wage payment modes, all workers were required to come to the payroll site on all paydays during each round — even when they were not being paid their wages. An MK100 show-up stipend, on top of any money workers were slated to receive, was provided to encourage attendance and defray workers' time cost.

This market day treatment was cross-randomized against another experiment that varied payment frequency: workers received their pay in four weekly installments or in a deferred lump-sum payment at the end of the month. The two variations in the timing of pay (the frequency of payments and the and temptation level of the environment when workers received the pay) were cross-randomized, creating four study arms in each round.

Table 1 presents the payment schedule in each round across the four payday weekends with

show-up stipends and wage disbursements per study arm. The market day and non-market day arms have an identical number of paydays in the lump-sum and weekly payment schemes. The total payment excluding the MK100 show-up stipend was MK2800 in round one and MK3200 in round two, because there were seven work days during the first round and eight days during the second round.

Workers in the study were randomly assigned to study arms in each round of the study, and the randomization for both rounds of the study was done prior to the baseline survey. The group assignments were not revealed to the workers until the beginning of each round of work. The randomization for the first round was stratified by village and gender, and the randomization for the second round was stratified on the round one assignment and village. We generally pool observations across rounds and the cross-randomized lump-sum treatment to improve statistical power.

We use three rounds of survey data: a baseline and a survey after each round of the study. The surveys after each round were conducted on the Monday immediately after the last payday of each round. The order in which workers were visited for the surveys was randomized by village, and workers were interviewed at their homes. The survey collected information on income, physical assets, saving, transfers, and details on the worker's expenditures since the first day of the final payday weekend. We also utilize brief survey questions asked of workers when they came to collect their pay; see Brune and Kerwin (2019) for a detailed discussion of the payroll survey data.

The random assignment produced a sample of workers that is balanced across study arms on observable characteristics. Table 2 shows balance tests for the main comparison we use in this paper, which is between all workers who were paid on Fridays (the control group) and all those

who were paid on Saturdays (the market-day payments treatment). We find no statistically-significant imbalance on any of the covariates in the table, and we also fail to reject the joint null hypothesis of zero difference on all covariates together (p=0.87).[1]

Our analytic sample is 70% female and 70% married; the average age is 40 and they have about 3.5 years of formal schooling on average. The average worker has received about MK3000 (USD $18.77) in cash income and spent about MK4000 (USD $25.02) since the previous Friday. Workers have received substantially more in loans than they have given out, and are also net beneficiaries of transfers. The average midline surveys took place 2.5 days after the last payday, and 74% of workers preferred the lump-sum wage payments.

1.1 Measures of Temptation Spending

Temptation spending has been defined in different ways in previous studies, and temptation goods are typically goods that are commonly perceived as harmful (Evans and Popova 2016). For instance, alcohol, tobacco, high-calorie savory foods, and sweets are commonly included in the definition of temptation spending. In general, temptation spending is defined as money "wasted" by the poor on things that policymakers would not prefer them to purchase.

This approach presumes that that perfectly competent adults cannot be trusted to make their own decisions, and that policymakers or people in other countries could do better on their behalf. At the same time, the poor very commonly identify categories of spending that they wish to reduce, and the expenditures that they — like most consumers — most often identify as problematic are alcohol and tobacco (Banerjee and Duflo, 2007).

[1] This table pools workers across rounds and across the lump-sum and weekly payment arms. We also find that the two study arms of interest are balanced when we analyze them separately by round (Appendix Table 1) as well as separately by weekly vs. lump-sum payments (Appendix Tables 2 and 3).

Motivated by Banerjee and Duflo's findings, we take an individualized, respondent-driven approach to categorizing temptation spending. We allow respondents to identify categories of expenditure that they themselves see as problematic, and compute the share of all expenditure that is deemed to be temptation spending. One promising approach is to ask people about goods that they are tempted into purchasing and match those categories to detailed survey data on actual expenditures. This allows us to classify purchases as temptation spending based on people's own perceptions of goods that are problematic purchases. Our household surveys include three different definitions of temptation goods: 1) purchases that the respondent commonly regrets after the fact; 2) goods that are commonly unplanned purchases; and 3) goods that the respondent is tempted into purchasing that they should not buy or that are wastes of money. For each respondent, we match the goods that they personally deem to be problematic with itemized lists of purchases they have made since the previous Friday, also from the household survey. We do this separately for the three definitions above.

We also use two other self-reports of temptation spending. The first is simply the respondents' own recall of the total amount of money they wasted. For the second, we ask, for every good in the itemized list, whether the purchase was planned beforehand, an approach first developed by Brune et al. (2017). Unplanned purchases are taken to be temptation spending in this case. The English translations of the exact survey questions we used for respondent self-reports of temptation spending are shown in Appendix Table 4. For three of the definitions, workers are asked to self-designate which goods count as temptation spending; we had options for the most common responses from pilot-testing the survey, and also an "other" category where workers could list up to three additional goods.[2] We present the frequency of each choice in Appendix

[2] Very few workers used all three "other" spaces on the survey.

Table 5. The most commonly-mentioned goods are gifts for children, clothes, and savory foods.

There is substantial heterogeneity across definitions: for example, gifts for children are frequently reported as something that respondents are tempted into buying or are wastes of money, but is not commonly regretted or unplanned. There is also substantial heterogeneity across individuals: "other" is the most common option for all three definitions, but the specific other purchases vary widely. Moreover, the specific other purchases mentioned show why definitions C and E may be problematic. Workers frequently regret buying expensive food (as opposed to food, which was mentioned separately), presumably because they feel they paid too much for it. That question also elicits mentions of bad business decisions and other expensive goods, likely for the same reason, whereas they do not appear under definition A and all and just once under definition E. Illnesses/medicine and funerals are unsurprisingly commonly unplanned purchases (definition E) but much less-commonly regretted or considered to be tempting. These patterns highlight the fact that purchases that are unplanned or regretted are not always temptation spending.

We supplement these subjective self-judgments of temptation goods with two objective measures drawn from the previous literature. First, following Evans and Popova (2016), we consider purchases of alcohol and tobacco to be temptation spending. Second, we use an expanded version of their definition, by including all goods that are mentioned as temptation goods in the studies they summarize and that also appear in our surveys' itemized lists of purchases. This adds donuts[3] and soda to their list. For each of these prescriptive definitions, we follow the same procedure described above – we match them to our itemized lists of purchases and compute total expenditures.

[3] Donuts are referred to in Chichewa as *mandasi*, which is sometimes also translated as fritters. They are typically made from sweet bread but not covered in sugar or frosting like American donuts.

Table 3 presents summary statistics and correlations between the various definitions of temptation spending. The recorded level of temptation spending varies substantially based on the definition we employ. Moreover, the various measures are only weakly correlated with one another: the only correlation coefficient that exceeds 0.25 is between "Alcohol and Tobacco" (Row F) and "Alcohol, Tobacco, Donuts, and Soda" (Row G) — an artifact of the partially-overlapping definitions.

Our preferred measures of temptation spending are purchases of goods the workers say they often waste money on or are tempted to buy (Row A, "Waste/Temptation") and self-reported aggregate money wasted (Row B, "Money wasted"). That is because, first, regretted purchases (Row C), unplanned purchases (Row D), and purchases that deviate from one's plans (Row E) often capture other mistakes and deviations from plans that are not conceptually equivalent to being tempted into wasting money. As discussed above, our workers often report regrets due to price fluctuations or quality — they recognize ex post that they overpaid for something. Unplanned purchases can result from a similar pattern: if something is available at a bargain price then people may deviate from plans and purchase it, but this is the result of re-optimization, not a mistake. Unplanned purchases also include medical expenses and funeral costs, which are not generally temptation goods.

Second, the common researcher-imposed definitions of temptation spending (Rows F and G) miss important categories of goods that the workers in our sample report being tempted into purchasing. These include fried meat and other savory foods (which is often available from vendors during market days) and clothing (both for personal use or as gifts to family members).

Our two preferred measures show non-trivial average levels of temptation spending — 3% of average income for Row A and 10% for Row B — and also have higher variances than the other

measures. Although we think the definitions on rows A and B are the best measures of temptation spending, we utilize all seven definitions to limit researcher degrees of freedom.[4] We report our main analyses for each definition, and focus primarily on a combined index of temptation spending. We do this by taking the first principal component of the seven individual temptation measures for the control (weekly payment) group, constructing predicted values for the entire sample, and normalizing to the control group. Since one of the seven measures (total money wasted) was collected only in round 2 of the study, we construct the index two ways: one that includes all seven outcomes but is only computed for round 2, and one that excludes the "total money wasted" variable and is computed for both rounds.

1.2 Market Days as a Source of Temptation

We chose market days as the tempting environment for our study based on extensive qualitative and descriptive work with people in the local area. Prior to running the experiment, we did open-ended interviews with people from the local area to ask them about which situations they find tempting. Based on their responses, we chose market days as a potentially-tempting environment and conducted a pilot test of the experiment at a market near our study site to refine our field procedures and ensure that the experiment was feasible. Participants in that pilot reported that they found the market highly tempting. We also collected data on perceptions of tempting environments from members of our study sample, which are presented in Appendix Table 5. As we discuss below in Section 4, workers find market days quite tempting.

