The Influence of Budgets on Consumer Spending

Marcel F. Lukas

Ray Charles "Chuck" Howard

Forthcoming, Journal of Consumer Research

Marcel F. Lukas (mfl3@st-andrews.ac.uk) is a Lecturer in Banking and Finance, School of Management, University of St. Andrews, The Gateway, North Haugh, St Andrews, KY16 9RJ, United Kingdom. Ray Charles "Chuck" Howard (corresponding author, rhoward@mays.tamu.edu) is an Assistant Professor of Marketing, Mays Business School, Texas A&M University, 4113 TAMU | 210 Olsen Blvd, College Station, Texas, 77843-4113. This article is based on the dissertation work of both authors. For valuable feedback on this research, the authors thank Dale Griffin, David Hardisty, Abigail Sussman, Johannes Boegershausen, Gizem Yalcin, participants at the 2019 Boulder Summer Conference on Consumer Financial Decision Making, and their JCR review team. For financial support, the authors thank the Dean's Office at Mays Business School. Supplementary materials are included in the web appendix accompanying the online version of this article.

Running Head: LUKAS AND HOWARD

Editors: Amna Kirmani and Andrew T. Stephen

Associate Editor: Ajay Kalra

ABSTRACT

Foundational research in marketing and behavioral economics has revealed a great deal about the psychology of budgeting. However, little is known about the extent to which budgets do (or do not) influence consumers' real-world spending. The present research addresses this gap in the literature using naturally occurring budgeting and spending data provided by a popular personal finance app in the UK, a field experiment conducted with members of a Canadian credit union, and a financial diary study conducted with consumers in the US. Budget compliance is generally weak because budgets are wildly optimistic. However, optimistic budgets do help consumers reduce their spending. Moreover, the influence of budgets on spending is surprisingly sticky: consumers continue to reduce their spending six months after setting a budget, even though spending remains over-budget. Impulsive consumers exhibit worse budget compliance than less impulsive consumers. However, counterintuitively, this is predominately because more impulsive consumers set lower budgets than less impulsive consumers, not because they spend more. Finally, we provide evidence that budgets influence spending across several theory-informing psychographic variables. Taken together, these findings show that budgets can be both wildly optimistic and highly influential, and that beliefs about the nature of consumers' budgets require updating.

Keywords: budgeting, mental accounting, reference points, planning fallacy, optimism, impulsiveness, temporal discounting

"If you and your family want financial security, following a budget is the ***only*** *answer."*

- Investopedia, the world's leading source of financial content on the web (Bell 2019)

"Budgeting doesn't work."

- Bill Harris, Former CEO of Intuit, parent company of Mint.com, the USA's largest personal finance service (Olen 2015)

Budgeting is commonly thought to be a practical and effective way for consumers to manage their money and improve their financial well-being. For example, financial advisors frequently invoke budgeting as a way for consumers to curb overspending, and therefore increase savings, decrease debt, and improve financial security (Bell 2019; Caldwell 2019; Credit Counselling Society 2019). However, a growing chorus of voices has begun to argue that budgeting simply doesn't work (Browne 2020; Elkins 2018; Mathera and Juarez 2021; Olen 2015). Surprisingly, the academic literature on consumer budgeting has little to offer this debate: for all we know about the psychology of budgeting, we know very little about the extent to which budgets do (or do not) influence spending in the wild (Zhang and Sussman 2018). The primary goal of the present research is to address this gap in the literature.

Our examination of budget influence utilizes two distinct modes of inquiry. In Study 1, we use naturally occurring data provided by a popular personal finance app to present a rich descriptive analysis of the relationship between consumers' self-generated budgets and their real-world spending behavior. The novel structure of the data allows us to observe each user's pre-budget spending (i.e., how much they spend *before* they set a budget), their budget, and their post-budget spending across three qualitatively different budget categories: dining and drinking, groceries, and fuel. The data also let us compare the spending behavior of consumers who use the app to set budgets against the spending behavior of similar consumers who use the app to track their expenses but do not set budgets. We then supplement this descriptive analysis with a pair of experiments that examine the causal influence of budgets on spending by tracking consumers' spending after they have been randomly assigned to set a budget or not (Study 2), or to make a relatively high or low budget (Study 3). We also use Study 2 to explore the relationship between budget setting, spending, and several psychographic variables of theoretical interest. Taken together, this multi-method approach contributes a deeper understanding of how budgets influence spending over time, between categories, and across psychographics. Our core finding is that budgets can be simultaneously wildly optimistic and highly influential.

The remainder of this article unfolds as follows: we first review the theory and evidence upon which we develop our hypotheses. We then present our studies. Finally, we discuss the implications of our work for theory and practice, and outline directions for future research.

THE BUDGETING PROCESS

Budgeting is a two-stage process. In stage one, consumers set a budget by earmarking a specified amount of money for a particular expense category over a predetermined length of time. For example, a consumer may set a grocery budget of $300 per month. In stage two, consumers track their spending against their budget. This requires that expenses be both "booked" (noticed) and "posted" (assigned) to the correct expense category (Heath 1995; Heath and Soll 1996; Thaler 1985).

In the present research we focus on the influence of budget setting on real-world spending, while controlling for expense tracking frequency. This focus is motivated by two observations. First, research on consumer budgeting has predominately investigated psychological factors that influence how consumers set budgets (e.g., Peetz and Buehler 2012; Ülkümen, Thomas, and Morwitz 2008) or how they track expenses (e.g., Cheema and Soman 2006; Heath and Soll 1996). In contrast, much less is known about the extent to which budgets influence real-world spending once they are set (Zhang and Sussman 2018). Second, budgeting apps have made expense tracking relatively easy, but setting a realistic budget remains a challenge for many consumers, even when they know how much money they have spent in the recent past (Howard et al. 2022; Peetz and Buehler 2009; see also Buehler, Griffin, and Peetz 2010). Therefore, focusing on the effect of budget setting on real-world spending fills an important gap in the consumer budgeting literature, and it can better inform consumer budgeting in practice.

We next discuss why consumers set budgets, and the factors that influence how much money they earmark for a budget.

THE PSYCHOLOGY OF BUDGETING

Consumers set budgets for many reasons. Often, budgets are set in an attempt to minimize or control spending. For example, consumers can use budgets as a self-control device that helps them avoid frivolous expenses (Thaler 1999; Thaler and Shefrin 1981), a way to meet a savings goal (Peetz and Buehler 2009; Soman and Cheema 2004), or a tool to prioritize certain expenses over others (Fernbach, Kan, and Lynch 2015). However, emerging evidence suggests that consumers sometimes set budgets in an attempt to maximize their spending, such as when they budget for gifts (Choe and Kan 2018). Consumers may also set budgets simply to gain an accurate view of their future expenses, independent of any desire to minimize or maximize their spending (Howard et al. 2022; Peetz et al. 2016).

