

How Important is Temptation Spending? Maybe Less than We Thought

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

March 11, 2021

[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 in rural Malawi. 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 examine the potential factors that may be driving these null results, and argue that future research may find a limited role for temptation in the economic decisions of the poor.

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

JEL Codes: D90, D91, O12

* Lasse Brune is a Postdoctoral Fellow at the Global Poverty Research Lab at Northwestern University; his email is lasse.brune@northwestern.edu. Jason Kerwin (corresponding author) is an Assistant Professor in the Department of Applied Economics at the University of Minnesota; his email is jkerwin@umn.edu. Qingxiao Li is a PhD Candidate in the Department of Applied Economics at the University of Minnesota; his email is lix5376@umn.edu. We are grateful to Marc Bellemare, Jeff Bloem, Audrey Dorélien, Anett John, Mike Kuhn, 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 theoretical work suggests that it does (Banerjee and Mullainathan 2010), and policymakers are often concerned that participants will mispend cash transfers on temptation goods (Harvey 2007, Ikiara 2009, Evans and Popova 2016). Temptation goods are typically defined by researchers to include goods that are commonly perceived as harmful (as in Evans and Popova 2016) or that the people themselves would prefer not to buy when asked at a different time (as in O'Donoghue and Rabin 1999). Since temptation goods are valued only in the moment, and not ahead of time, that money is as good as wasted. 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 (see Evans and Popova, 2016, for an in-depth literature review on the impact of cash transfers on temptation good expenditures). The disconnect between the theoretical literature and the evidence from cash transfers raises the question of how important temptation really is in the financial lives of people living in poor countries.

In this study, we study the relevance of temptation by attempting to experimentally vary workers' exposure to temptation at the time that they receive cash payments and examine whether and how workers' spending differs, using a field experiment in southern Malawi. We do this by having treatment-group workers pick up their pay at the location of a weekly market *on market day*. In contrast, control-group workers picked up their pay at the same market location but on the day before market day. Market days at local markets are commonly identified by participants of our study as a highly-tempting environment. Thus, the intervention was designed to induce the treatment group to have cash on hand in a tempting environment, while the control group had cash on hand in an otherwise-equivalent environment without the source of temptation. The experiment held transaction costs, such as time and transportation costs, constant, by requiring having each

worker come to the payment location at the local market both on market day and the day before.¹

Our results do not provide strong evidence in favor of the typical temptation narrative. Direct exposure to the tempting market environment at the time of payment induces no appreciable changes in expenditure, nor in temptation spending in particular. 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 substantially from the typical definitions and that the magnitude of temptation spending varies substantially across definitions. For example, if we directly ask workers how much money they wasted or were tempted into spending that they should not have spent, temptation spending is 20 times larger than the conventional approaches that focus on alcohol and tobacco. Crucially, the different ways of measuring temptation spending do not alter the interpretation of for our core finding: we cannot reject that spending decisions are unaffected by the variation in the receipt of payment in a tempting environment that our experiment induces. A caveat to this finding is that the magnitude

¹ 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.

of the point estimate is quite sensitive to the definition used, ranging from 3 percent of average control-group temptation spending up to 23 percent.

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 variation in payment timing did not induce sufficient variation in temptation. 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 most compelling is that the treatment was simply not very intense. Specifically, we believe that we exposed workers to one of the most-tempting environments they commonly encounter, and that the treatment was implemented successfully—but that this exposure has intrinsically small effects on behavior. If the upper bound on the practical effect of temptation exposure on spending is small, this implies that policymakers and economists may both be overly-concerned about the role of temptation in driving spending and savings decisions in developing countries.

This paper contributes to the empirical literature on temptation 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.²

These results provide important insights into the potential role of temptation in the economic lives of the poor. Given our findings, we argue that researchers should not expect strong effects of temptation exposure on temptation spending. 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 savory snacks, gifts for their children, clothes, and food. Estimates of temptation effects that focus on alcohol and tobacco alone may be downward biased. Relying on individuals' own determination

² Our intervention could also be framed as inducing a waiting period before people could make consumption choices. Previous research has shown that waiting periods induce more patient choices (Imas, Kuhn, and Mironova 2018) and healthier choices (Brownback, Imas, and Kuhn 2019) by facilitating a shift from heuristic to deliberative processing. In our study, the control group had to wait to make purchases at the market day, while the treatment group could make them immediately. Thus one channel for increasing temptation spending is that the treatment could have caused more-heuristic choices.

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.

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.³ 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 (PPP 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. Our experiment

³ Attrition is balanced across treatment and control: see Panel 2 of the Appendix Table 1 for the balance tests after attrition in round 2.

occurred during the lean season of the year, which begins around maize planting and continues until the harvest. At this time of the year, money is typically tight and jobs are scarce. 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 all paid at the site of the largest local market. They were randomly assigned, independently by round, to receive their wages either on market day (Saturday) or the day before market day (Friday); all payments were made at the same location. 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. The weekly payment was MK700 in round one and MK800 in round two; correspondingly, the lump-sum payment was MK2800 in the first round and MK3200 in the second. To ensure transaction costs, such as transportation and time costs, were held constant across wage payment modes, workers assigned to Saturday paydays were also asked to come to the payroll site on Fridays, and vice versa. 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. Equalizing transaction costs is key to interpreting differences in spending between Friday-payment workers and Saturday-payment workers as an effect of the differences in the degree of temptation of the environment. Since Saturday-payment workers are at the market on market days to get their pay, their marginal cost of purchasing a good during the market is just the price of the good. A worker who was not at the market would have to pay not just the price of the good but also the transportation and time cost of getting to the market. By requiring Friday-payment workers

to show up during the market day as well, this difference in the overall cost of goods is eliminated.

The 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 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,

household-level expenditure, physical assets, saving, transfers, cash on hand, and details on the worker's expenditures since the first day of the final payday weekend.⁴ 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 baseline (pre-treatment) variables 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 ($p=0.91$).^{5,6} We also fail to reject the null when we compare Friday and Saturday workers who received their pay in a lump sum ($p=0.97$) or weekly ($p=0.79$).