Weekly market days are common across rural Africa. Markets in Malawi are held at trading

[4] Simmons et al. (2011) discuss how researcher choices over things like variable definitions can lead to false-positive findings.

centers that contain a few fixed businesses and have a large number of spaces for other vendors to come in and sell additional goods on the market day. In the local area where we ran the experiment, there are seven of these trading centers, and typically each one holds two market days per week. Market days are often the only feasible option for people living in rural Malawi to buy common consumption goods. These days tend to offer a fairly stark contrast to ordinary days in rural Malawi. They are typically lively, noisy affairs with many goods on offer, presenting environments that try to tempt consumers into spending their money. Anecdotally, people in Mulanje District often describe market days as tempting situations, in which excitement can cause them to purchase things they would rather not.

Our survey data (Appendix Table 6) confirms that people find markets tempting: for a free-response question about situations that are tempting or in which respondents may waste money, 37% of all respondents volunteered market days as a tempting situation, by far the most common response (Panel A). Multiple-choice questions (Panel B) show the same pattern: 69% of people said that market days are more tempting than the day before market days, and 66% of people said having a lot of cash on hand at the trading center was more tempting than having it on hand elsewhere.

These answers suggest that payments during market days could exacerbate temptation-based psychological savings constraints, by inducing people to spend money on tempting goods that they would prefer to save. Panel D confirms that markets are an important part of life in the area, with the typical person reporting they went to the market six times in the past month. Saturdays are the most common days that people visit the market (32% of all visits), although other trading centers do hold market days on Fridays and so 26% of visits happen on Fridays.

We compare payments during the market day to payments at the same site the day before,

when the market does not take place. We chose the day before — Friday — as the alternate day for several reasons. First, it was logistically simpler to manage payments on two consecutive days than on non-adjacent ones; Sunday was not an option because the vast majority of our sample goes to church on Sunday mornings. Second, using the day before the market ensured that all respondents had the liquid cash needed to make purchases at the market — if we had paid the control group on a later day, then for the first week they would not have had any money to spend at the market on Saturday. Third, and most important, if the control group was paid after the Saturday group, then any differences in temptation could simply be a function of having the money for a shorter period. The control group does not collect their wages in the tempting market day environment, thus the money will not "burn a hole in their pocket" in the sense of Fudenberg and Levine (2006) unless they keep the money and come back to the market again the other day or find another market to spend the wage right away.

The location and timing of the payroll was specifically chosen to maximize the likelihood that people would be exposed to temptation goods. The market at Mwanamulanje happens only on Wednesdays and Saturdays (with Saturdays having the larger market out of the two days), and principally in the morning, which is when people were paid. Shops are still open on Fridays, and there are some mobile vendors, but the majority of market activity happens on Saturdays.

2 Econometric Strategy

To estimate the mean effects of the exposure to a tempting environment on expenditure and temptation spending, we estimate regressions of the following form:

$$Y_{ir} = \alpha + \beta T_{ir} + \boldsymbol{\gamma}'\mathbf{X}_{ir} + \varepsilon_{ir} \quad (1)$$

where *i* denotes worker and r denotes the round of survey. The outcome of interest for worker *i* in

round r is Y_{ir}. The treatment variable T_{ir} equals 1 if the worker receives wages in a tempting environment (the market day) and 0 otherwise. The vector $\mathbf{X}_{ir}$ is a set of controls, which comprise stratification cell dummies, two household financial variables[5], indicators for the day-of-week of the exogenously-assigned (first attempted) interview date, and the baseline values of the outcome variable. Finally, ε_{ir} is a mean-zero error term.

We cluster standard errors at the worker level when we use pooled data from both rounds to account for the statistical dependence of outcome measures for the same worker across two rounds. This means our standard errors are arguably conservative, since treatment status is randomized within-worker (Abadie et al. 2017). The stratification cells are defined separately by round to control for round fixed effects.

We supplement our main regressions with heterogeneous treatment effect analyses. We do this by examining how the treatment effect varies by various baseline characteristics, such as gender and whether the worker's household has children enrolled in school. We construct indicators for each level of the baseline characteristic W_i. For example, if W_i is gender, W_{0i} is an indicator for being male and W_{1i} is an indicator for being female. We then run the following regression:

$$Y_{ir} = \alpha + \theta_0 T_{ir} W_{0i} + \theta_1 T_{ir} W_{1i} + \delta W_{0i} + \boldsymbol{\gamma}' \mathbf{X}_{ir} + \varepsilon_{ir} \quad (2)$$

where we omit the main effect of the indicator for the highest level (female in this example). Here θ_0 is the treatment effect for males and θ_1 is the treatment effect for females. We also present formal tests of the null hypothesis that the two treatment effects are equal. Given the many tests we conduct in these heterogeneous treatment effect analyses (and the many other possible tests we

[5] The household financial variables are an index of physical asset and livestock ownership using principal component analysis (PCA) and total spending out of income received since the past Friday. Both variables were measured before the randomized assignment.

could conduct) our results should be interpreted with caution due to potential multiple testing issues.

3 Effects of Exposure to a Tempting Environment

Table 4 shows the effect of the market day payment treatment on spending. Panel A shows pooled results across both workers paid weekly and workers paid in deferred lump sums. Panel B presents the results using only the weekly payment group. Panel C examines treatment effects just for workers in the lump-sum treatment group. There are no substantive differences across the three approaches, so our discussion focuses on Panel A.

The market day payment treatment leads to large shifts in the exact timing of expenditure. Column 1 shows that combined spending on Friday and Saturday drops by MK756 (USD $4.73) for the market day treatment group relative to the workers who are paid on non-market day. In the presence of liquidity constraints, this is to be expected: workers paid on Friday have had an additional day to spend their income. Taking into account the difference in income timing, the market day treatment induces no meaningful changes in total expenditure: workers spend a similar amount immediately upon receiving their income (Columns 2 and 3) and have statistically indistinguishable total income, remaining cash holdings, and total spending between the previous Friday and the survey date.

The next logical question is whether exposing workers to a tempting environment – the market day – induced any changes in temptation spending in particular. Table 5 shows the results of this analysis. It shows that the market day payment treatment does not substantially change temptation spending for any of the temptation spending measures that we use. The estimates in all columns are statistically insignificant, and the signs for the effect of market day payment on

temptation spending measures are mostly negative. The only positive treatment effect estimates for temptation spending measures are regretted purchases (which rises by 10%) and alcohol and tobacco (which rises by 23%). The other temptation spending measures show negative effects ranging from 3% to 21% of the control-group mean. We focus on the PCA index measure (which uses data from both rounds and thus omits self-reported money wasted) which prevents issues arising from researcher degrees of freedom. It allows us to rule out all but the smallest treatment effects: the upper bound of the 95% confidence interval for the combined index across both rounds is 0.15 SDs.

We can see from Appendix Table 7 that on average men have higher temptation spending levels than women, with the exception of self-reported total money "wasted". Thus, we test for treatment effect heterogeneity by gender. Table 6 shows that there is some evidence that men are more affected by the treatment based on our preferred measure of temptation spending (goods respondent says they waste money on or are tempted to buy). However, this gender heterogeneity is very limited in both spending levels and treatment effects. The treatment effect for women is insignificant, and men report MK114 (USD $0.70) more temptation spending than female respondents when they receive wages in a tempting environment.

Our results imply a limited average treatment effect of exposure to a tempting environment on spending among our sample of workers. In addition, the market day payment treatment does not have appreciable impacts on temptation spending, irrespective of the choice of definition.

4 Mechanisms for the Null Effect of Temptation

The finding of the null effect of temptation is consistent with previous research on cash transfers (Evans and Popova 2016). However, it runs contrary to a prominent strain of theoretical

work (Gul and Pesendorfer 2001; Fudenberg and Levine 2006; Ozdenoren, Salant, and Silverman 2012; Banerjee and Mullainathan 2010) which argues that temptation plays key role in the spending of the poor. Moreover, the previous evidence simply shows that very little of cash transfers is spent on temptation goods; we find that even direct exposure to temptation seems to lead to little change in behavior. Here we discuss the potential mechanisms that may provide explanation to the null results in our study.

Statistical power

A first possibility is that our intervention might have had meaningful effects that are simply too small for us to detect. The fourth row of Table 5 presents the minimal detectable effect (MDE) size on temptation spending at 80% power[6] based on our estimated standard errors. We have 80% power to detect 0.17 to 0.28 SDs changes in temptation spending, and our MDEs for our preferred PCA index of temptation is 0.18 SDs. For our preferred temptation spending measures, we find MDEs of 0.19 SDs for goods respondents say they waste money on or are tempted to buy and 0.28 SDs for self-reported aggregate money "wasted", with similar values for the other goods. These values correspond to 54% and 61% changes relative to the control-group mean. One estimate of the effect of exposure to temptation on temptation spending comes from Wansink et al. (2006), who varied whether a candy dish was located next to or further from office workers.[7] They find increases in candy consumption of nearly 100%. While this is arguably a stronger intervention than the one we conducted, our study is reasonably well-powered based on their estimated treatment effects.

[6] The minimal detectable effect at 80% power is 2.8 times the standard error divided by the control group standard deviation (Ioannidis et al. 2017).

[7] Some of Wansink's research has been retracted due to poor scientific practice and data that does not match the published results. However, we are aware of no issues that have been raised with this specific paper.