In addition to the many reasons *why* consumers set budgets, there are numerous factors that influence *how much* money they budget. Ülkümen, Thomas, and Morwitz (2008) present a metacognitive account of budget setting that proposes the amount of money consumers budget is inversely related to the amount of confidence they have in their budget accuracy. Peetz and Buehler (2009) explain budget setting with a motivated cognition framework in which the amount of money consumers budget is negatively associated with how strongly they endorse a savings goal. Peetz and Buehler (2012) show that construal level can influence the amount of money consumers budget, such that a more abstract construal leads to higher budgets and a more concrete construal leads to lower budgets. The amount of money consumers budget may also be influenced by cognitive accessibility, meaning the ease or difficulty with which certain expenses come to mind when a budget is being set (Howard et al. 2022; Sussman and Alter 2012).

Given the multitude of factors that influence why consumers set budgets and how much money they budget, it is reasonable to ask *what exactly is a budget?* This is the question we address next.

BUDGETS AS REFERENCE POINTS

Academics and practitioners have defined budgets variously as estimates (Ganti 2021; Ülkümen, Thomas, and Morwitz 2008), plans (Credit Counseling Society 2021; Novemsky and

Kahneman 2005), predictions (InCharge.org 2021; Peetz and Buehler 2009), and targets (Caldwell 2019; Soman and Cheema 2004). Although related, each of these definitions is distinct. Estimates are "an approximate calculation," plans are "an intention or decision about what one is going to do," predictions try to correctly "make known beforehand" what will happen in the future, and targets are an "an objective or result toward which efforts are directed" (Oxford English Dictionary 2021). This variability in how past research defines budgets accurately reflects the reality that budgets can represent different things to different people. However, it also begs the question, *is there a superordinate construct that captures the different conceptualizations of what a budget represents?*

We believe the answer to this question is 'yes.' Specifically, we propose that regardless of why or how a budget is created, it subsequently becomes a *reference point* or benchmark against which future spending can be tracked (also see Choe and Kan 2021). This proposition is consistent with past research showing that reference points are derived from many different constructs, including expectations (Abeler et al. 2011; Kőszegi and Rabin 2006), plans (Allen et al. 2016), and goals (Heath, Larrick, and Wu 1999). It is also consistent with work demonstrating that once a numeric value is brought to mind it can strongly influence a consumer's decision making, regardless of why or how the value was generated. For example, research on "arbitrary anchors" has shown that consideration of random numbers like the last two digits of a person's social security number changes willingness to pay for consumer goods (Ariely, Loewenstein, and Prelec 2003). Similarly, research on "mere goals" has shown that generating a subjective goal like being able to do 25 push-ups (why not 24 or 26?!) serves as a reference point that influences subsequent behavior (Heath, Larrick, and Wu 1999).

We next discuss the optimistic nature of consumers' budgets.

BUDGET OPTIMISM

We define budget optimism as the difference between average past spending and budgeted future spending. So, for example, if a consumer spends an average of $250 per month on entertainment over the past three months, then sets a budget of $150 for next month, their degree of budget optimism is quantified as $100 or 40%. This definition of budget optimism follows from the general definition of optimism as an expectation that the future will be better than the past (Sharot 2011), because in the context of budgeting "better" is commonly thought to mean spending less and saving more (Bell 2019; Caldwell 2019; Credit Counselling Society 2019). This definition of budget optimism is also consistent with the argument that the most realistic predictor of future behavior is often the mean of relevant past behavior (Kahneman and Tversky 1979), because the belief that future spending will be better than average also matches the general definition of optimism.

Importantly, past research offers many reasons to believe that budgets tend to be optimistic. Mental budgeting theory argues that budgets are often set as a self-control device (Heath and Soll 1996). This suggests that budgets will be optimistic because the need to exercise self-control implies that certain past expenses should be avoided in the future. Research on planning fallacies also indicates that budgets will be optimistic, because budgets are often a plan for future spending (Lynch et al. 2010; Novemsky and Kahneman 2005), and plans do not adequately incorporate all relevant past behavior (Buehler, Griffin, and Ross 1994). Following a similar logic, research on expense prediction has shown that consumers predict their spending will be lower in the future than it was in the past, even when they are fully aware of how much they spent in the past at the time of prediction (Howard et al. 2022). This again suggests that budgets will be optimistic, because budgets are at least partially a prediction of future spending (Peetz and Buehler 2009).

We next discuss competing hypotheses regarding the relationship between budget optimism and budget compliance.

BUDGET COMPLIANCE

We define budget compliance as the difference between budgeted spending for a given month and the amount of money a consumer subsequently spends during that month. So, for example, if a consumer budgets $150 for entertainment next month, then ends up spending $200, their budget compliance is quantified as −$50 or −33.3%.

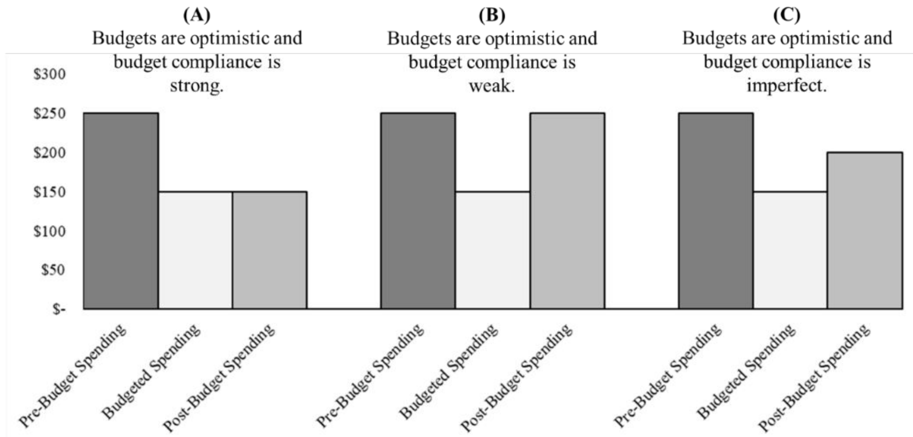
Mental budgeting theory asserts that when expenses are easily booked and posted – as is the case when using a personal finance app – budgets will be "inflexible" (Heath and Soll 1996). This characterization of budgets is most directly supported by lab studies showing that budgets constrain individuals' spending and investment decisions by decreasing their perceived disposable income (Heath 1995; Heath and Soll 1996). Taken together, these findings lead to the hypothesis that consumers will generally comply with their budgets (**H1a**). Given the expectation that budgets are optimistic, H1a implies that consumers will achieve compliance by lowering their spending to the amount they have budgeted. This possibility is illustrated in Panel A of Figure 1.

In contrast, research on planning fallacies and prediction biases suggests that optimistic budgets will only result in weak budget compliance. For example, studies have shown that consumers predict they will spend less in the future than in the past, but on average they end up spending approximately the same amount as before (Howard et al. 2022; Peetz and Buehler 2009). This pattern of results has also been observed for related phenomena like project completion times: planned project completion times are highly optimistic as compared to similar past projects, but actual completion times often end up being no different than past projects (Buehler, Griffin, and Peetz 2010). Taken together, these findings lead to the hypothesis that consumers will spend significantly more than they budget (**H1b**). This possibility is illustrated in Panel B of Figure 1.