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 (PPP USD \$18.77) in cash income and spent about MK4,000 (PPP 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.

⁴ The cash on hand variable was measured as the household's remaining cash holdings out of income received starting since the Friday before the survey interview, and so is not necessarily equal to the household's entire savings.

⁵ This table pools workers across rounds; we also find that the two study arms of interest are balanced when we analyze them separately by round (Appendix Table 1).

⁶ The difference in income across study arms is statistically-insignificant ($p=0.503$) but somewhat large in magnitude: 17% higher for the Friday payment group. We therefore control for baseline income in our main analyses.

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 otherwise-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—like most consumers—commonly identify categories of spending that they wish to reduce (Banerjee and Duflo, 2007).

In an attempt to expand on the frameworks used in previous studies, we explore a number of different ways to capture temptation spending that builds on the respondent-driven approach in Banerjee and Duflo (2007). We use detailed survey instruments that both allow us to see how spending levels and the results in our experiment vary using different approaches and shed light on their relative merits. Table 3 overviews the seven temptation outcomes we use, listing key characteristics as well advantages and disadvantages.

We use five definitions that are based on respondents’ own categorizations. For three of those additional definitions (A, C, and E) we first allow respondents to identify categories of expenditure that they themselves typically see as problematic. In second step we match actual spending from an itemized expenditure module with each respondent’s categories to compute measures of temptation spending, separately for each definition. For this category-elicitation approach our instruments included the follow definitions of temptation goods: goods that the respondent is tempted into purchasing that they should not buy or that are a waste of money

(definition A); purchases that the respondent commonly regrets after the fact (definition C); and goods that are commonly unplanned purchases (definition E).⁷

We also use two other respondent-driven measures of temptation spending. For definition B we the ask the respondent to report the total amount of money they “wasted”, where the latter term is the English translation the Chichewa expression that, based on extensive piloting, best captures the spirit of temptation goods. Definition D focuses on unplanned purchases, and is captured by asking about every good in the itemized list of purchases, whether the purchase was planned beforehand (following Brune et al., 2017). Unplanned purchases are a proxy for temptation spending in this case. The English translations of the exact survey questions we used for all respondent self-reports of temptation spending are shown in Appendix Table 2.

There are several empirical patterns worth highlighting when considering the types of goods that are included in the different definitions of temptation spending (see Appendix Table 3 for a detailed breakdown). First, the items that respondents categorize as temptation goods differ starkly from the narrow definitions of only harmful goods that are often used to capture temptation spending. For example, under definition A, sweets and alcohol are mentioned by 11 percent and 6 percent of respondents, respectively. In contrast, three of the top four categories are gifts for children (15 percent), clothes or shoes (14 percent) and food (11 percent). Second, the items included in respondent’s own categorizations differ across definitions. For example, gifts for children is commonly reported under definition A. However, this category is not commonly reported as regretted (definition C) or unplanned (definition E).

Third, categorizing purchases based on what respondents consider to be common

⁷ 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. Very few workers used all three “other” spaces on the survey. The survey module where workers categorized temptation goods was conducted after we elicited actual purchases, to avoid potential underreporting due to priming.

temptation goods can lead to important miscategorizations: items that are typically considered temptation spending in given category are not necessarily temptation goods all or even most of the time. For example, *all* spending on food or clothes, commonly mentioned as categories for definitions A, C and E, would be classified as temptation spending for respondents who list this category. Obviously some purchases of clothing and food are necessities and do not fit a reasonable definition of temptation spending; what respondents clearly mean is that they sometimes make impulse purchases of these goods, rather than that all purchases of them are temptation spending. Conversely, if a respondent does not consider an item a common temptation good, no purchases of the good will be coded as temptation spending, even if some actually are. This inclusion/exclusion error is a downside that does not apply the direct elicitation of spending under definition B. The main disadvantage from using definition B is susceptible to ex-post rationalizations: unlike the categorizations of common temptation goods, elicitation of total temptation spending is only done after the fact.

Fourth, the definitions that rely on eliciting regretted and unplanned goods have an important limitation in that not all purchases that are unplanned or regretted are temptation spending. For example, healthcare expenses and funerals are commonly listed under unplanned purchases (definition E) but are not plausible common temptation goods and are in fact not listed as tempting or a waste of money (under definition A). Under definition C, respondents frequently regret buying “expensive food” (as opposed to just “food”, which was mentioned separately), presumably because they feel they paid too much for it.

We supplement the 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 (definition F). 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 and soda to their list (definition G).⁸

Table 4 presents summary statistics and correlations between the definitions of temptation spending we use in the analysis. The recorded level of temptation spending varies substantially based on the definition. 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 overlapping definitions. The low pair-wise correlations suggest that the different types of approaches are potentially capturing different aspects of the same concept or, alternatively, are mis-measuring the target concept in mostly uncorrelated ways. If the different approaches captured different aspects or if the approaches were unbiased but noisy measures of the same concept, the best approach empirically would be to aggregate the different measures. To the extent that one approach is capturing the concept of temptation spending better than others, we must rely on conceptual arguments about which definition seems most fitting.

Based on the discussion above on the items included under the different definitions, we suggest that temptation spending is best captured by self-reported aggregate money wasted (definition B, “Money wasted”), and to a lesser extent, purchases of goods in categories that the workers say they listed separately as often waste money on or are tempted to buy (definition A, “Waste/Temptation”). First, regretted purchases (definition C) and unplanned purchases (D and E) often capture other mistakes and deviations from plans that are not conceptually equivalent to being tempted into wasting money. Second, the common researcher-imposed definitions of

⁸ We categorize all fried sweet breads as donuts.

temptation spending (F and G) miss important categories of goods that the workers in our sample report being tempted into purchasing such as savory snacks, gifts for children, and clothing.

Our temptation measures show non-trivial average levels of temptation spending. Notably, definition B shows an average temptation spending level of MK306, which is 11 percent of average baseline income from Table 2 (MK2,680). Temptation spending under definition A is 4 percent of average income, while the other definitions are less than 3 percent.