An alternative way of assessing our statistical power follows an argument made by Evans and Popova (2016), who point out that cash transfer studies which find null effects on temptation spending do find significant treatment effects on other outcomes. This suggests that power should be less of a concern. The same logic applies to our study: while we find null effects on temptation spending, the intervention did have large and statistically-significant effects on the exact timing of expenditure (Table 4, column 1). We also find that our MDEs are within the range of MDEs for 80% power in the studies reviewed in Evans and Popova (2016), which they assess to be reasonably well-powered, albeit for a different intervention (they study changes in the level of income received while we study changes in the timing of income receipt). Overall we conclude that limited statistical power is not the main issue driving our null results: our data would let us detect treatment effects that are large enough to be of interest and consistent with the literature.

Treatment intensity

Along similar lines, our treatment simply might not have been intense enough to produce appreciable changes in behavior. The available evidence suggests this is unlikely. Appendix Table 6 presents survey evidence from our sample of workers on the tempting nature of market days. Market days are the situation workers most commonly report as one in which they waste money or are tempting into spending (Panel A). Out of the workers who report any situation as being tempting, 61% choose market, and this number is by far the most common response (about four times more-frequent than the second-most-common selection). Panel B shows that large majorities of respondents find the market day more tempting than the day before the market day (69%) and are more tempted to spend money they will later regret when they have cash in their pocket at a trading center as opposed to elsewhere (66%). In addition, the market day is considered much more

tempting than the night before (which is a common night for drinking in our setting) 74% of our sample found the market day to be more tempting. Temptation spending is also self-perceived to be a major driver of waste: 42% of workers report it as a reason they waste money (Panel C).

Substitution bias

One specific reason our treatment might have been weaker than expected is substitution bias – while we exposed workers to an environment that was fairly tempting, control-group workers could have chosen to substitute toward other temptation spending opportunities.[8] Our treatment was designed around the market at Mwanamulanje Trading Centre, which operates on Saturdays (with a smaller market day on Wednesdays). However, there are a number of other nearby trading centers that do have market days on Fridays, which was the alternate day on which workers received their pay. It is possible that the workers who are assigned to a low-temptation environment on payday (Friday at Mwanamulanje) simply substitute toward other sources of temptation, such as the market days happening elsewhere. Evidence of this potential substitution is found in our baseline survey data, presented in Appendix Table 6: while Saturdays are the most-common day that members of our sample go to the market prior to the experiment, Fridays are nearly as common (26% vs. 32% of all market visits), and 42% of all visits happen on a different day of the week entirely. The Mwanamulanje market does not occur on Fridays, so this indicates that workers are frequently visiting other markets for market day.

The likelihood of this possibility is mitigated somewhat by the fact that the control-group workers, who were paid on Fridays, were required to come to the market on Saturday even though

[8] Here we are using "substitution bias" in the sense of Heckman et al. (2000), where treatment effect estimates from randomized experiments are attenuated because the control group gets access to the treatment or to a close substitute for it. "Substitution bias" also has a separate definition in the literature on price indices.

they were not being paid their wages. This reduces the time available for workers to seek out other markets, making it less likely for workers to seek out alternative market to consume the day before.

At the same time, the mandatory attendance at all paydays could have led to another form of substitution, over time instead of across space: if workers primarily save their earnings by holding cash on their persons, then workers who are paid on non-market day may simply hold onto the cash and face the same temptations as those paid on market day.[9] This explanation helps reconcile our results with the workers' own evaluations of markets as being extremely tempting, and with the fairly high levels of temptation spending we observe (1-10% of total expenditure, depending on the definition we use). At the same time, if people substitute toward other temptation spending opportunities, then temptation spending is conceptually quite different from how it is typically conceived in economic models. It is hard to reconcile the active seeking of temptations with dual-self style theoretical frameworks in which temptations are valued only by the instantaneous, current self.

Under-reporting

Another reason for the small measured effects of the treatment is under-reporting of temptation spending. There is evidence that socially-undesirable behaviors are misreported in surveys (Mathiowetz et al. 2002), and it is very common for people to under-report spending on alcohol and tobacco. For instance, cigarette smoking is significantly underreported compared with cigarette sales figures (Warner 1978) and survey reports of alcohol use are less than half of retail sales in the United States (Cook 2007). It is unlikely that under-reporting would be systematically

[9] The survey did not distinguish between savings kept at home and carried on one's person, so we cannot assess how common this is. Anecdotally, however, the workers in our sample commonly carry at least some of their savings on their person.

related to treatment status in our setting: respondents did not know the intent of the study and had no incentive to alter their responses based on their treatment status. Still, if temptation spending is sufficiently under-reported across the board, this would cause our coefficient estimates to converge to zero. This is also implausible: some of our measures of temptation spending comprise 10% of overall money spent.

In addition, our preferred measure of temptation allows respondents to self-designate spending as problematic (Rows A and B of Table 5), and is this less focused on purchases that respondents would be likely to under-report. We observe higher spending on this temptation spending measure than we do on alcohol, tobacco, donuts, and soda. This mitigates the concerns that our results are driven primarily by under-reporting.

Pre-committed spending plans

Alternatively, the small average treatment effects could mask important heterogeneity. Both models of temptation (e.g. Gul and Pesendorfer 2001) and dual-self models of self-control (e.g. Fudenberg and Levine 2006) imply that if workers are aware of their self-control problems, they should demand commitment devices. If workers in our sample are aware of the temptation of the market, they need not actually succumb to temptation; instead, they can find ways to constrain their behavior through commitment devices (Bryan et al. 2010). In particular, since workers know their wage payment schedule, it is possible that workers may pre-commit to spending plans ahead of time, which would reduce the scope for temptation spending. This could be done in two ways: workers might have made promises to friends or family members or agreements with vendors, or workers might have mental accounts (Thaler 1985) that drive them to spend the money in the planned-upon fashion. Both mechanisms could constrain workers' spending decisions and mute

any temptation spending effects.

We cannot directly test the mental account mechanism, but we do have a potential empirical test of the other mechanism (binding promises to friends or family members). As a proxy for the extent of binding promises to kin, we use the presence of children in school: workers with children in school are likely to have more pre-planned expenditures such as school fees.[10] The results in Table 7 show that there are stark differences in treatment effects between respondents with and without children in school. Specifically, respondents with children in school experience zero treatment effects on any outcome, while respondents with no children in school have substantial treatment effects for our main PCA index of temptation spending. They experience a 0.33 SD increase in temptation spending as a result of the treatment (Column 1), and this effect is significantly different from zero at the 0.10 level. The treatment effects for workers with and without children in school are also significantly different from one another at the 0.05 level.

The differences in treatment effects are driven by our most-preferred measure of temptation spending, which allows respondents to self-designate temptation goods (Column 3). This rises by MK115 for workers without children in school, which is nearly 100% of the control-group mean. However, we see meaningful effects on just one of the other six measures of temptation spending. There is also no evidence of an effect on the alternate index, which includes self-reported money wasted and thus can be constructed only for round 2; we have somewhat lower power for this test due to the smaller sample size. These results imply that pre-committed spending plans could be an important factor that limits the effect of temptation on the spending decisions of the poor. They also highlight the importance of how we measure temptation spending: the results differ

[10] In principle, parents with children in school could have more exposure to temptation through requests from their kids. However, the exogenous variation temptation we are using – exposure to the market – should generally affect only the workers themselves and not their children.

substantially depending on which outcome we choose, and if we relied on the standard temptation spending measures that focus on alcohol and tobacco, we would miss this pattern of heterogeneity entirely.

Peer effects

Another factor that could mask the effects of the treatment is peer effects: workers picked up their pay in a queue with all other workers, including those who were not paid (but who still appeared to receive a small attendance incentive). Interactions with these peers might lead all workers to make similar temptation spending decisions. For example, workers might all go buy beers or snacks together, or workers exposed to temptation might instead follow their peers and not spend money on tempting goods. Evidence on peer effects in temptation spending suggests that the spillovers are most likely to be positive (Chuang 2016). This would cause the workers to behave more similarly to one another, leading to attenuated treatment effects. Peer effects have previously been documented in this context: Brune, Chyn, and Kerwin (2019) find evidence of workplace peer effects at an agricultural firm in the same district of Malawi.

Our study design allows us to examine one potential source of peer effects, via the order in which workers queued up to receive their pay. Workers had assigned ID numbers, and the sign-in sheet was in order by number. To speed up the payroll process, the workers typically queued in the order of their names on the payroll sheet, which was sorted by village and then alphabetically by last name. Thus workers were exposed to the same neighbors in line throughout the study, and those neighbors had randomly-assigned treatment statuses. We can estimate peer effects from line neighbors by including the average treatment status of the workers ahead of and behind you in line as an additional variable in our regression equation:

$$Y_{ir} = \alpha + \beta T_{\mathrm{ir}} + \delta \bar{T}_{-\mathrm{i,r}} + \boldsymbol{\gamma}' \mathbf{X}_{ir} + \varepsilon_{ir} \quad (3)$$

where $\bar{T}_{-\mathrm{i,r}} = (T_{(\mathrm{i}-1),\mathrm{r}} + T_{(\mathrm{i}+1)\mathrm{r}})/2$ is the average treatment status[11] of the workers *i-1* and *i+1* in round *r*, and takes the values 0, 0.5, or 1.[12]

We find no evidence that peer effects are driving our substantive results. The estimates of equation (3) in Table 8 reveal no statistically-significant effect of peers' treatment status on workers' temptation spending decisions. More importantly, the point estimates for the effect of the workers' own treatment status are essentially unchanged relative to the estimates of equation (1) from Table 5. An important limitation of this analysis is that it relies on the assumption that, if peer effects exist, they operate at least in part through the workers one interacts with in line. We have no measures of other social networks such as workers' friends, extended families, or neighbors. A further limitation is that this test does not cover the case where peer pressure leads to no responses to the treatment whatsoever. If that is true then workers would not respond to their peers' treatment status.