Note that the budget compliance hypothesis offered by mental budgeting theory implies that budget optimism can be a good thing; if budgets are inflexible, then a lower budget will lead to lower spending and higher saving. However, research on planning fallacies and prediction biases suggests that budget optimism should be avoided because it leads to inaccurately low budgets, and ultimately, weaker budget compliance. This is highlighted by the fact that most interventions in the planning fallacy and expense prediction literatures are designed to remove optimism by increasing plans and predictions (e.g., Buehler, Griffin, and Ross 1994; Howard et al. 2022; Peetz and Buehler 2012; Peetz et al. 2015; Putnam-Farr and Ghosh 2018).

In Panel C of Figure 1 we illustrate a third possibility, which is that consumers' budgets influence their spending even when budgets are overly optimistic and budget compliance is imperfect. We next discuss this possibility in greater detail.

FIGURE 1: POSSIBLE OUTCOMES REGARDING PRE-BUDGET SPENDING, BUDGETED SPENDING, AND POST-BUDGET SPENDING



NOTES: **Budget optimism** is the difference between pre-budget spending and budgeted spending. The greater the difference, the more optimistic a budget is said to be. **Budget compliance** is the difference between budgeted spending and post-budget spending. Strong budget compliance is post-budget spending ≈ budgeted spending (Panel A); the greater the difference between budgeted spending and post-budget spending the weaker compliance is said to be.

BUDGET INFLUENCE

Whether or not consumers strictly comply with their budgets, it is theoretically possible for budgets to influence their spending. To illustrate this possibility, consider a consumer who typically spends $250 per month on entertainment, sets a budget of $150 for the next month, then ends up spending $200. All else equal, this consumer's budget has influenced their spending even though they have not strictly complied with their budget.

The possibility that budgets influence spending even when consumers exceed their budget leads to several hypotheses. The first is that lower, more optimistic budgets are associated with lower spending in the target week or month (**H2**). As an example, consider two consumers named Jane and John who both spend an average of $250 per month on entertainment. Now imagine that Jane sets an entertainment budget of $150 for the month of August while John sets a budget of $200. At the end of the month Jane has spent $175 on entertainment and John has spent $200. All else equal, Jane's more optimistic budget is associated with lower spending than John's less optimistic budget, *even though Jane's budget compliance is worse.*

To better understand the nature of budgets it is also informative to consider their longitudinal influence on spending. For example, what will happen to Jane and John's spending behavior in September, given their spending behavior in August? Past research suggests that budgets should exert less influence on spending over time (Choe and Kan 2021), and that spending more than budget in one month may actually lead consumers to abandon their budget altogether in the next month (Soman and Cheema 2004). In tandem, these findings lead to two longitudinal hypotheses. The first is that the association between budgets and spending weakens over time (**H3a**). The second is that weaker budget compliance in one month is associated with higher spending in the next month (**H3b**).

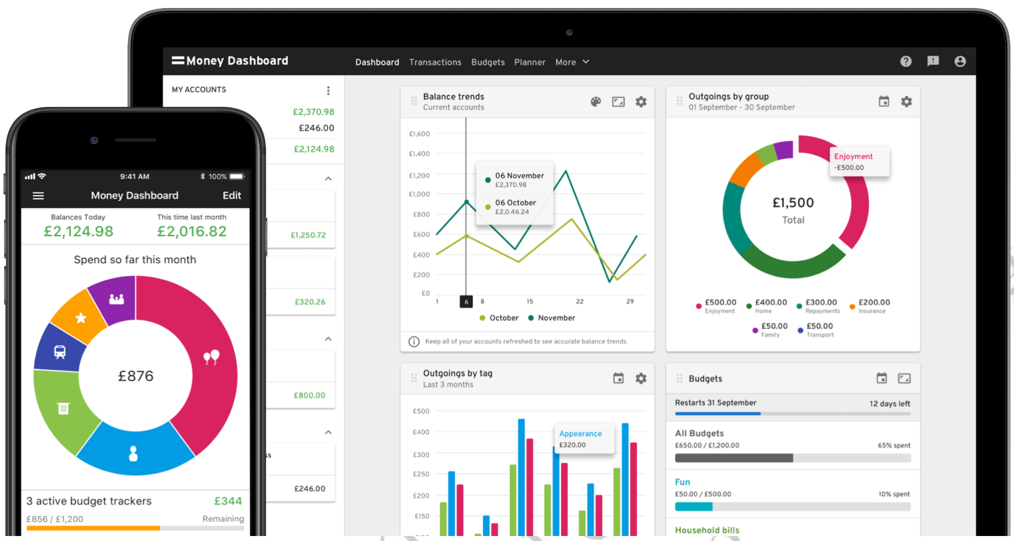
The final hypothesis we test regarding budget influence is that consumers who set a budget subsequently spend less money than consumers who do not set a budget (**H4**). This follows logically from the expectation that budgets are both optimistic and influential. In contrast, if budgets are merely optimistic, as some past research implies, then similar consumers should display similar spending whether one has set a budget or not.

We next test H1−H4 in Study 1.

STUDY 1: PERSONAL FINANCE APP DATA

In Study 1 we test H1–H4 using data provided by Money Dashboard (MDB), a personal finance app in the UK with approximately 70,000 active users at the time of writing. The primary function of MDB is to provide users with a holistic overview of their financial situation. To accomplish this, MDB collects and combines all transactional information across all financial accounts for each user. So, for example, if a user has two credit cards, a chequing account, and a savings account across different providers, MDB will aggregate all inflows and outflows across these cards and accounts and present the user with up-to-the-minute information on when, where, and how they are spending their money. FIGURE shows the user interface for the mobile and web application.

FIGURE 2: MONEY DASHBOARD INTERFACE



Spending Data

The data set includes all user transactions – more than 350 million – between January 2014 and December 2016. Each transaction is automatically assigned to a spending category by MDB (e.g., “Groceries”), and includes a merchant tag (e.g., “Tesco”) and time stamp. One novel feature of the data is that MDB’s terms of service allow us to observe each user’s transaction history for at least three months before they download the app and set their first budget.

Cash spending represents 2.90% of total spending in our sample, and it appears in the dataset as ATM withdrawals. So, although we do not observe exactly what users spend their cash on, we do observe exactly how much they withdraw. This allows us to estimate cash spending in each budget category using consumer spending data from the UK Office for National Statistics (ONS). For example, we use the ONS data to estimate cash spending on groceries by (a) calculating the proportion of total cash spending that the average UK consumer dedicates to grocery items (e.g., milk, eggs, bread, and so on), and (b) multiplying the amount each MDB user withdrawals from ATMs month by that proportion.

Budgeting Data

The MDB budgeting function allows users to set budgets for expenses in multiple categories. For example, a user may set a monthly budget of ₤150 for dining and drinking, ₤300 for groceries, and/or ₤100 for fuel. (Most consumers in the dataset budget for a single category. See Table 1.) MDB then automatically tracks transactions in these categories and allows users to observe their spending against their budget. Users do not receive any push notifications about their budget compliance, and they have to manually login to track their expenses. However, when they do login, they are presented with a very salient illustration of their budget and remaining funds in each category, as shown in FIGURE .