Although we think the definitions A and B are the best measures of temptation spending, since we do not have a pre-analysis plan, we use all seven definitions to limit researcher degrees of freedom (see for example Simmons et al., 2011, who discuss how researcher choices over variable definitions can lead to false-positive findings.). We report our main analyses separately for each temptation spending measure, but 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 (non-market day payment) group, generating index values for the entire sample and normalizing to the control group's mean and standard deviation. The weights of the components of the index are calculated using the control group only so that variable construction is unaffected by the treatment. 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

Market days are a common institution in rural Africa. They bring together traders and customers at fixed locations and on a regular fixed schedule. The locations, called trading centers,

contain a few permanent 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, with one of the two days being a bigger event than the other. The largest market day in a given region is typically on Saturday. Vendors will travel long distances to the most-important market days, usually arriving the night before so they can set up their stalls early in the morning. As an illustration of what a market day in southern Malawi looks like, Appendix Figure 1 shows a picture of a market day at another trading center in the region, outside of our study site.⁹

In Malawi, market days are very tempting environments. They are typically lively, noisy affairs with many goods on offer, and salespeople trying to convince customers to browse their wares. Clothes are put out on display and vendors cook fragrant foods. The environment, with many sources of temptation, is a fairly stark contrast to ordinary days in rural Malawi. These temptations are also not completely avoidable: market days are often the only feasible option for people living in rural Malawi to buy common consumption goods. 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.

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

⁹ To protect the confidentiality of our research subjects' responses, we do not include any pictures of the participants in the study.

our field procedures and ensure that the experiment was feasible. Participants in that pilot reported that they found the market highly tempting.

Survey data from our sample of workers (Appendix Table 4) 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 variables¹⁰, indicators for the day-of-week of the exogenously-assigned (first attempted) interview date, baseline income, and the baseline values

¹⁰ The household financial variables are an index of physical assets 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.

of the outcome variable.¹¹ Finally, ε_{ir} is a mean-zero error term.¹²

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.

3 Effects of Receiving Pay in a Tempting Environment

We begin by showing that our intervention substantially shifted both when workers had cash on hand and the timing of expenditures. In Table 5, we show the effect of being paid on Saturday on workers' cash on hand before the wage payments (columns 1 and 2) and after the payments (columns 3 and 4). Prior to the wage payments, there is no appreciable difference in cash-on-hand on Fridays, whereas treatment-group workers have MK140 less on hand on Saturdays—presumably because the Friday-payment workers keep some of their pay with them and bring it back to the market the next day. The post-payment cash-on-hand results confirm that we did in fact induce treatment-group workers to have more cash during the market day: they have MK644 more cash on hand on Saturdays, and MK791 less on Fridays. Table 6 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

¹¹ We dummy out missing values of the baseline covariates.

¹² We cannot rule out potential spillovers between groups because workers in our sample can interact with each other. We can rule out any within-household spillovers as we restrict eligibility to a single person per household. We also show in Section 4 that there is no evidence of spillovers between workers who were next to one another in the payroll queue.

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 MK814 (PPP USD \$5.08) 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 7 shows the results of this analysis.¹³ 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 9%) and alcohol and tobacco (which rises by 23%). The other temptation spending measures show negative effects ranging from 3% to 22% 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

¹³ Appendix Table 5 shows the same results without any control variables. The results are substantively identical.

effects: the upper bound of the 95% confidence interval for the combined index across both rounds is 0.15 SDs.

Our results imply a null 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 (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 receiving cash in a more tempting environment 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 7 presents the minimum detectable effect (MDE) size on temptation spending at 80% power based on our estimated standard errors.¹⁴ We have 80% power to detect 0.17 to 0.28 SDs changes in temptation spending, and our MDEs for the aggregate index of temptation is 0.18 SDs. For our temptation spending measures, we find MDEs of 0.19

¹⁴ 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).

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.¹⁵ They find increases in candy consumption of nearly 100%. While this may be a more tightly-controlled and stronger intervention than the one we conducted, our study is arguably well-powered based on their estimated treatment effects.

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 6, 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

¹⁵ 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.

Along similar lines, our treatment simply might not have been intense enough to produce appreciable changes in behavior. The available evidence suggests this is somewhat unlikely. Appendix Table 4 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 days, 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).

While the survey responses above strongly suggest that getting paid on market days is a more tempting environment than getting paid the prior day, the receipt of the MK100 show-up stipend for both treatment and control on market days might have weakened the treatment. It is possible that workers from both study arms spent the show-up stipend on temptation goods, satisfying their demand (this could be the case, in particular, due a mental accounting effect that made the “unearned” income psychologically more available for spending). For all but two of our temptation spending measures the average in the control group is in fact smaller than the show-up stipend, which could support this view. However, the two measures for which average spending is higher than stipend amount are the ones that we argue are conceptually superior in Section 1.1.

More importantly, the show-up stipend is a very small amount of money: just 3 percent of wages for the lump-sum payment group and 12 percent for the weekly payments group. If temptation purchases are easily satisfied by such small amounts of spending, this finding still supports our overall conclusion that receipt of cash in a more tempting environment does not have practically meaningful effects on temptation spending.

A different way of thinking about treatment intensity is that even if our treatment was as strong as it could be, receiving cash in a tempting environment might simply have small effects on spending decisions. If so, our small treatment effects thus have important implications: they imply that even the most-tempting situation that a person could practically be exposed to will not strongly affect their behavior. This means that policymakers should worry less about the potentially temptation-enhancing aspects of a program's design, and that economists should focus, at the margin, on other explanations for low savings rates in developing countries.

Substitution bias

One specific reason our treatment might have been weaker than expected is “substitution bias.” Heckman et al. (2000) define substitution bias as an attenuation in the estimates of treatment effect that occurs because the control group gets access to the treatment or to a close substitute.¹⁶ In our case, the intervention exposed treatment-group workers to an environment that was fairly tempting, but control-group workers could have chosen to substitute toward other temptation spending opportunities. 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

¹⁶ “Substitution bias” also has a separate definition in the literature on price indices.

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 4: 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 day does not occur on Fridays, so this suggests that workers are visiting other markets for market day. There are three other major markets within easy walking distance of Mwanamulanje. None of them hold an official market day on Fridays, but all have some permanent vendors that are open every day. Moreover, workers may also travel longer distances to other markets that do hold market days on Fridays.