Costly resistance to temptation

One prediction of some models of temptation is that people can resist actually purchasing tempting goods, but must pay a utility cost to do so (Gul and Pesendorfer 2001). This implies workers in our study may have been able to forgo the tempting items they faced during the market day. Models of finite willpower (Ozdenoren et al. 2012) suggest that people can overcome temptations by drawing on a limited well of self-restraint in order to control their impulsive

[11] To keep the first and last members of the line in the sample, we use the treatment status of their only neighbor. The results are also robust using two workers ahead of and two workers behind in the line (Appendix Table 8).
[12] This test is not affected by the exclusion bias problem of Caeyers and Fafchamps (2016) because treatment status is randomly assigned and thus independent of any worker characteristics.

behavior. This theory predicts that people who exert greater self-control in consumption problem will exhibit less self-control in subsequent activities. A similar implication holds if the worker simply pays a utility cost and the cost is convex.

A partial empirical test of this prediction is possible. Workers who have more other exposure to temptation should be less able to resist the temptation of the market day, and vice versa. If we use temptation spending levels as a proxy for temptation exposure, this implies that the treatment should have higher effects on workers at the top of the distribution of temptation spending. Under the assumption of rank preservation, we can test this via quantile regressions. Appendix Figure 1 shows the results for our preferred index of temptation spending, which uses data from both rounds of the experiment. We see no evidence of systematic differences in treatment effects across quantiles of the outcome distribution: none of the quantiles show a statistically-significant effect and the point estimates fluctuate between positive and negative. While this is not a high-powered test, it does not provide evidence that costly self-control plays a role in driving our results.

5 Conclusion

This paper examines the importance of temptation on spending by studying the effect of exposure to temptation on spending decisions. Our data comes from a randomized field experiment in Malawi that varied the temptation level of the environment in which workers received cash payments by having some workers receive their pay during the major local market day, which is identified as the most-tempting local environment.

Our experiment suggests that the exact context in which workers are paid may not be an important consideration for designing payment systems. We find results that are consistent with,

and extend, the findings from the literature on cash transfer studies: that literature shows that receiving additional income does not raise temptation spending, while our results show that direct the specific timing and environment of income receipt does not raise temptation spending.

We discuss a range of potential factors that lead to the null effects of exposure to temptation on spending decisions. The evidence is the strongest for two: pre-committed spending plans and substitution. Pre-committed spending plans could act as a constraint that prevents people from buying temptation goods. Consistent with this, we see significant treatment effects on workers with no children in school, and zero effects for workers who have children in school and thus must save up for school fees and other school costs like uniforms.[13] Substitution bias in this setting implies that people look for same level of impulse spending irrespective of variations in environmental exposure to temptation. This could happen across space or over time: there were other markets that workers could have opted to visit to buy temptation goods on the same day, and untreated workers could also have hung onto their payments and brought them to the market the next day.

The results in this paper suggest that researchers and policymakers should be less concerned about the importance of exposure to temptation in driving consumption and savings choices in sub-Saharan Africa. While substitution bias may have driven our null results, the setting of our study is fairly typical for many rural areas in Malawi and other countries in sub-Saharan Africa, where there are very often trading centers whose market days cover most days of the week, located within distances that can be traveled in reasonable times. Even people paid on a day when there is no market happening anywhere could also choose to hold onto their cash and spend it on

[13] Both rounds of the study took place close to the beginning of the academic year, when such expenditures would need to be made.

temptation goods when the next market day happens. There is also some evidence that pre-committed spending plans may limit the effect of temptation on spending choices, although this comes from just one of many potential treatment effect heterogeneity analyses. Even if this finding helps explain our results, such pre-committed plans are very common in developing countries. People need to save for school fees, for farm inputs, and for investments in their houses; if such spending plans are an effective line of defense against temptation then it cannot drive changes in spending for any significant fraction of the population. Thus, the findings of our study suggest that the specific location or day of income receipt is not a major driver of spending decisions in a broad range of settings in rural Africa.

We also show the deficits of some temptation spending measures and recommend measurements that use respondents' own determination of which purchases are temptation spending to estimate the temptation effects. Spending on tempting goods such as alcohol and tobacco are widely used in studies as the definitions of temptation spending. However, these commonly-used definitions may conceal important patterns. People may be tempted to spend on goods other than these conventional ones, which would result in downward-biased estimates of temptation levels as well as treatment effects. We show that respondents' own designations of temptation goods can be quite different from the conventional definitions and could provide a better estimate of actual temptation spending. Our preferred measure of temptation spending is eight times higher than the level of spending on tobacco and alcohol – and 11 times higher for women. The crucial heterogeneity we document, where treatment effects are evident for workers without children in school but not for other workers, is not present if we restrict our analysis to alcohol and tobacco.

A few factors limit the strength of the conclusions we can draw from our study. While the evidence best supports two of the mechanisms, we cannot convincingly rule out the other four. Our sample is reasonably representative of low-income households in the local region that we study; it is possible that the treatment effects would differ for other populations of people in Malawi or in other parts of sub-Saharan Africa or the developing world. However, there is little evidence of treatment effect heterogeneity within our sample by any observable other than children enrolled in school, which is consistent with a role for pre-committed spending plans. This mitigates concerns about external validity somewhat. While replicating this study in other contexts would be valuable, it would also be worthwhile for future research to examine other potential drivers of self-control problems aside from temptation spending.

References

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. (2017). *When Should You Adjust Standard Errors for Clustering?* National Bureau of Economic Research.

Aker, J. C. (2017). Comparing Cash and Voucher Transfers in a Humanitarian Context: Evidence from the Democratic Republic of Congo. *The World Bank Economic Review*, 31(1), 44-70.

Banerjee, A. V., & Duflo, E. (2007). The Economic Lives of the Poor. *Journal of Economic Perspectives*, *21*(1), 141–168.

Banerjee, A. V., & Mullainathan, S. (2010). The Shape of Temptation: Implications for the Economic Lives of the Poor. BREAD Working Paper No. 257, Cambridge, MA.

Brune, L., Giné, X., Goldberg, J., & Yang, D. (2017). Savings Defaults and Payment Delays for Cash Transfers: Field Experimental Evidence from Malawi. *Journal of Development Economics*, *129*, 1–13.

Brune, L., Chyn, E., & Kerwin, J. T. (2019). *Peers and Motivation at Work: Evidence from a Firm Experiment in Malawi* (Working Paper).

Brune, L., & Kerwin, J. T. (2019). Income Timing and Liquidity Constraints: Evidence from a Randomized Field Experiment. *Journal of Development Economics*, *138*, 294–308.

Bryan, G., Karlan, D., & Nelson, S. (2010). Commitment Devices. *Annual Review of Economics*, *2*(1), 671–698.

Caeyers, B., & Fafchamps, M. (2019). Exclusion Bias in the Estimation of Peer Effects. NBER Working Paper, No. 22565.

Chuang, Y. (2018). *Self control or social control Peer effects on temptation consumption*. Working Paper.

Cook, P. (2007). *Paying the tab: The costs and benefits of alcohol control*. Princeton University Press.

Dasso, R., & Fernandez, F. (2014). *Temptation Goods and Conditional Cash Transfers in Peru*. Working Paper.

Dupas, P., & Robinson, J. (2013). Why Don't the Poor Save More? Evidence from Health Savings Experiments. *American Economic Review*, *103*(4), 1138–1171.

Evans, D. K., & Popova, A. (2017). Cash Transfers and Temptation Goods. *Economic Development and Cultural Change*, *65*(2), 189–221.

Evans, W. N., & Moore, T. J. (2011). The short-term mortality consequences of income receipt. *Journal of Public Economics*, *95*(11–12), 1410–1424.

Fudenberg, D., & Levine, D. K. (2006). A Dual-Self Model of Impulse Control. *American Economic Review*, *96*(5), 1449–1476.

Gul, F., & Pesendorfer, W. (2001). Temptation and Self-Control. *Econometrica*, *69*(6), 1403–1435.

Harvey, P. (2007). Cash-based Responses in Emergencies. *IDS Bulletin*, *38*(3), 79–81.

Heckman, J., Hohmann, N., Smith, J., & Khoo, M. (2000). Substitution and Dropout Bias in Social Experiments: A Study of an Influential Social Experiment. *The Quarterly Journal of Economics*, *115*(2), 651–694.

Ikiara, G. K. (2009). *Political Economy of Cash Transfers in Kenya* (p. 28). Overseas Development Institute.

Ioannidis, J. P. A., Stanley, T. D., & Doucouliagos, H. (2017). The Power of Bias in Economics Research. *The Economic Journal*, *127*(605), F236–F265.

Mathiowetz, N. A., Brown, C., & Bound, J. (2002). Measurement Error in Surveys of the Low-Income Population. In *Studies of welfare populations: Data collection and research issues* (pp. 157–195). National Academy Press.

O'Donoghue, T., & Rabin, M. (1999). Doing It Now or Later. *American Economic Review*, *89*(1), 103–124.