FIGURE 3: MONEY DASHBOARD BUDGET INTERFACE

In total, the dataset includes 9,292 budgets set by 5,696 users for whom we can observe three months of pre-budget spending, budgeted spending, and six months of post-budget spending. We focus our analysis on the three most popular budget categories: Dining and Drinking (n = 2,190), Groceries (n = 2,415), and Fuel (n = 1,011).[1] Table 1 summarizes users' demographic, budgeting and account data.

[1] Note: The next most popular budget category is clothing (n = 399). Our hypotheses cannot be tested in the clothing category because of missing data (most consumers in the dataset do not spend money on clothes every month). Beyond clothing, the remaining budget categories in the dataset are highly idiosyncratic (e.g., "Domestic supplies", n = 11; "Hobbies or Activities", n = 10; "Vehicle Running Costs", n = 2), which means (a) few people set budgets in these categories, and (b) it is not possible to accurately map spending in these categories to the corresponding budget. These issues do not apply to dining and drinking, groceries, and fuel, which is why we focus our analyses on these categories. In Study 2 we examine total discretionary spending, and in Study 3 we examine total spending.

TABLE 1: MONEY DASHBOARD USER PROFILE STATISTICS

	Mean	Median	St. Dev.
Age	36.2	34	9.67
Annual salary (£ 000's)	28.4	25	14.10
Logins per month	5.8	4	6.41
# of budget categories per user	1.63	1	0.71
# of accounts linked per user	4.69	4	3.54
% Male	68.2%	N/A	0.47
% England	85.6%	N/A	0.35
% Scotland	9.6%	N/A	0.30
% Wales	4.8%	N/A	0.21

Analysis and Results

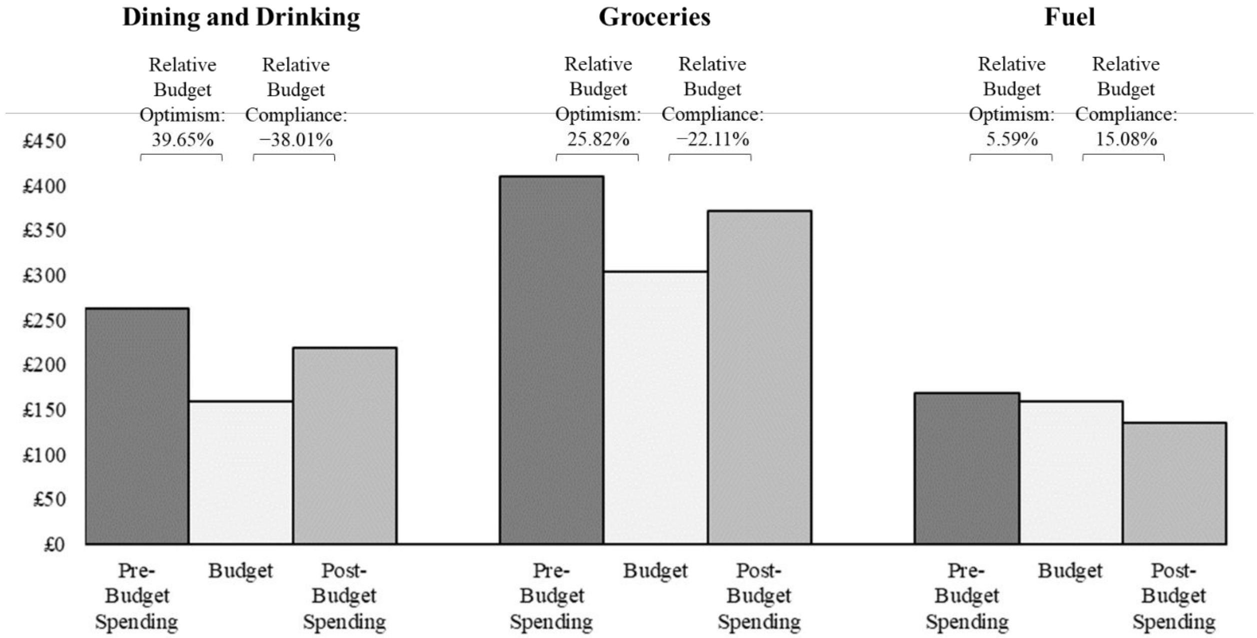
Budget Optimism. We quantify budget optimism as the difference between a consumer's mean monthly spending in a given category over the three months *before* they set a budget, and the amount they subsequently budget for that category. A set of within-subject t-tests confirm the expectation that budgets are generally optimistic. However, the relative degree of budget optimism varies between categories. On average, budgeted spending is 39.65% lower than mean pre-budget spending for dining and drinking (Mean difference = £104.10, SD = 256.53, $t(2,189) = 18.98$, $p < .001$), 25.82% lower for groceries (Mean difference = £105.83, SD = 272.58, $t(2,414) = 19.08$, $p < .001$), and 5.59% lower for fuel (Mean difference = £9.43, SD = 117.56, $t(1,010) = 2.5510$, $p = .011$). Therefore, if more optimistic budgets are associated with weaker budget compliance, as the planning fallacy suggests, then it should be the case that compliance is weakest for dining and drinking and strongest for fuel.

Hypothesis 1: Budget Compliance. We quantify budget compliance as the difference between a consumer's budget for a given category and the amount they spend in that category during the first month after the budget is set. Within-subject t-tests reveal that actual spending is

38.01% higher than budgeted spending for dining and drinking (Mean difference = £60.22, SD = 204.73, $t(2,189) = 13.77$, $p < .001$), 22.11% higher than budgeted spending for groceries (Mean difference = £67.22, SD = 260.98, $t(2,414) = 12.66$, $p < .001$), and 15.08% lower than budgeted spending for fuel (Mean difference = -£24.02, SD = 120.33, $t(1,010) = 6.35$, $p < .001$).

To make relative budget optimism and compliance easily comparable, Figure 4 plots mean pre-budget spending, budgeted spending, and post-budget spending for each of the three categories. The descriptive results for dining and drinking and groceries bear a striking resemblance to Panel C in Figure 1, whereas the results for Fuel more closely resemble Panel A in Figure 1. To test the robustness of these results we examined the percentage change in the price of goods in each budget category during our observation period (2014 to 2016), as reported by the UK Office of National Statistics. The change in the price of goods related to dining and drinking and groceries was negligible: prices related to dining and drinking rose by 2.00% and prices related to groceries fell by 2.53%. However, the price of fuel fell dramatically, by 15.59%, which is remarkably close to the difference between the amount consumers budgeted and spent on fuel (15.08%). This suggests that if the price of fuel had not fallen serendipitously, then budgeted and actual spending on fuel would be approximately equal, and the results for fuel would even more closely match the hypothetical outcome illustrated in Panel A of Figure 1.

FIGURE 4: MEAN PRE-BUDGET SPENDING, BUDGET, AND POST-BUDGET SPENDING IN EACH CATEGORY OF STUDY 1



NOTES: Relative budget optimism is calculated as (mean pre-budget spending − budget)/(mean pre-budget spending). Relative budget compliance is calculated as (budget – post-budget spending)/(budget). All statistical tests for budget optimism and compliance were performed using absolute mean differences. We present relative differences (i.e., percentage changes) in this figure to make the results easily comparable across categories.