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 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. As a partial test of this possibility, we look at treatment effect heterogeneity by the time it would take workers to travel to another market. Specifically, we compute the perimeter of a triangular path that starts at each worker's home, goes to the payroll site, then to the closest alternative market to their home, and then back home again. This variable is included in our balance tables and there are no significant differences across study arms. We examine treatment effect heterogeneity by this variable in Appendix Table 6. There are no statistically-significant differences in the treatment effects for any of the outcome measures, so we find no evidence in favor of the substitution bias mechanism.

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.¹⁷ 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). The cash on hand results discussed above suggest some role for this mechanism, although Friday-payment workers do not bring large amounts of cash with them to the market.

Even if substitution bias does explain our null results, they have important theoretical implications. If people take intentional actions to seek out 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

¹⁷ 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.

sales in the United States (Cook 2007). It is unlikely that under-reporting would be systematically 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, two of our temptation measures allow respondents to self-designate spending as problematic (Rows A and B of Table 4), and are thus 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.

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 in order 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:

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

where $\bar{T}_{-i,r} = (T_{(i-1),r} + T_{(i+1),r})/2$ is the average treatment status¹⁸ of the workers $i-1$ and $i+1$ in round r , and takes the values 0, 0.5, or 1.¹⁹

We find no evidence that peer effects are driving our substantive results. The estimates of equation (3) in Appendix Table 7 reveal no statistically-significant effect of peers' treatment status on workers' temptation spending decisions.²⁰ 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 7. 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

¹⁸ 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 to several alternative specifications: using two workers ahead of and two workers behind in the line, controlling separately for the treatment statuses of the workers ahead and behind the focal worker, and controlling separately for the treatment statuses of the 10 workers ahead of the focal worker (results available upon request).

¹⁹ 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.

²⁰ One of the individual components has a significant difference at the 0.1 level; this is likely due to random sampling variation.

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.²¹ 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 2 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.

²¹ The use of temptation spending levels as a proxy for outside temptation exposure comes with the caveat that it may capture other factors as well. For example, people who differ in their baseline temptation spending may also have different willpower stocks.

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 directly that 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 most-plausible explanation, given the evidence is that the intensity of the treatment was low. However, our treatment was implemented well and captures one of the most-tempting situations that the people in our sample ever face. This suggests that the role exposure to temptation in driving consumption and savings choices in sub-Saharan Africa may be less of a concern than researchers and policymakers have heretofore thought. More broadly, the findings of our study imply 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 self-reported measure of temptation spending is eight times higher than the level of spending on tobacco and alcohol—and 11 times higher for

women.

A few factors limit the strength of the conclusions we can draw from our study. While the evidence best supports one of the mechanisms, we cannot convincingly rule out the other six. Also, although 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. 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.
- Brownback, A., Imas, A., & Kuhn, M. (2019). *Behavioral Food Subsidies* (Working Paper).
- 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.
- Imas, A., Kuhn, M., & Mironova, V. (2016). *Waiting to choose* (Working Paper)
- 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		Sun	Mon	Tue	Wed	Thu	Fri
	N	Fri	Sat	Fri	Sat	Fri	Sat	Fri	Sat						
Weekly Payments, Fridays	169	850	100	850	100	850	100	850	100						
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						
	$\Sigma=689$									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 nearest 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). Each worker worked for two weeks during each round of the program but collects their pay in four weekly installments. The weekly payment is MK700 in round one and MK800 in round two. All subjects received an MK100 show-up stipend. 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

	Lump Sum						Weekly						Fri. vs Sat. Bal Test.	
	Friday			Saturday			Friday			Saturday			Balance test	Balance test
	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	<i>p</i> -value	<i>p</i> -value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
<i>Background characteristics</i>														
Male	0.33	0.47	177	0.34	0.47	173	0.28	0.45	174	0.34	0.47	176	0.376	--
Married	0.70	0.46	176	0.66	0.48	169	0.70	0.46	171	0.73	0.44	172	0.879	--
Age (Years)	39.83	15.55	177	40.35	15.41	173	40.09	16.00	174	39.26	14.43	176	0.890	--
Years of Education Completed	3.60	3.36	176	3.58	3.11	172	3.24	3.08	173	3.66	3.08	173	0.408	--
Midline survey date (days after Sunday)	2.50	1.15	177	2.47	1.13	168	2.61	1.16	169	2.57	1.15	175	0.701	--
Prefers lump sum wage payments	0.73	0.45	177	0.72	0.45	172	0.72	0.45	173	0.77	0.42	176	0.520	--
Roundtrip Distance to Closest Two Markets (mi.)	12.30	1.12	153	12.33	1.12	147	12.20	1.25	149	12.27	1.18	147	0.592	--
<i>Financial outcomes (in units of MK unless noted)</i>														
Income received since past Friday	2,680	4,671	177	2,360	3,185	173	3,803	12,044	174	3,165	5,994	176	0.503	0.826
Remaining cash holdings out of income received	603	1,770	177	532	2,743	173	540	1,766	174	893	3,476	176	0.555	0.159
Total spending since Friday	3,955	4,588	177	3,290	3,319	173	4,027	5,148	174	3,553	4,116	176	0.144	0.137
Asset Ownership (PCA)	0.14	2.81	177	0.04	2.62	173	-0.31	2.09	174	0.13	3.10	176	0.410	--
Loans received in past month	3,416	9,886	177	2,435	7,091	173	1,790	4,714	174	4,121	17,798	176	0.524	0.283
Loans made in past month	800	3,441	177	571	1,645	173	562	1,636	174	993	4,211	176	0.571	0.678
Transfers received in past month	802	2,197	177	768	1,410	173	994	2,648	174	851	2,269	176	0.597	0.623
Transfers made in past month	779	2,322	177	464	1,360	173	509	2,493	174	633	2,562	176	0.421	0.634
<i>p</i> -value from joint significance of 15 covariates:	0.97						0.79						0.91	

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. Roundtrip Distance to Two Closest Markets measures the distance for the worker to walk from home to the payroll site, then to the closest other market to their home, and then back home again.