Ozdenoren, E., Salant, S. W., & Silverman, D. (2012). Willpower and the optimal control of visceral urges. *Journal of the European Economic Association*, *10*(2), 342–368.

Sadoff, S., Samek, A., & Sprenger, C. (2019). Dynamic Inconsistency in Food Choice: Experimental Evidence from Two Food Deserts. *The Review of Economic Studies*, forthcoming.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-Positive Psychology Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychological Science*, *22*(11), 1359–1366.

Somville, V., & Vandewalle, L. (2018). Saving by Default: Evidence from a Field Experiment in Rural India. *American Economic Journal: Applied Economics*, *10*(3), 39–66.

Spears, D. (2011). Economic Decision-Making in Poverty Depletes Behavioral Control. *The B.E. Journal of Economic Analysis & Policy*, *11*(1).

Thaler, R. (1985). Mental Accounting and Consumer Choice. *Marketing Science*, *4*(3), 199–214.

Toussaert, S. (2018). Eliciting Temptation and Self-Control Through Menu Choices: A Lab Experiment. *Econometrica*, *86*(3), 859–889.

Wansink, B., Painter, J. E., & Lee, Y.-K. (2006). The office candy dish: Proximity's influence on estimated and actual consumption. *International Journal of Obesity*, *30*(5), 871–875.

Warner, K. E. (1978). Possible Increases in the Underreporting of Cigarette Consumption. *Journal of the American Statistical Association*, *73*(362), 314–318.

White, J. S., & Basu, S. (2016). Does the benefits schedule of cash assistance programs affect the purchase of temptation goods? Evidence from Peru. *Journal of Health Economics*, *46*, 70–89.

Table 1

Average Pay Schedule Across Rounds 1 and 2

Average Payment amounts (MK)

		Weekend #1		Weekend #2		Weekend #3		Weekend #4							
	N	Fri	Sat	Fri	Sat	Fri	Sat	Fri	Sat	Sun	Mon	Tue	Wed	Thu	Fri
Weekly Payments, Fridays	169	850	100	850	100	850	100	850	100		Days of survey visits				
Weekly Payments, Saturdays	175	100	850	100	850	100	850	100	850						
Lump Sum Payment, Fridays	177	100	100	100	100	100	100	3100	100						
Lump Sum Payment, Saturdays	168	100	100	100	100	100	100	100	3100						
	Σ=689							Days covered by survey questions							

Notes: This table presents the average payment schedule (combining rounds 1 and 2) across the four payday weekends with show-up stipends and wage disbursements per study arm; we round the payments to the nears MK10. Sample includes 359 respondents who participated in at least one round of the work program and have data from at least one data source for that round (either the payday data, the survey, or both). All money amounts are in Malawian Kwacha (MK); during the study period the market exchange rate was approximately MK400 to the US dollar, and the PPP exchange rate was approximately MK160 to the US dollar.

Table 2

Balance of Background Characteristics and Financial Outcomes

	Friday Payment Group			Saturday Payment Group			Balance test *p*-value (levels)	Balance test *p*-value (logs)†
	Mean	SD	N	Mean	SD	N		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Background characteristics								
Male	0.30	0.46	351	0.34	0.47	351	0.376	--
Married	0.70	0.46	347	0.70	0.46	347	0.879	--
Age (Years)	39.96	15.75	351	39.80	14.92	351	0.890	--
Years of Education Completed	3.42	3.23	349	3.62	3.09	349	0.408	--
Midline survey date (days after Sunday)	2.55	1.15	346	2.52	1.14	346	0.701	--
Prefers lump sum wage payments	0.73	0.45	350	0.75	0.44	350	0.520	--
Financial outcomes (in units of MK unless noted)								
Income received since past Friday	3,237	9,110	351	2,766	4,821	349	0.503	0.826
Remaining cash holdings out of income received	572.01	1765.49	351	713.97	3135.19	349	0.555	0.159
Total spending since Friday	3,991	4,867	351	3,422	3,739	349	0.144	0.137
Asset Ownership (PCA)	-0.08	2.49	351	0.08	2.87	349	0.410	--
Loans received in past month	2,610	7,797	351	3,285	13,596	349	0.524	0.283
Loans made in past month	682	2,700	351	784	3,209	349	0.571	0.678
Transfers received in past month	897	2,430	351	810	1,890	349	0.597	0.623
Transfers made in past month	645	2,409	351	549	2,055	349	0.421	0.634
p-value from joint significance of 14 covariates:							0.87	

Notes: Observations in the table are worker-rounds of data (two rounds per worker). Sample includes 359 respondents who participated in at least one round of the work program and have data from at least one data source for that round (either the payday data, the survey, or both). All money amounts are in Malawian Kwacha (MK); during the study period the market exchange rate was approximately MK400 to the US dollar, and the PPP exchange rate was approximately MK160 to the US dollar.

Itemized expenditure data does not include all purchases, and so these estimates are likely to be a lower bound. Asset purchases are measured since the previous survey, a period of approximately two months. Loans are measured since November 1st in round 1 and since January 1st in round 2, a period of approximately one month. Transfers are measured over the month leading up to the survey interview.

† Column 8 presents balance tests for the logs of the financial variables instead of the levels; we use the inverse hyperbolic sine transformation of Burbidge et al. (1988) instead of the log function because our data contains zeroes.

Table 3

Comparison of Available Definitions of Temptation Spending

	Panel A - Summary Statistics			Panel B - Pairwise Correlations					
	Mean	SD	N						
A. Waste/ Temptation	116.55	343.91	689	**A. Waste/ Temptation**					
B. Money Wasted (Round 2 Only)	305.85	685.04	346	0.25	**B. Money Wasted**				
C. Regrets	44.59	181.55	689	0.16	0.13	**C. Regrets**			
D. Unplanned Purchases	49.57	121.38	689	-0.01	0.00	0.10	**D. Unplanned Purchases**		
E. Against Plans	62.21	287.56	689	0.16	0.02	0.05	-0.03	**E. Against Plans**	
F. Alcohol and Tobacco	14.33	46.79	689	0.07	0.01	-0.01	0.09	-0.02	**F. Alcohol and Tobacco**
G. Alcohol, Tobacco, Donuts, and Soda	65.28	89.35	689	0.13	0.23	0.03	0.19	0.06	0.61

Notes: Sample includes 359 respondents who participated in at least one round of the work program and have data from at least one data source for that round (either the payday data, the survey, or both). Correlations are estimated using pooled data for rounds 1 and 2 except for the "Money Wasted" variable (Row B) which exists only in Round 2. All money amounts are in Malawian Kwacha (MK).

Table 4

Effects of Market Day Wage Payment on Expenditures

	(1)	(2)	(3)	(4)	(5)	(6)
	Payday survey panel - Spending at market on the four payday weekends			**Household survey data**		
Dependent variable:	Amount spent on Friday and Saturday, all Weekends (MK)	Amount spent on payday (MK)	(Spending on payday) / (Income received)	Income received since last Friday (MK)	Remaining cash out of income received since last Friday (MK)	Total spending since Friday from itemized expenditure data (MK)
Panel A - Lump sum and weekly payment group pooled						
Market Day Payment	**-812.5*** **	**-26.91**	**0.0136**	**17.88**	**-92.03**	**128.0**
	(113.5)	(89.43)	(0.0290)	(194.9)	(76.05)	(161.3)
Dependent variable mean, control group (Friday paydays)	3,293	1,688	0.622	3,081	579.2	3,147
Number of observations	689	689	689	689	689	689
Panel B - Lump sum payment group only						
Market Day Payment	**-756.4*** **	**-25.77**	**0.0372**	**161.6**	**-162.5**	**193.3**
	(171.9)	(119.8)	(0.0421)	(230.4)	(107.2)	(237.8)
Dependent variable mean, control group (Friday paydays)	3,068	1,247	0.534	3,753	670.6	3,341
Number of observations	345	345	345	345	345	345
Panel C - Weekly payment group only						
Market Day Payment	**-810.5*** **	**-25.76**	**-0.0026**	**0.0280**	**-29.81**	**52.49**
	(153.5)	(114.9)	(0.0385)	(258.0)	(114.0)	(233.9)
Dependent variable mean, control group (Friday paydays)	3,530	2,151	0.714	2,378	483.5	2,944
Number of observations	344	344	344	344	344	344

Notes: Sample includes 359 respondents who participated in at least one round of the work program and have data from at least one data source for that round (either the payday data, the survey, or both). Regressions are run on pooled data from round one and round two. Boldface type indicates the treatment variable of interest. 1 USD was worth approximately MK400 at market exchange rates and MK160 at PPP exchange rates during the study period. All regressions control for stratification cell fixed effects, an index of baseline asset ownership based on first principal components, indicators for the number of days after the weekend the interview occurred, baseline total spending and (if available) the baseline value of the outcome variable. Heteroskedasticity-robust standard errors, clustered by worker, in parentheses: * $p<0.1$; ** $p<0.05$; *** $p<0.01$