To further examine the relationship between budget optimism and budget compliance, we next performed panel regression analysis using the following model:

$$BudgetCompliance_{ic} = \alpha_0 + \beta_0 * Budget_{ic} + \beta_1 * PreBudget_{ic} + X'_i\beta + \varepsilon_i \quad (1)$$

where $BudgetCompliance_{ic}$ is defined as the difference between budgeted spending and actual spending for each user i in category c, $Budget_{ic}$ is the budget that each user i sets for category c, $PreBudget_{ic}$ is mean monthly spending for each user i in category c over the three months before they set a budget, and X' is a vector of control variables defined in Table 2. Conceptually, the right-hand side of this equation represents budget optimism because when mean pre-budget spending is held constant, a lower budget represents a more optimistic budget.

TABLE 2: STUDY 1 REGRESSION VARIABLE DETAILS

Notation	Variable	Definition
X'	Vector of control variables	Logins per month (squared), number of budgets per user, age (squared), gender, salary, month-of-the-year, and country, as described below.
	Logins per month (squared)	Sum of daily logins per month. Every day a user logs into their money dashboard account a login is registered. Hence, login ranges from 0 (no logins) to 31 (daily logins) per month. The squared term is included to account for the possibility of a non-linear relationship between login frequency and budget compliance.
	# of budgets	The number of budgets a user sets in the app. For example, if a user sets a grocery budget this variable is equal to one, if a user sets a grocery budget and a fuel budget this variable is equal to two, and so on.
	Age (squared)	The birth year of users is provided. Age is calculated as 2016 - birth year. The squared term is included to account for the possibility of a non-linear relationship between age and budget compliance.
	Gender identifier	This dummy variable is equal to one if the user is male and zero otherwise.
	Salary groups	Dummy variable for salary groups. Groups range from 0 to 10k, up to over 80k. We use the 20k to 30k group as our baseline because this group includes the largest number of users (38.27% of the sample).
	Month-of-the-year-FE	Dummy variable for month and year of budget creation.
	Country-FE	Dummy variables for Scotland and England, Wales is the baseline group.

The results of our budget compliance regression analysis are presented in Table 3. The positive coefficients for budget indicate that lower, more optimistic budgets are associated with weaker budget compliance. In tandem, the positive coefficients for Logins and the negative coefficients for Logins Squared suggest a quadratic relationship between login frequency and spending, such that the positive marginal effect of login frequency on spending diminishes as the number of logins increases.

TABLE 3: BUDGET COMPLIANCE REGRESSION RESULTS IN STUDY 1

	(1) Dining and drinking	(2) Groceries	(3) Fuel
Budget	0.877***	0.780***	0.870***
	(0.0411)	(0.0414)	(0.0486)
PreBudget	-0.378***	-0.483***	-0.341***
	(0.0690)	(0.0327)	(0.0472)
Logins	7.332***	9.680***	2.497**
	(1.537)	(1.868)	(1.267)
Logins squared	-0.182***	-0.281***	-0.0573
	(0.0568)	(0.0786)	(0.0525)
# of budgets	-0.690	0.816	4.747**
	(2.105)	(2.646)	(2.081)
Observations	2,185	2,403	1,010
Adjusted R-squared	0.496	0.388	0.421

NOTES: The dependent variable for all models is **budget compliance**, which is defined as budgeted spending minus actual spending in the relevant category. Hence, negative values indicate weaker compliance (i.e., higher overspending vs. budget). **Budget** is the amount of money a consumer budgets for a given category. **PreBudget** is mean spending in a given category over the three months before a budget is set. Control variables are defined in Table 2. Robust standard errors in parentheses. *$p < .05$, **$p \leq 0.01$, ***$p \leq 0.001$.

Taken together, the results of the preceding analyses indicate that more optimistic budgets are associated with weaker budget compliance. The model-free results presented in Figure 4 show that this is true between categories, and the regression results presented in Table 3 show this is true on a within-subject basis as well. However, the dining and drinking and groceries results in Figure 4 closely match the hypothesized outcomes in Panel C of Figure 1, which suggests that budgets can influence spending even when they are optimistic and budget compliance is imperfect. As preliminary evidence of this, within-subject t-tests reveal that post-budget spending is lower than pre-budget spending in both of these budget categories (p's < .001; see Figure 5). Our next analysis builds on this result.

Hypothesis 2: Budget Influence. To test the hypothesis that lower, more optimistic budgets are associated with lower spending, we performed panel regression analysis using the following model:

$$Spend_{ic} = \alpha_0 + \beta_0 * Budget_{ic} + \beta_1 * PreBudget_{ic} + X'_i\beta + \varepsilon_i \quad (2)$$

where the dependent variable $Spend_{ic}$ is the amount that each user *i* spends in category *c* during the first month after setting their budget, and the independent variables are defined as in Model 1. The results of this analysis are presented in Table 4. The positive coefficients for Budget provide support for H2 by demonstrating that lower budgets are associated with lower spending, even after controlling for mean pre-budget spending. Specifically, the budget variable coefficients indicate that a £10 decrease in budget is associated with a £1.23 decrease in dining and drinking spending, a £2.20 decrease in grocery spending, and a £1.30 decrease in spending on fuel, controlling for the other variables in the model.

TABLE 4: BUDGET INFLUENCE REGRESSION RESULTS IN STUDY 1

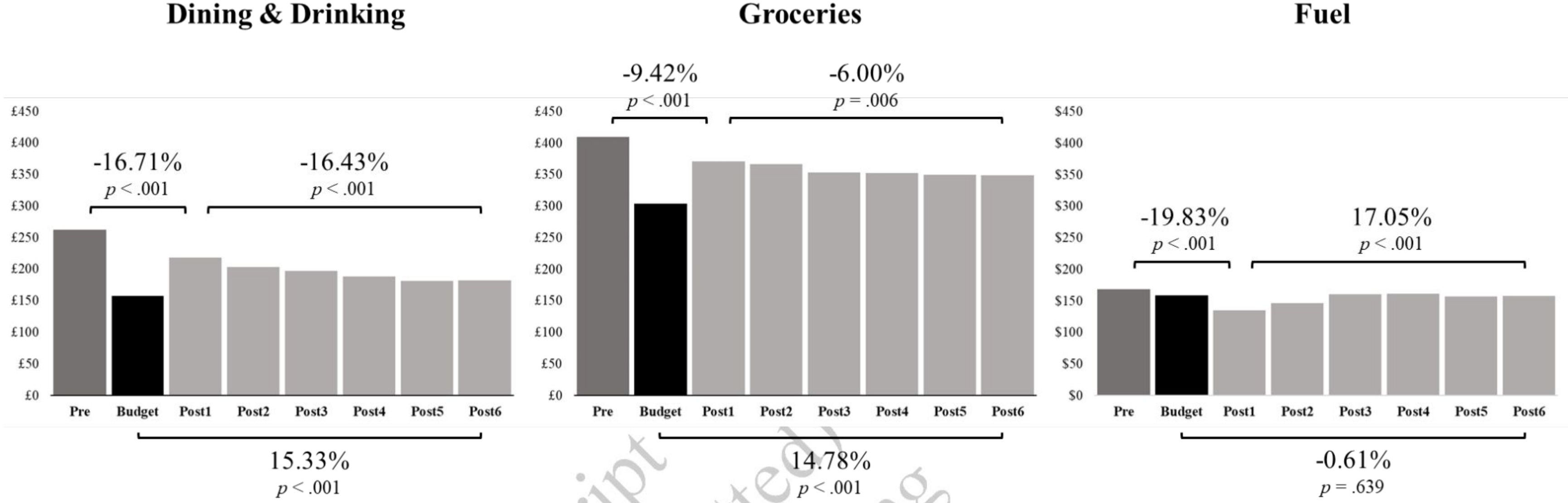
	(1) Dining and drinking	(2) Groceries	(3) Fuel
Budget	0.123***	0.220***	0.130***
	(0.0411)	(0.0414)	(0.0486)
PreBudget	0.378***	0.483***	0.341***
	(0.0690)	(0.0327)	(0.0472)
Logins	-7.332***	-9.680***	-2.497**
	(1.537)	(1.868)	(1.267)
Logins squared	0.182***	0.281***	0.0573
	(0.0568)	(0.0786)	(0.0525)
# of budgets	0.690	-0.816	-4.747**
	(2.105)	(2.646)	(2.081)
Observations	2,185	2,403	1,010
Adjusted R-squared	0.496	0.388	0.421