Table 3
Temptation Variable Construction, Advantages, and Disadvantages

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	A. Waste/ Temptation Goods	B. Money Wasted	C. Frequently Regretted	D. Unplanned Purchases	E. Commonly Against Plans	F. Alcohol and Tobacco	G. Alcohol, Tobacco, Donuts, and Soda
Based on respondent categorization	X	X	X	X	X		
Uses itemized spending within pre-specified categories	X		X	X	X	X	X
Timing of elicitation of categorization, by survey round	2	n/a	Baseline	n/a	Baseline	n/a	n/a
Data availability, by survey round	1 & 2	2 only	1 & 2	1 & 2	1 & 2	1 & 2	1 & 2
Advantages							
Individualized measure	X	X	X	X	X		
Focuses on harmful goods						X	X
Disadvantages							
Paternalistic						X	X
Only elicits most common categories of goods	X		X	X			
Can categorize non-tempt. purchases as tempt.				X [†]		X	X
Spending measure can include non-tempt. goods	X		X	X	X	X	X
Ex-post rationalization		X		X			
Undercounting of items in total		X					

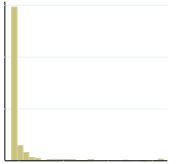
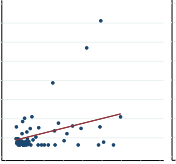
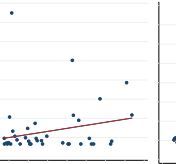
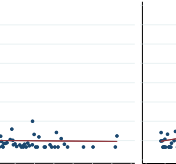
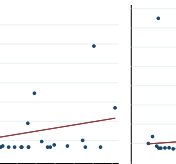
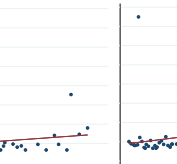
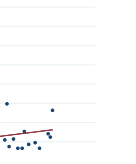
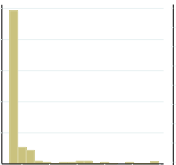
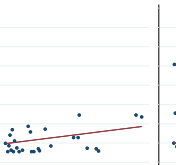
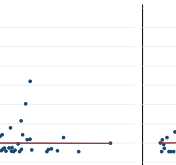
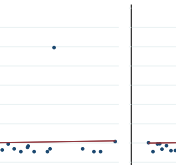
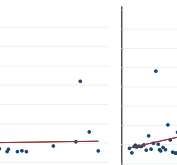
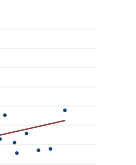
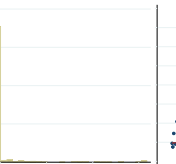
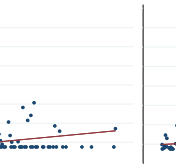
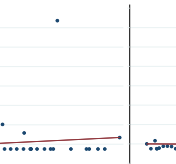
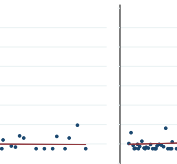
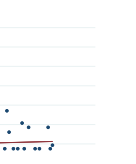

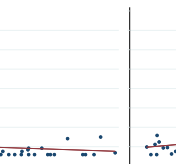
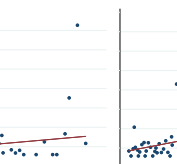
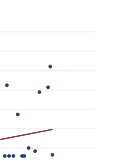
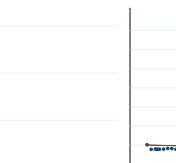
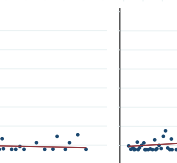
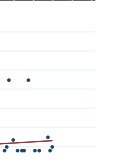
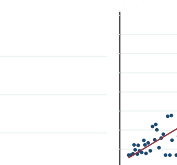
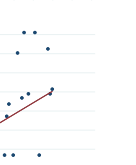

Notes: † Examples include health care spending and funerals.

Table 4

Comparison of Available Definitions of Temptation Spending

Panel A: Summary Statistics

Panel B: Pairwise Correlations

	Mean (MK)	SD (MK)	N	A. Waste/ Temptation Goods	B. Money Wasted [†]	C. Frequently Regretted	D. Unplanned Purchases	E. Commonly Against Plans	F. Alcohol and Tobacco	G. Alcohol, Tobacco, Donuts, and Soda
A. Waste/ Temptation Goods	116.55	343.91	689							
B. Money Wasted (round 2 only)	305.85	685.04	346	0.25						
C. Frequently Regretted	44.59	181.55	689	0.16	0.13					
D. Unplanned Purchases	49.57	121.38	689	-0.01	0.00	0.10				
E. Commonly Against Plans	62.21	287.56	689	0.16	0.02	0.05	-0.03			
F. Alcohol and Tobacco	14.33	46.79	689	0.07	0.01	-0.01	0.09	-0.02		
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.

Details for Panel B: All variables are standardized. Upper right triangle shows binned scatterplots, diagonal shows histograms, lower left triangle shows Pearson correlation coefficients.

Table 5
Effects of Market Day Wage Payment on Cash on Hand

<u>Dependent variable:</u>	(1)	(2)	(3)	(4)
	Cash on Hand Before Payment		Cash on Hand After Payment	
	Friday	Saturday	Friday	Saturday
Market Day Payment	-26.6 (20.8)	-140.4*** (24.5)	-791.2*** (22.0)	644.0*** (26.3)
Control-group mean	121.7	246.0	988.3	346.0
Number of observations	686	686	686	686

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 income, 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