Table 5

Effects of Market Day Wage Payment on Temptation Spending

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	PCA indices of temptation spending		**Measures of temptation spending (MK)**						
Dependent variable:	Preferred Index, Omitting Col. 2	Including Col. 2	Goods respondent says they waste money on or are tempted to buy	Self-reported total of money "wasted"	Goods respondent often regrets purchasing	Unplanned Purchases	Goods respondent often buys in violation of prior plans	Alcohol and Tobacco	Alcohol, Tobacco, Donuts, and Soda
Rounds Available	1 & 2	2 only	1 & 2	2 only	1 & 2	1 & 2	1 & 2	1 & 2	1 & 2
Market Day Payment	**-0.00596**	**-0.12916**	**-10.04**	**-47.65**	**4.073**	**-4.652**	**-14.81**	**2.983**	**-2.020**
	(0.0785)	(0.1275)	(25.37)	(71.39)	(16.07)	(8.286)	(20.10)	(3.230)	(5.781)
Dependent variable mean, control group (Friday paydays)	-0.0124	0.0000	131.2	324.3	42.69	52.66	69.06	12.91	67.19
Control-group SD	1.234	1.378	382.7	719.7	178.6	122.6	321.4	44.27	93.67
MDE on temptation spending at 80% power (SDs)	0.178	0.259	0.186	0.278	0.252	0.189	0.175	0.204	0.173
Number of observations	689	346	689	346	689	689	689	689	689

Notes: Sample includes 359 respondents who participated in at least one round of the work program and have data from at least one data source for that round (either the payday data, the survey, or both). Regressions are run on pooled data from round one and round two. Boldface type indicates the treatment variable of interest. 1 USD was worth approximately MK400 at market exchange rates and MK160 at PPP exchange rates during the study period. All regressions control for stratification cell fixed effects, an index of baseline asset ownership based on first principal components, indicators for the number of days after the weekend the interview occurred, baseline total spending and (if available) the baseline value of the outcome variable. Heteroskedasticity-robust standard errors, clustered by worker, in parentheses: * $p<0.1$; ** $p<0.05$; *** $p<0.01$.

Table 6

Treatment Effect Heterogeneity by Gender

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	PCA indices of temptation		**Measures of temptation spending (MK)**						
Dependent variable:	Preferred Index, Omitting Col. 4	Including Col. 4	Goods respondent says they waste money on or are tempted to buy	Self-reported total of money "wasted"	Goods respondent often regrets purchasing	Unplanned Purchases	Goods respondent often buys in violation of prior plans	Alcohol and Tobacco	Alcohol, Tobacco, Donuts, and Soda
Rounds Available	1 & 2	2 only	1 & 2	2 only	1 & 2	1 & 2	1 & 2	1 & 2	1 & 2
(Market Day Payments) X (Male)	**0.137**	**0.116**	**63.15**	**57.64**	**13.58**	**-8.383**	**-77.17**	**8.169**	**11.13**
	(0.200)	(0.291)	(62.09)	(107.3)	(28.65)	(19.94)	(54.72)	(8.490)	(14.40)
(Market Day Payments) X (Female)	**-0.0846**	**-0.264****	**-50.44****	**-93.66**	**1.143**	**-3.668**	**11.02**	**0.328**	**-8.255**
	(0.0751)	(0.134)	(24.12)	(88.72)	(19.08)	(8.994)	(16.83)	(3.058)	(5.800)
Male	0.302	0.145	-2.483	-122.9	-19.73	16.70	77.93	13.48**	12.06
	(0.183)	(0.243)	(60.90)	(96.88)	(24.52)	(19.38)	(55.82)	(6.844)	(12.18)
Difference in Treatment Effects	**0.222**	**0.380**	**113.59***	**151.30**	**12.44**	**-4.71**	**-88.19**	**7.84**	**19.39**
	(0.218)	0.320	(67.32)	(135.17)	(34.20)	(22.58)	(57.44)	(9.22)	(15.78)
Number of observations	674	331	674	331	674	674	674	674	674

Notes: Sample includes 359 respondents who participated in at least one round of the work program and have data from at least one data source for that round (either the payday data, the survey, or both). Regressions are run on pooled data from round one and round two. Boldface type indicates the treatment variable of interest. 1 USD was worth approximately MK400 at market exchange rates and MK160 at PPP exchange rates during the study period. All regressions control for stratification cell fixed effects, an index of baseline asset ownership based on first principal components, indicators for the number of days after the weekend the interview occurred, baseline total spending and (if available) the baseline value of the outcome variable. Heteroskedasticity-robust standard errors, clustered by worker, in parentheses: * $p<0.1$; ** $p<0.05$; *** $p<0.01$.

Table 7

Treatment Effect Heterogeneity by Whether Worker Has Children in School

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	of temptation spending		**Measures of temptation spending (MK)**						
Dependent variable:	Preferred Index, Omitting Col. 4	Including Col. 4	Goods respondent says they waste money on or are tempted to buy	Self-reported total of money "wasted"	Goods respondent often regrets purchasing	Unplanned Purchases	Goods respondent often buys in violation of prior plans	Alcohol and Tobacco	Alcohol, Tobacco, Donuts, and Soda
Rounds Available	1 & 2	2 only	1 & 2	2 only	1 & 2	1 & 2	1 & 2	1 & 2	1 & 2
(Market Day Payment) X (Children in School)	**-0.0828**	**-0.179**	**-40.10**	**-60.67**	**8.847**	**-14.18**	**-18.71**	**2.094**	**-5.731**
	(0.0898)	(0.154)	(28.78)	(83.66)	(19.71)	(9.048)	(24.25)	(3.480)	(6.843)
(Market Day Payment) X (No Children in School)	**0.330***	**0.0638**	**115.6****	**37.01**	**-15.23**	**36.64***	**5.620**	**7.209**	**14.23**
	(0.170)	(0.198)	(51.66)	(140.4)	(14.44)	(21.08)	(24.56)	(8.709)	(11.83)
Children in School	0.134	-0.243	83.03**	208.5**	35.56*	8.399	61.38	-2.227	0.770
	(0.145)	(0.257)	(42.08)	(100.5)	(18.71)	(15.13)	(44.12)	(6.384)	(10.84)
Difference in Treatment Effects	**-0.413****	**-0.243**	**-155.67*****	**-97.68**	**24.08**	**-50.81****	**-24.33**	**-5.12**	**-19.97**
	(0.196)	(0.257)	(59.57)	(168.30)	(25.33)	(23.18)	(36.34)	(9.46)	(14.20)
Number of observations	689	346	689	346	689	689	689	689	689

Notes: Sample includes 359 respondents who participated in at least one round of the work program and have data from at least one data source for that round (either the payday data, the survey, or both). Regressions are run on pooled data from round one and round two. Boldface type indicates the treatment variable of interest. 1 USD was worth approximately MK400 at market exchange rates and MK160 at PPP exchange rates during the study period. All regressions control for stratification cell fixed effects, an index of baseline asset ownership based on first principal components, indicators for the number of days after the weekend the interview occurred, baseline total spending and (if available) the baseline value of the outcome variable. Heteroskedasticity-robust standard errors, clustered by worker, in parentheses: * $p<0.1$; ** $p<0.05$; *** $p<0.01$.