NOTES: The dependent variable for all models is **spending** in the relevant budget category during the month after a budget is set. **Budget** is the amount of money a consumer budgets for a given category. **PreBudget** is mean spending in a given category over the three months before a budget is set. Robust standard errors in parentheses. *$p <$ .05, **$p \leq 0.01$, ***$p \leq 0.001$.

In tandem, our tests of H1 and H2 support the proposition that budgets can be simultaneously wildly optimistic and yet also influential. We next test H3a and H3b to determine if the influence of budgets on spending persists over time.

Hypothesis 3: Longitudinal Budget Influence. Figure 5 and Table 5 summarize the relationship between mean pre-budget spending over the three months before a budget is set ("Pre"), budgeted spending ("budget"), and post-budget spending over the six months after a budget is set ("Post1," "Post2," and so on). Post-budget spending in the month after budget creation is significantly lower than pre-budget spending in all three categories, as revealed by within-subject t-tests. Furthermore, spending continues to decrease for dining and drinking and groceries, such that it is significantly lower six months after budget creation than it was in the first month after budget creation, *even though budget compliance remains relatively weak*. Spending on fuel displays the opposite trend, which is consistent with the fact that demand for a good tends to rise when its price drops. Collectively, these model-free results suggest that actual spending moves toward budgeted spending over time, regardless of whether the initial outcome was spending more than budget (as is the case for dining and drinking and groceries) or spending less than budget (as is the case for fuel).

FIGURE 5: MEAN PRE-BUDGET SPENDING, BUDGETED SPENDING, AND POST-BUDGET SPENDING IN STUDY 1



NOTES: "Pre" represents mean pre-budget spending over the three months before a budget is set. Budget represents the budget set for each category. "Post1," "Post2," and so on represent post-budget spending over the six months after a budget is set. Results were obtained using t-tests to compare absolute mean differences. We present relative differences (i.e., percentage changes) in this figure to make the results comparable across categories.

TABLE 5: MEAN DIFFERENCE BETWEEN PRE-BUDGET SPENDING AND POST-BUDGET SPENDING IN MONTHS 1–6 AFTER BUDGET CREATION IN STUDY 1

	Panel A: First month					Panel B: Second month				
	£	%	SD	n	*p*-value	£	%	SD	n	*p*-value
Dining and drinking	43.87	16.71%	219.56	2190	< 0.001	58.73	22.37%	236.21	2190	< 0.001
Groceries	38.61	9.42%	252.71	2415	< 0.001	43.32	10.57%	263.94	2415	< 0.001
Fuel	33.45	19.83%	115.55	1011	< 0.001	21.73	12.89%	118.83	1011	< 0.001
	Panel C: Third month					Panel D: Fourth month				
	£	%	SD	n	*p*-value	£	%	SD	n	*p*-value
Dining and drinking	65.40	24.91%	248.97	2190	< 0.001	73.85	28.13%	256.92	2190	< 0.001
Groceries	56.79	13.86%	282.14	2414	< 0.001	57.44	14.02%	286.83	2414	< 0.001
Fuel	8.31	4.92%	117.55	1010	< 0.050	6.80	4.03%	116.54	1010	< 0.100
	Panel E: Fifth month					Panel F: Sixth month				
	£	%	SD	n	*p*-value	£	%	SD	n	*p*-value
Dining and drinking	80.75	30.75%	265.14	2190	< 0.001	79.80	30.39%	273.23	2189	< 0.001
Groceries	60.34	14.72%	297.67	2414	< 0.001	60.90	14.86%	307.50	2412	< 0.001
Fuel	11.09	6.58%	119.33	1010	< 0.010	10.40	6.16%	118.70	1010	< 0.010

To directly test H3a, we performed panel regression analysis using the following model:

$$Spend_{ict} = \alpha_0 + \beta_0 * Budget_{ic} + \beta_1 * PreBudget_{ic} + X'_i\beta + \varepsilon_i \quad (3)$$

where $Spend_{ict}$ is the amount user *i* spends in category *c* in month *t*, and all other variables are as described in models (1) and (2) above. The results of this analysis are presented in Table 6. The persistently positive coefficients for Budget suggest that the influence of budgets on spending does *not* decay over time (H3a). In fact, the association between budgets and spending tends to strengthen month-over-month, while the association between pre-budget spending and post-budget spending weakens.

TABLE 6: REGRESSION RESULTS FOR H3A IN STUDY 1

	(1) Dining and drinking Month 1	(2) Dining and drinking Month 2	(3) Dining and drinking Month 3	(4) Dining and drinking Month 4	(5) Dining and drinking Month 5	(6) Dining and drinking Month 6
Budget	0.123***	0.134***	0.143***	0.142***	0.220***	0.207***
	(0.0411)	(0.0410)	(0.0400)	(0.0369)	(0.0334)	(0.0344)
PreBudget	0.378***	0.295***	0.234***	0.181***	0.129***	0.0770***
	(0.0690)	(0.0724)	(0.0604)	(0.0432)	(0.0332)	(0.0251)
Observations	2,185	2,185	2,185	2,185	2,185	2,184
Adjusted R-squared	0.432	0.360	0.321	0.283	0.246	0.176
	(1) Groceries Month 1	(2) Groceries Month 2	(3) Groceries Month 3	(4) Groceries Month 4	(5) Groceries Month 5	(6) Groceries Month 6
Budget	0.220***	0.285***	0.310***	0.343***	0.361***	0.421***
	(0.0414)	(0.0387)	(0.0374)	(0.0408)	(0.0383)	(0.0395)
PreBudget	0.483***	0.444***	0.352***	0.341***	0.298***	0.229***
	(0.0327)	(0.0289)	(0.0286)	(0.0290)	(0.0252)	(0.0279)
Observations	2,403	2,403	2,402	2,402	2,402	2,400
Adjusted R-squared	0.503	0.476	0.417	0.424	0.392	0.367
	(1) Fuel Month 1	(2) Fuel Month 2	(3) Fuel Month 3	(4) Fuel Month 4	(5) Fuel Month 5	(6) Fuel Month 6
Budget	0.130***	0.222***	0.238***	0.238***	0.233***	0.194***
	(0.0486)	(0.0502)	(0.0556)	(0.0448)	(0.0394)	(0.0369)
PreBudget	0.341***	0.302***	0.327***	0.316***	0.235***	0.224***
	(0.0472)	(0.0343)	(0.0413)	(0.0406)	(0.0339)	(0.0322)
Observations	1,010	1,010	1,009	1,009	1,009	1,009
Adjusted R-squared	0.261	0.262	0.307	0.295	0.227	0.221