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		
<u>Dependent variable:</u>	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)
<u>Panel A - Lump sum and weekly payment group pooled</u>						
Market Day Payment	-814.2*** (113.3)	-28.0 (89.5)	0.013 (0.029)	17.9 (194.9)	-93.6 (75.9)	127.1 (161.3)
Control-group mean	3,293.0	1,688.0	0.622	3,081.0	579.2	3,147.0
Number of observations	689	689	689.0	689	689	689
<u>Panel B - Lump sum payment group only</u>						
Market Day Payment	-758.4*** (171.6)	-23.7 (120.0)	0.038 (0.042)	161.6 (230.4)	-160.9 (107.0)	186.3 (237.0)
Control-group mean	3,068.0	1,247.0	0.534	3,753.0	670.6	3,341.0
Number of observations	345	345	345.0	345	345	345
<u>Panel C - Weekly payment group only</u>						
Market Day Payment	-810.9*** (153.5)	-26.1 (115.1)	-0.003 (0.039)	-3.9 (253.9)	-30.3 (114.7)	53.0 (234.3)
Control-group mean	3,530.0	2,151.0	0.714	2,378.0	483.5	2,944.0
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 income, 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
Effects of Market Day Wage Payment on Temptation Spending

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	PCA indices of temptation spending measures (columns 3 to 9)		Measures of temptation spending (MK)						
<u>Dependent variable:</u>	Omitting column 4	Including column 4	A. Waste/Temptation Goods	B. Money Wasted	C. Frequently Regretted	D. Unplanned Purchases	E. Commonly Against Plans	F. Alcohol and Tobacco	G. Alcohol, Tobacco, Donuts, and Soda
	1 & 2	2 only	1 & 2	2 only	1 & 2	1 & 2	1 & 2	1 & 2	1 & 2
Market Day Payment	-0.008	-0.129	-9.6	-49.2	4.0	-5.1	-14.9	3.0	-2.2
	(0.078)	(0.126)	(25.4)	(71.2)	(16.0)	(8.4)	(20.1)	(3.2)	(5.7)
Control-group mean	-0.012	0.000	131.2	324.3	42.7	52.7	69.1	12.9	67.2
Control-group SD	1.234	1.378	382.7	719.7	178.6	122.6	321.4	44.3	93.7
MDE on temptation spending at 80% power (SDs)	0.177	0.256	0.186	0.277	0.251	0.191	0.175	0.204	0.171
Number of observations	689	346	689	346	689	689	689	689	689

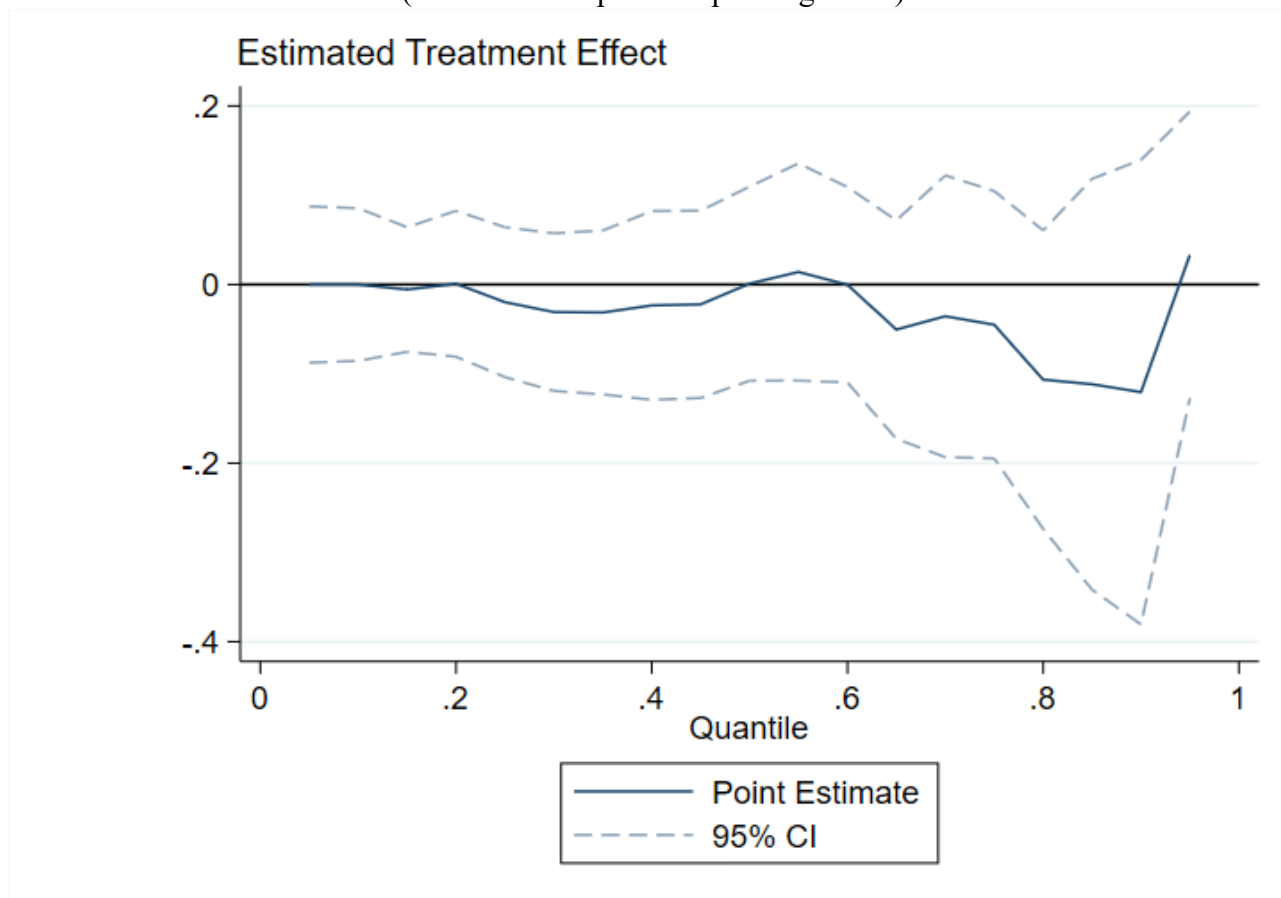
Notes: Sample includes 359 respondents who participated in at least one round of the work program and have data from at least one data source for that round (either the payday data, the survey, or both). The dependent variable in Column 1 is the PCA index constructed using data from both rounds, and thus it excludes the variable in Column (4) since it is only available in round 2. 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 income, 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
Example of a Market Day in Southern Malawi



Notes: Photograph of a market day in southern Malawi. To preserve the confidentiality of our subjects, this picture comes from a separate market in a nearby district that is not in our study area. Faces are blurred to protect privacy. Photo credit: Jason Kerwin.