Table 8

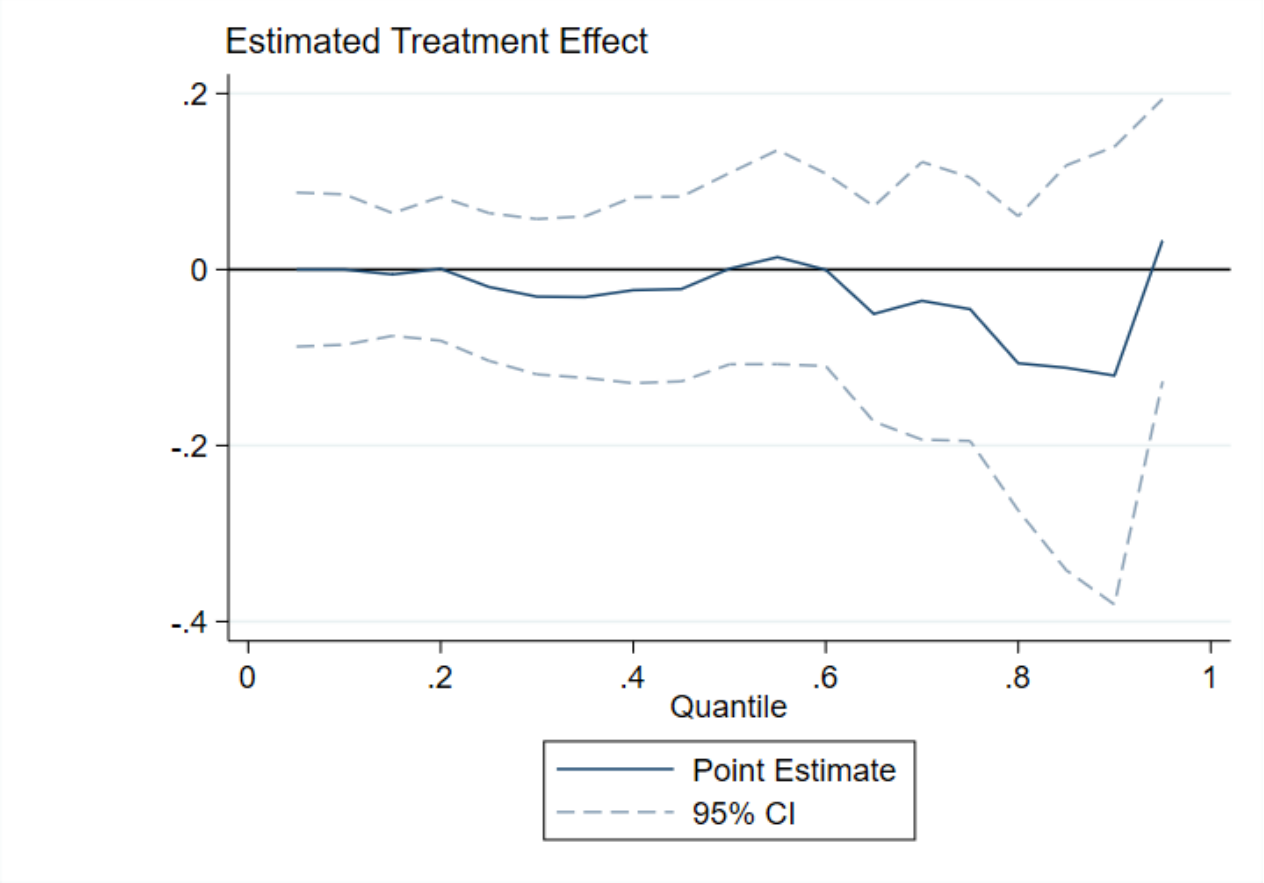
Peer Effects on Temptation Spending

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	PCA indices of temptation spending		**Measures of temptation spending (MK)**						
Dependent variable:	Preferred Index, Omitting Col. 4	Including Col. 4	Goods respondent says they waste money on or are tempted to buy	Self-reported total of money "wasted"	Goods respondent often regrets purchasing	Unplanned Purchases	Goods respondent often buys in violation of prior plans	Alcohol and Tobacco	Alcohol, Tobacco, Donuts, and Soda
Rounds Available	1 & 2	2 only	1 & 2	2 only	1 & 2	1 & 2	1 & 2	1 & 2	1 & 2
Market Day Payment	**-0.00984**	**-0.155**	**-9.707**	**-56.22**	**0.406**	**-5.508**	**-16.53**	**3.134**	**-1.946**
	(0.0814)	(0.132)	(26.16)	(71.07)	(15.79)	(8.427)	(20.77)	(3.323)	(5.876)
Average Peer Treatment Status	**-0.0195**	**-0.289**	**20.43**	**-96.09**	**-16.28**	**21.74***	**-21.31**	**-2.734**	**-4.983**
	(0.120)	(0.192)	(39.37)	(104.7)	(21.31)	(12.93)	(33.08)	(4.549)	(8.691)
Dependent variable mean, control group (Friday paydays)	-0.0209	0.0000	131.2	324.3	42.69	52.66	69.06	12.84	65.95
Control-group SD	1.228	1.378	383.7	719.7	179.1	122.9	322.3	44.34	90.83
Number of observations	684	346	684	346	684	684	684	684	684

Notes: Sample includes 359 respondents who participated in at least one round of the work program and have data from at least one data source for that round (either the payday data, the survey, or both). Regressions are run on pooled data from round one and round two. Boldface type indicates the treatment variable of interest. All regressions control for stratification cell fixed effects, an index of baseline asset ownership based on first principal components, indicators for the number of days after the weekend the interview occurred, baseline total spending and (if available) the baseline value of the outcome variable.

Heteroskedasticity-robust standard errors, clustered by worker, in parentheses: * $p<0.1$; ** $p<0.05$; *** $p<0.01$.

Appendix Figure 1
Quantile Treatment Effects on Temptation Spending
(Preferred Temptation Spending Index)



Notes: Sample includes 359 respondents who participated in at least one round of the work program and have data from at least one data source for that round (either the payday data, the survey, or both). All money amounts are in Malawian Kwacha (MK); during the study period the market exchange rate was approximately MK400 to the US dollar, and the PPP exchange rate was approximately MK160 to the US dollar. Outcome is the PCA index from Table 5 that omits column 4 and uses data from both rounds of the study.

Appendix Table 1

Balance of Background Characteristics and Financial Outcomes by Intervention Round

	Friday Payment Group			Saturday Payment			Balance
	Mean	SD	N	Mean	SD	N	test
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A - Round 1							
Background characteristics							
Male	0.33	0.47	175	0.31	0.47	175	0.819
Married	0.73	0.45	173	0.67	0.47	171	0.214
Age (Years)	38.98	15.33	175	40.78	15.34	175	0.274
Years of Education Completed	3.54	3.32	173	3.50	3.00	174	0.899
Survey date (days after Sunday)	2.69	1.27	171	2.56	1.18	172	0.319
Prefers lump sum wage payments	0.76	0.43	174	0.71	0.46	175	0.238
Financial outcomes (in units of MK unless noted)							
Income received since past Friday	3,464	9,503	175	2,540	4,011	175	0.237
Remaining cash holdings out of income received	610	2,055	175	676	2,958	175	0.807
Total spending since Friday	4,160	5,138	175	3,255	3,340	175	0.052
Asset Ownership (PCA)	-0.01	2.67	175	0.01	2.71	175	0.947
Loans received in past month	2,743	8,216	175	3,150	13,366	175	0.731
Loans made in past month	888	3,510	175	578	2,295	175	0.328
Transfers received in past month	980	2,686	175	727	1,505	175	0.279
Transfers made in past month	451	1,648	175	744	2,702	175	0.221
p -value from joint significance of 14 covariates:			0.32				
Panel 2 - Round 2							
Background characteristics							
Male	0.28	0.45	176	0.36	0.48	174	0.148
Married	0.67	0.47	174	0.72	0.45	170	0.303
Age (Years)	40.93	16.14	176	38.82	14.46	174	0.197
Years of Education Completed	3.31	3.13	176	3.74	3.18	171	0.199
Midline survey date (days after Sunday)	2.42	1.02	175	2.49	1.09	171	0.582
Prefers lump sum wage payments	0.69	0.46	176	0.79	0.41	173	0.036
Financial outcomes (in units of MK unless noted)							
Income received since past Friday	3,011	8,723	176	2,993	5,520	174	0.982
Remaining cash holdings out of income received	535	1,425	176	752	3,312	174	0.427
Total spending since Friday	3,823	4,589	176	3,591	4,105	174	0.618
Asset Ownership (PCA)	-0.15	2.29	176	0.16	3.03	174	0.284
Loans received in past month	2,477	7,378	176	3,421	13,862	174	0.428
Loans made in past month	477	1,499	176	991	3,915	174	0.106
Transfers received in past month	815	2,148	176	893	2,213	174	0.737
Transfers made in past month	838	2,971	176	354	1,035	174	0.042
p -value from joint significance of 14 covariates:			0.14				

Notes: Sample includes 359 respondents who participated in at least one round of the work program and have data from at least one data source for that round (either the payday data, the survey, or both). All money amounts are in Malawian Kwacha (MK); during the study period the market exchange rate was approximately MK400 to the US dollar, and the PPP exchange rate was approximately MK160 to the US dollar.

Appendix Table 2

Balance of Background Characteristics and Financial Outcomes for Weekly Payment

	Friday Payment Group			Saturday Payment Group			Balance test
	Mean	SD	N	Mean	SD	N	*p* -value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Background characteristics							
Male	0.28	0.45	174	0.34	0.47	176	0.272
Married	0.70	0.46	171	0.73	0.44	172	0.439
Age (Years)	40.09	16.00	174	39.26	14.43	176	0.612
Years of Education Completed	3.24	3.08	173	3.66	3.08	173	0.232
Midline survey date (days after Sunday)	2.61	1.16	169	2.57	1.15	175	0.744
Prefers lump sum wage payments	0.72	0.45	173	0.77	0.42	176	0.290
Financial outcomes (in units of MK unless noted)							
Income received since past Friday	3,803	12,044	174	3,165	5,994	176	0.618
Remaining cash holdings out of income received	540	1,766	174	893	3,476	176	0.202
Total spending since Friday	4,027	5,148	174	3,553	4,116	176	0.357
Asset Ownership (PCA)	-0.31	2.09	174	0.13	3.10	176	0.114
Loans received in past month	1,790	4,714	174	4,121	17,798	176	0.215
Loans made in past month	562	1,636	174	993	4,211	176	0.264
Transfers received in past month	994	2,648	174	851	2,269	176	0.624
Transfers made in past month	509	2,493	174	633	2,562	176	0.325
p -value from joint significance of 12 covariates:		0.73					

Notes: Sample includes 359 respondents who participated in at least one round of the work program and have data from at least one data source for that round (either the payday data, the survey, or both). All money amounts are in Malawian Kwacha (MK); during the study period the market exchange rate was approximately MK400 to the US dollar, and the PPP exchange rate was approximately MK160 to the US dollar.