NOTES: The dependent variable in each model is spending in month t, where t ranges from Month One after a budget is set to Month Six. Budget and PreBudget are defined as in the preceding analyses. The coefficients for Logins, Logins Squared, and # of Budgets are substantively identical to those in Table 4, so they are omitted here for space. Robust standard errors in parentheses. *$p < .05$, **$p \leq 0.01$, ***$p \leq 0.001$.

The preceding analysis examines the relationship between budgets and spending over time, and it finds that the association between these two variables is remarkably persistent. We next examine the relationship between budget *compliance* and spending over time (H3b), because past research suggests that failure to comply with a budget in one period should lead to higher spending in the next period (Soman and Cheema 2004). To test this hypothesis, we performed a panel regression analysis identical to Model (3), but with budget compliance at time t-1 in place of the budget variable. So, this model examines whether compliance in month one is associated with spending in month two, whether compliance in month two is associated with spending in month three, and so on. The results of this analysis are presented in Table 7. Supporting H3b, the negative coefficients for Compliance indicate that weaker compliance in month t-1 is associated with higher spending in month t. We next test H4 by comparing the spending behavior of budgeters with similar non-budgeters.

TABLE 7: REGRESSION RESULTS FOR H3B IN STUDY 1

	(1) Month 2 Spending Dining and drinking	(2) Month 3 Spending Dining and drinking	(3) Month 4 Spending Dining and drinking	(4) Month 5 Spending Dining and drinking	(5) Month 6 Spending Dining and drinking
$Compliance_{t-1}$	-0.313***	-0.289***	-0.298***	-0.283***	-0.350***
	(0.0296)	(0.0271)	(0.0215)	(0.0258)	(0.0232)
PreBudget	0.220***	0.191***	0.154***	0.127***	0.0809***
	(0.0640)	(0.0505)	(0.0350)	(0.0310)	(0.0237)
Observations	2,185	2,185	2,185	2,185	2,184
Adjusted R-squared	0.446	0.397	0.374	0.311	0.282
	(1) Month 2 Spending Groceries	(2) Month 3 Spending Groceries	(3) Month 4 Spending Groceries	(4) Month 5 Spending Groceries	(5) Month 6 Spending Groceries
$Compliance_{t-1}$	-0.234***	-0.269***	-0.241***	-0.291***	-0.333***
	(0.0273)	(0.0239)	(0.0226)	(0.0241)	(0.0240)
PreBudget	0.441***	0.350***	0.376***	0.329***	0.280***
Observations	2,403	2,402	2,402	2,402	2,400
Adjusted R-squared	0.490	0.441	0.431	0.415	0.398

	(1) Month 2 Spending Fuel	(2) Month 3 Spending Fuel	(3) Month 4 Spending Fuel	(4) Month 5 Spending Fuel	(5) Month 6 Spending Fuel
$\text{Compliance}_{t\text{-}1}$	0.00416	-0.124***	-0.106***	-0.148***	-0.188***
	(0.0327)	(0.0330)	(0.0266)	(0.0301)	(0.0280)
PreBudget	0.364***	0.382***	0.369***	0.284***	0.273***
	(0.0372)	(0.0463)	(0.0430)	(0.0370)	(0.0369)
Observations	1,010	1,009	1,009	1,009	1,009
Adjusted R-squared	0.253	0.312	0.293	0.236	0.258

NOTES: Robust standard errors in parentheses. $^*p < .05$, $^{**}p \leq 0.01$, $^{***}p \leq 0.001$.

Hypothesis 4: Budgeters vs. Non-budgeters. To test H4 we performed a propensity score matching analysis (Rosenbaum and Rubin, 1983; Imbens and Rubin, 2015) to compare the spending of MDB users who set budgets to the spending of similar MDB users who do not set budgets. Analytically, this parallels the approach taken by marketing researchers who have used secondary data of a similar nature to investigate the listening behavior of music streaming platform users (Datta, Knox and Bronnenberg, 2018). For budgeters, our dependent variable is the difference between their pre- and post-budget spending. For non-budgeters, the dependent variable is the same difference over the same time period, but without a budget having been created. So, for example, if a user sets a budget on May 1st, we compare the change in their spending from April to May against the change in spending of similar non-budgeters over the same period.

Mirroring our regression analyses above, we match budgeters and non-budgeters on login frequency, salary, age, gender, and country of residence. We then estimate differences between budgeters and non-budgeters for all of these matching variables (Leuven and Sianesi 2003). If matching is successful, these variables should not differ significantly from each other (p >.10). Table 8 shows that this condition is fulfilled. However, even though our estimators are not biased, we also report the results of a nearest-neighbour analysis which uses bias-corrected matching estimators as a robustness test (Abadie and Imbens 2011).

TABLE 8: STUDY 1 PROPENSITY SCORE BIAS ANALYSIS

Variable	**Treated**	**Control**	**%Bias**	***p*-value**
Age	37.34	37.06	2.8	0.426
Salary	28.43	28.18	1.7	0.620
Logins	7.47	7.28	3.1	0.422
Male	0.69	0.71	-4.4	0.184
England	0.85	0.85	-2.0	0.553
Scotland	0.10	0.10	0.6	0.868
Wales	0.05	0.04	2.7	0.447

As can been seen in Table 9, the results of our propensity score matching analysis support H4: budgeters decreased their spending to a larger degree than similar non-budgeters across all three budget categories.

TABLE 9: PROPENSITY SCORE MATCHING RESULTS FOR H4 IN STUDY 1

Delta Spending	(1) Dining and Drinking		(2) Groceries		(3) Fuel	
Treatment PSM	39.73*** (4.264)		34.34*** (6.289)		20.35** (9.872)	
Treatment NN		42.57*** (4.666)		34.45*** (6.595)		22.32** (10.60)
Observations	70,476	63,467	80,840	76,358	54,931	42,355

NOTES: Coefficients represent average treatment effects. The first row shows results for the propensity score matching analysis. The second row shows results for the nearest neighbor analysis that makes use of the bias adjusted estimator developed by Abadie and Imbens (2011). The dependent variable is the month-to-month difference in spending before and after a budget was set. The selected matching variables are age, gender, salary groups, number of logins per month, month of the year, country (Wales, England, Scotland). Robust standard errors in parentheses. *$p < .05$, **$p \leq 0.01$, ***$p \leq 0.001$.