Appendix Figure 2
 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

	Lump Sum						Weekly						Fri. vs Sat. balance test	
	Friday			Saturday			Friday			Saturday			<i>p</i> -value (13)	<i>p</i> -value (logs)† (14)
	Mean (1)	SD (2)	N (3)	Mean (4)	SD (5)	N (6)	Mean (7)	SD (8)	N (9)	Mean (10)	SD (11)	N (12)		
Panel A - Round 1														
Male	0.33	0.47	87	0.31	0.47	86	0.32	0.47	88	0.31	0.47	89	0.819	--
Married	0.71	0.46	87	0.62	0.49	84	0.74	0.44	86	0.71	0.46	87	0.214	--
Age (Years)	37.87	13.75	87	41.03	15.98	86	40.08	16.76	88	40.53	14.77	89	0.274	--
Years of Education Completed	3.65	3.42	86	3.63	3.06	86	3.44	3.24	87	3.38	2.95	88	0.899	--
Survey date (days after Sunday)	2.55	1.25	87	2.53	1.16	83	2.83	1.28	84	2.58	1.20	89	0.319	--
Prefers lump sum wage payments	0.78	0.42	87	0.67	0.47	86	0.75	0.44	87	0.74	0.44	89	0.238	--
Roundtrip Distance to Closest Two Markets (mi.)	12.27	1.15	78	12.37	1.08	71	12.17	1.30	78	12.30	1.13	71	0.408	--
Income received since past Friday	2,525	4,515	87	2,298	3,590	86	4,392	12,598	88	2,774	4,387	89	0.237	0.168
Remaining cash holdings out of income received	497	1,672	87	702	3,804	86	721	2,379	88	651	1,821	89	0.807	0.812
Total spending since Friday	3,967	4,894	87	3,154	3,717	86	4,351	5,389	88	3,352	2,948	89	0.052	0.110
Asset Ownership (PCA)	0.16	2.98	87	0.08	2.97	86	-0.18	2.33	88	-0.06	2.44	89	0.947	--
Loans received in past month	3,518	10,181	87	2,401	7,287	86	1,976	5,599	88	3,874	17,345	89	0.731	0.567
Loans made in past month	1,014	4,563	87	506	1,374	86	763	2,007	88	647	2,929	89	0.328	0.103
Transfers received in past month	1,028	2,896	87	717	1,135	86	932	2,477	88	737	1,798	89	0.279	0.492
Transfers made in past month	700	2,276	87	596	1,765	86	204	435	88	886	3,374	89	0.221	0.073
<i>p</i> -value from joint significance of 15 covariates:			0.75						0.63				0.38	
Panel 2 - Round 2														
Male	0.32	0.47	90	0.36	0.48	87	0.24	0.43	86	0.36	0.48	87	0.148	--
Married	0.70	0.46	89	0.69	0.46	85	0.65	0.48	85	0.75	0.43	85	0.303	--
Age (Years)	41.72	16.97	90	39.68	14.89	87	40.10	15.28	86	37.95	14.04	87	0.197	--
Years of Education Completed	3.56	3.33	90	3.53	3.16	86	3.05	2.91	86	3.95	3.20	85	0.199	--
Midline survey date (days after Sunday)	2.46	1.05	90	2.41	1.09	85	2.39	0.99	85	2.56	1.09	86	0.582	--
Prefers lump sum wage payments	0.68	0.47	90	0.77	0.42	86	0.70	0.46	86	0.80	0.40	87	0.036	--
Roundtrip Distance to Closest Two Markets (mi.)	12.33	1.10	75	12.30	1.16	76	12.22	1.20	71	12.24	1.22	76	0.968	--
Income received since past Friday	2,829	4,838	90	2,421	2,746	87	3,201	11,492	86	3,565	7,286	87	0.982	0.295
Remaining cash holdings out of income received	706	1,863	90	364	830	87	356	695	86	1,140	4,591	87	0.427	0.023
Total spending since Friday	3,944	4,298	90	3,424	2,888	87	3,696	4,898	86	3,758	5,050	87	0.618	0.552
Asset Ownership (PCA)	0.13	2.65	90	0.00	2.24	87	-0.45	1.82	86	0.31	3.66	87	0.284	--
Loans received in past month	3,317	9,648	90	2,469	6,933	87	1,599	3,617	86	4,374	18,347	87	0.428	0.345
Loans made in past month	594	1,791	90	636	1,881	87	356	1,113	86	1,347	5,200	87	0.106	0.294
Transfers received in past month	583	1,159	90	819	1,642	87	1,056	2,825	86	967	2,672	87	0.737	0.984
Transfers made in past month	854	2,377	90	332	766	87	821	3,501	86	375	1,252	87	0.042	0.258
<i>p</i> -value from joint significance of 15 covariates:			0.39						0.28				0.18	

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. Roundtrip Distance to Two Closest Markets measures the distance for the worker to walk from home to the payroll site, then to the closest other market to their home, and then back home again.

Appendix Table 2
Temptation Questions

A. Expenditures that are wastes of money or that respondent 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 frequently 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 list)

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	

Notes: 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 3

Self-reported Temptation Goods, by Definition of Temptation

	(1) Share reporting this good	(2)	(3)	(4)	(5)	(6)	(7)
	Share reporting this good			Share purchasing this good			
	A. Waste/ Temptation Goods	C. Frequently Regretted	E. Commonly Against Plans	B. Money Wasted [†]	D. Unplanned Purchases	F. Alcohol and Tobacco	G. Alcohol, Tobacco, Donuts, and Soda
Savory snacks	0.44	0.18	0.14	0.09	0.31	0.00	0.00
Gifts for children	0.15	0.03	0.01	0.01	0.00	0.00	0.00
Clothes or shoes	0.14	0.13	0.07	0.07	0.01	0.00	0.00
Food	0.11	0.06	0.13	0.51	0.49	0.00	0.00
Sweets	0.11	0.02	0.02	0.16	0.22	0.00	0.82
Alcohol	0.06	0.04	0.03	0.00	0.02	0.09	0.09
Illness/Medicine	0.03	0.01	0.20	0.03	0.00	0.00	0.00
Other gifts	0.01	0.01	0.03	0.01	0.15	0.00	0.00
Expensive food	0.00	0.07	0.00	0.00	0.00	0.00	0.00
Bad business decisions	0.00	0.00	0.05	0.00	0.00	0.00	0.00
Funerals	0.00	0.00	0.05	0.00	0.02	0.00	0.00
Tobacco	0.00	0.00	0.00	0.01	0.01	0.18	0.18
Other	0.11	0.16	0.09	0.10	0.16	0.00	0.00
N	346	349	349	346	359	359	359