Appendix Table 3

Balance of Background Characteristics and Financial Outcomes for Lump Sum Payment

	Friday Payment Group			Saturday Payment Group			Balance test
	Mean	SD	N	Mean	SD	N	*p* -value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Background characteristics							
Male	0.33	0.47	177	0.34	0.47	173	0.879
Married	0.70	0.46	176	0.66	0.48	169	0.331
Age (Years)	39.83	15.55	177	40.35	15.41	173	0.742
Years of Education Completed	3.60	3.36	176	3.58	3.11	172	0.950
Midline survey date (days after Sunday)	2.50	1.15	177	2.47	1.13	168	0.786
Prefers lump sum wage payments	0.73	0.45	177	0.72	0.45	172	0.861
Financial outcomes (in units of MK unless noted)							
Income received since past Friday	2,680	4,671	177	2,360	3,185	173	0.502
Remaining cash holdings out of income received	603	1,770	177	532	2,743	173	0.795
Total spending since Friday	3,955	4,588	177	3,290	3,319	173	0.151
Asset Ownership (PCA)	0.14	2.81	177	0.04	2.62	173	0.721
Loans received in past month	3,416	9,886	177	2,435	7,091	173	0.264
Loans made in past month	800	3,441	177	571	1,645	173	0.428
Transfers received in past month	802	2,197	177	768	1,410	173	0.871
Transfers made in past month	779	2,322	177	464	1,360	173	0.123
p -value from joint significance of 12 covariates:		0.96					

Notes: Sample includes 359 respondents who participated in at least one round of the work program and have data from at least one data source for that round (either the payday data, the survey, or both). All money amounts are in Malawian Kwacha (MK); during the study period the market exchange rate was approximately MK400 to the US dollar, and the PPP exchange rate was approximately MK160 to the US dollar.

Appendix Table 4

Evidence of the Tempting Nature of Market Days

A. Expenditures that are wastes of money or that repsondent is tempted into buying

What are things that you sometimes waste money on, or that you are sometimes tempted to buy but should not spend money on? List as many things as you can think of:

B. Own calculation of total money wasted/tempted into spending

[**LAST FRIDAY/LAST SATURDAY/SINCE SUNDAY**], how much money did you waste or how much money were you tempted into spending that you should not have spent?

C. Expenditures that the respondent commonly regrets

Do you ever spend money and then later regret it?	0 - No 1 - Yes **[Yes --> a.]**
a. What do you spend the money on? List as many things as you can think of: **Interviewer: Mark <u>all</u> that apply**	

D. Unplanned purchases on an individual expenditure basis (asked for each good purchase from an itemized lis

Did you plan to buy [ITEM] or did you decide to buy only when you saw it?

E. Expenditures that are commonly unplanned

Now think about the plans you make for spending your money. Do you ever make plans or budgets for spending, but then spend money on things you didn't plan on?	0 - No 1 - Yes **[Yes --> a.]**
a. When this happens, what are the unplanned items? List as many things as you can think of: **Interviewer: Mark <u>all</u> that apply**	

<u>Notes:</u> This table presents the survey questions that were used for respondent self-reports of temptation spending. Questions were asked in Chichewa, the local language.

Appendix Table 5

Self-reported temptation goods, by question

	(1)	(2)	(3)
	Share reporting this good		
	A. Wastes of money/tempted to buy	C. Commonly regretted purchases	E. Commonly unplanned purchases
Gifts for children	0.15	0.03	0.01
Other relish*	0.15	0.10	0.08
Clothes or shoes	0.14	0.13	0.07
Fried fish	0.10	0.03	0.01
Biscuits	0.07	0.01	0.01
Alcohol	0.06	0.04	0.03
Fresh fish	0.06	0.02	0.03
Chips (french fries)	0.05	0.01	0.01
Sweets	0.04	0.01	0.01
Fried Meats	0.03	0.00	0.00
Gifts for sex partner/spouse	0.01	0.00	0.02
Gifts for others	0.00	0.01	0.01
Other	0.31	0.36	0.49
Food	0.08	0.05	0.12
Fritters/Donuts	0.05	0.02	0.01
Illness/Medicine	0.03	0.01	0.20
Fish	0.03	0.01	0.01
Expensive food	0.00	0.07	0.00
Bad business decisions	0.00	0.05	0.00
Funerals	0.00	0.00	0.05
N	346	349	349

Notes: This table shows the share of respondents who mentioned each good as an example of a temptation good based on the three different definitions, which correpond to the questions from Appendix Table 4. We list the most-common "other" options, but not a complete list due to space constraints.

Sample includes 359 respondents who participated in at least one round of the work program and have data from at least one data source for that round (either the payday data, the survey, or both).

*"Relish" is a Malawian English word for the side dish that accompanies the typical staple, *nsima* (cornmeal porridge). It is typically made of meat, fish, or vegetables.

Appendix Table 6

Evidence of the Tempting Nature of Market Days

Panel A - Tempting Situations

"What are situations in which you waste money or are tempted to spend money that you would rather not spend?"

(Options not read aloud, response coded by enumerator)	Mean	Obs.
Market days	0.37	346
Going to the trading center in general (not just market days)	0.09	346
State holidays	0.05	346
Going to the Boma	0.01	346
Seeing your sex partner	0.01	346
Friday nights	0.00	346
Other	0.17	346
No response/no situation mentioned	0.39	346

Panel B - Comparisons of which situation is more tempting

Which situation makes you more tempted to spend money you will later regret?		A is more Tempting	B is more Tempting	No Difference	Obs.
A. Market day or	B. Day before market day	0.69	0.15	0.16	346
A. Having cash in pocket at trading center or	B. Having cash in pocket elsewhere	0.66	0.16	0.18	345
A. Friday or	B. Saturday	0.25	0.51	0.24	346
A. Market or	B. The night before the market	0.74	0.15	0.11	345

Panel C - Reasons you sometimes waste money

Reasons you sometimes waste money (Options not read aloud, response coded by enumerator)	Mean	Obs.
Tempted to buy things I should not	0.42	346
Lack of plans	0.25	346
Buying things on impulse	0.18	346
Drinking	0.06	346
Nothing good to do with the money	0.06	346
Relatives beg for money	0.05	346
Friends beg for money	0.03	346
Other	0.27	346

Panel D - Visits to the market per month

	Mean	Obs.
Total	6.08	365
% on Fridays	0.26	252
% on Saturdays	0.32	252
% on other days of the week	0.42	252

Notes: This table presents the survey questions and the response results indicating the tempting nature of market days. Sample includes 359 respondents who participated in at least one round of the work program and have data from at least one data source for that round (either the payday data, the survey, or both).

Appendix Table 7

Temptation Spending by Gender

	Mean			Std. dev.			10th percentile			Median			90th percentile			Obs.		
Temptation Spending	All	Male	Female	All	Male	Female	All	Male	Female	All	Male	Female	All	Male	Female	All	Male	Female
Waste/ Temptation	117	152	100	344	439	293	0	0	0	0	0	0	260	310	260	689	204	470
Money Wasted	306	240	332	685	552	723	0	0	0	25	0	50	800	780	900	346	98	233
Regrets	45	46	45	182	169	190	0	0	0	0	0	0	40	100	11	689	204	470
Unplanned Purchases	50	67	43	121	145	111	0	0	0	0	0	0	156	250	130	689	204	470
Against Plans	62	107	45	288	395	229	0	0	0	0	0	0	50	160	0	689	204	470
Alcohol and Tobacco	14	28	9	47	67	34	0	0	0	0	0	0	40	100	20	689	204	470
Alcohol, Tobacco, Donuts, and Soda	65	91	55	89	115	74	0	0	0	40	60	40	160	230	120	689	204	470

Notes: Sample includes 359 respondents who participated in at least one round of the work program and have data from at least one data source for that round (either the payday data, the survey, or both). All money amounts are in Malawian Kwacha (MK); during the study period the market exchange rate was approximately MK400 to the US dollar, and the PPP exchange rate was approximately MK160 to the US dollar.

Appendix Table 8

Peer Effects on Temptation Spending (Using Four Closest Neighbors Instead of Two)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	PCA indices of temptation spending		**Measures of temptation spending (MK)**						
Dependent variable:	Preferred Index, Omitting Col. 4	Including Col. 4	Goods respondent says they waste money on or are tempted to buy	Self-reported total of money "wasted"	Goods respondent often regrets purchasing	Unplanned Purchases	Goods respondent often buys in violation of prior plans	Alcohol and Tobacco	Alcohol, Tobacco, Donuts, and Soda
Rounds Available	1 & 2	2 only	1 & 2	2 only	1 & 2	1 & 2	1 & 2	1 & 2	1 & 2
Market Day Payment	**-0.00384**	**-0.126**	**-10.97**	**-45.97**	**1.017**	**-5.675**	**-16.34**	**3.418**	**-1.410**
	(0.0812)	(0.127)	(25.82)	(71.05)	(15.39)	(8.292)	(20.50)	(3.334)	(5.912)
Average Peer Treatment Status	**0.148**	**0.0718**	**-3.221**	**32.95**	**-9.316**	**32.05****	**-30.43**	**3.906**	**7.705**
	(0.137)	(0.195)	(40.63)	(110.9)	(23.73)	(13.18)	(41.69)	(5.539)	(10.04)
Dependent variable mean, control group (Friday paydays)	-0.0209	0.0000	131.4	324.3	42.88	52.91	69.46	12.84	65.95
Control-group SD	1.228	1.378	383.7	719.7	179.1	122.9	322.3	44.34	90.83
Number of observations	684	346	684	346	684	684	684	684	684

Notes: Sample includes 359 respondents who participated in at least one round of the work program and have data from at least one data source for that round (either the payday data, the survey, or both). Regressions are run on pooled data from round one and round two. Boldface type indicates the treatment variable of interest. All regressions control for stratification cell fixed effects, an index of baseline asset ownership based on first principal components, indicators for the number of days after the weekend the interview occurred, baseline total spending and (if available) the baseline value of the outcome variable.

Heteroskedasticity-robust standard errors, clustered by worker, in parentheses: * $p<0.1$; ** $p<0.05$; *** $p<0.01$.