Notes: Columns 1-3 show the share of respondents who mentioned the good as a temptation good based on each definition. Columns 4-7 show the share of respondents for whom the good was listed as one of the temptation goods they purchased according to each definition. Column 4 includes data just for round 2; columns 5-7 are for any purchase across rounds 1 and 2.

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 4

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 in past month

	Mean	Obs.
Number of visits	6.08	365
Among respondents with any visit:		
% 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 5
Effects of Market Day Wage Payment on Temptation Spending (Without Controls)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	PCA indices of temptation spending measures (columns 3 to 9)			Measures of temptation spending (MK)					
<u>Dependent variable:</u>	Omitting column 4	Including column 4	A. Waste/ Temptation Goods	B. Money Wasted	C. Frequently Regretted	D. Unplanned Purchases	E. Commonly Against Plans	F. Alcohol and Tobacco	G. Alcohol, Tobacco, Donuts, and Soda
	1 & 2	2 only	1 & 2	2 only	1 & 2	1 & 2	1 & 2	1 & 2	1 & 2
Market Day Payment	-0.034 (0.089)	-0.167 (0.134)	-29.2 (27.4)	-55.0 (71.8)	5.0 (16.0)	-5.6 (8.6)	-13.9 (19.9)	2.7 (3.6)	-3.2 (6.5)
Control-group mean	-0.012	0.000	131.2	324.3	42.7	52.7	69.1	12.9	67.2
Control-group SD	1.234	1.378	382.7	719.7	178.6	122.6	321.4	44.3	93.7
MDE on temptation spending at 80% power (SDs)	0.201	0.272	0.201	0.279	0.250	0.197	0.173	0.231	0.194
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). The dependent variable in Column 1 is the PCA index constructed using data from both rounds, and thus it excludes the variable in Column (4) since it is only available in round 2. 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 income, 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 Table 6
Treatment Effect Heterogeneity by Roundtrip Distance to Two Closest Markets

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	PCA indices of temptation spending measures (columns 3 to 9)			Measures of temptation spending (MK)					
<u>Dependent variable:</u>	Omitting column 4	Including column 4	A. Waste/ Temptation Goods	B. Money Wasted	C. Frequently Regretted	D. Unplanned Purchases	E. Commonly Against Plans	F. Alcohol and Tobacco	G. Alcohol, Tobacco, Donuts, and Soda
	1 & 2	2 only	1 & 2	2 only	1 & 2	1 & 2	1 & 2	1 & 2	1 & 2
Rounds Available									
Market Day Payment	0.318	0.498	-22.6	44.3	24.2	-13.3	-52.8	20.2	18.4
	(0.492)	(0.606)	(107.9)	(419.3)	(72.4)	(41.9)	(95.7)	(22.1)	(36.9)
<i>q</i> -value	0.916	0.916	0.916	0.916	0.916	0.916	0.916	0.916	0.916
Market Day Payment X Roundtrip Distance to Two Closest Markets	0.071	0.146	-1.1	27.6	4.8	-2.7	-7.2	3.8	4.3
	(0.111)	(0.129)	(22.3)	(95.1)	(15.9)	(9.2)	(19.7)	(5.1)	(8.3)
<i>q</i> -value	0.869	0.869	0.962	0.869	0.869	0.869	0.869	0.869	0.869
Number of observations	573	281	573	281	573	573	573	573	573

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). The dependent variable in Column 1 is the PCA index constructed using data from both rounds, and thus it excludes the variable in Column (4) since it is only available in round 2. 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 income, baseline total spending and (if available) the baseline value of the outcome variable. Roundtrip Distance to Two Closest Markets measures the distance for the worker to walk from home to the payroll site, then to the closest other market to their home, and then back home again. Heteroskedasticity-robust standard errors, clustered by worker, in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table 7
Peer Effects on Temptation Spending

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	PCA indices of temptation spending measures (columns 3 to 9)		Measures of temptation spending (MK)						
<u>Dependent variable:</u>	Omitting column 4	Including column 4	A. Waste/ Temptation Goods	B. Money Wasted	C. Frequently Regretted	D. Unplanned Purchases	E. Commonly Against Plans	F. Alcohol and Tobacco	G. Alcohol, Tobacco, Donuts, and Soda
	1 & 2	2 only	1 & 2	2 only	1 & 2	1 & 2	1 & 2	1 & 2	1 & 2
Market Day Payment	-0.011 (0.081)	-0.154 (0.131)	-9.5 (26.2)	-57.0 (71.0)	0.4 (15.8)	-5.7 (8.5)	-16.6 (20.8)	3.1 (3.3)	-2.1 (5.8)
Average Peer Treatment Status	-0.012 (0.122)	-0.294 (0.192)	18.6 (39.1)	-90.6 (105.7)	-16.0 (21.4)	23.3* (13.1)	-21.0 (33.3)	-2.8 (4.6)	-4.1 (8.8)
Dependent variable mean, control group (Friday paydays)	-0.0210	0.000	131.4	324.3	42.90	52.90	69.50	12.80	66.0
Control-group SD	1.228	1.378	383.7	719.7	179.1	122.9	322.3	44.3	90.8
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). The dependent variable in Column 1 is the PCA index constructed using data from both rounds, and thus it excludes the variable in Column (4) since it is only available in round 2. 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 income, baseline total spending and (if available) the baseline value of the outcome variable. Heteroskedasticity-robust standard errors, clustered by worker, in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.