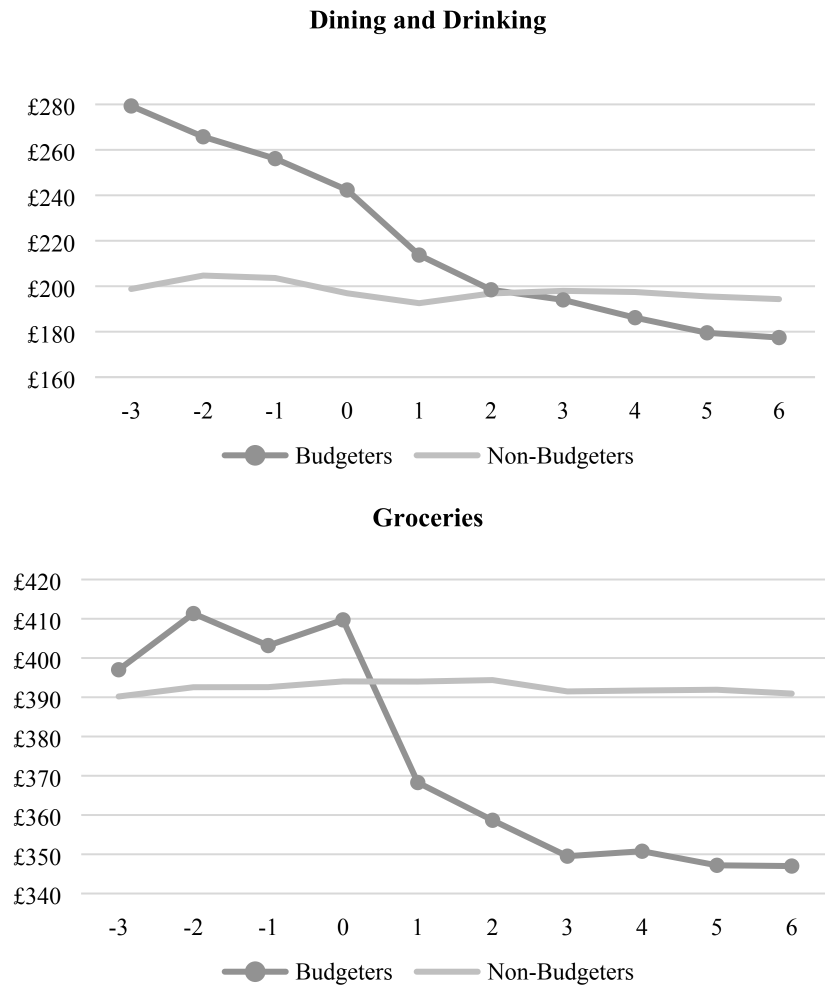
One limitation of our propensity score matching analysis is that it could hide regression toward to the mean. For example, it could be the case that consumers set a budget after a particularly high month of spending, and that the results in Table 9 merely represent a return to normal or average spending amongst budgeters. To examine this possibility, we performed a size-matched analysis as suggested in Barber and Odean (2002). For this size-matched analysis we identified the 'nearest neighbour' for each individual with a budget in a given category as in the matching analysis above. In addition to the matching approach above, we also match on total pre-budget spending per month per category. We then calculate mean spending for the size-matched individuals for three months before and six months after a budget was set.

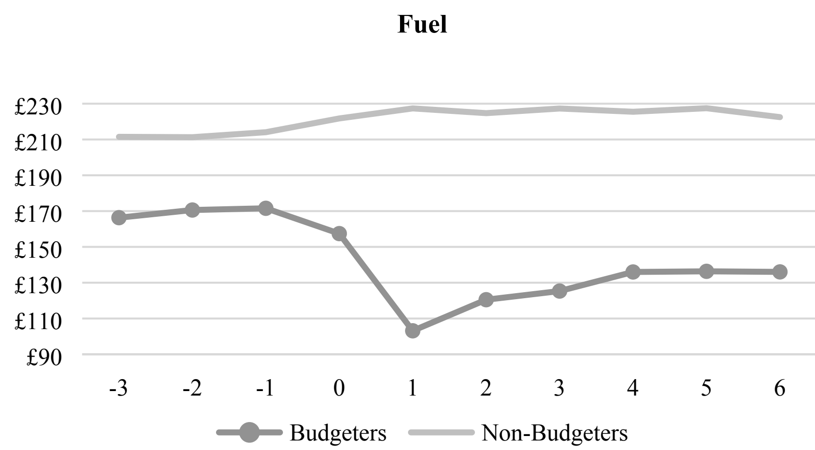
Figure 6 plots the results of our size-matched analysis. Point zero on the x-axis represents when a budget is set, −3 to 0 represent the three months before a budget is set, and 0 to 6 represent the six months after a budget is set. The results for groceries and fuel very clearly show that the PSM results are not a result of regression toward the mean. The results for dining and drinking are more nuanced. Budgeters show a clear downward trend in pre-budget spending, but the decrease in spending during the first month after a budget is set (i.e., from point 0 to 1) is significantly larger than the average monthly decrease observed in pre-budget spending (i.e., from point −3 to 0). In other words, the slope of the line from 0 to 1 is significantly steeper than the slope of the line from −3 to 0 ($t(1,433) = -6.75$, $p < .001$). This is a particularly compelling test of budget influence because marginal behavior change typically becomes more difficult over time. Just as dieters have a harder time losing the last ten pounds than the first ten, consumers should have a harder time decreasing their spending by yet *another* $10 after they have done so for several months already, because each $10 decrement represents a larger percentage change in spending vis-à-vis an increasingly tight budget constraint.[2] However, despite the fact that spending on dining and drinking was already trending downward, spending in the month after a

[2] As an example, consider a consumer who typically spends $200 per month, then decreases their spending by $10 every month for 10 months. In month one, their marginal change in spending is 5%. By the end of month ten, when their monthly spending is down to $100 per month, another $10 decrease represents a marginal change of 10%.

budget was set decreased approximately 2.5 times *more* than the average monthly decrease before a budget was set.

FIGURE 6: LONGITUDINAL RESULTS FOR H4 IN STUDY 1





NOTES: The dependent variable in each graph of Figure 6 is spending in the relevant category. Point zero on the x-axis represents when a budget is set, −3 to 0 represent the three months before a budget is set, and 0 to 6 represent the six months after a budget is set.

Discussion

Study 1 offers several valuable insights regarding the relationship between budgets and spending. Budget compliance is shown to vary across categories, such that budget compliance is weaker when budget optimism is stronger (H1b). However, more optimistic budgets are associated with lower spending across all three budget categories (H2), suggesting that budgets can influence spending even when budget compliance is weak. We also observe that budget influence persists over time (H3a), and that stronger budget compliance in one month is associated with lower spending in the next month (H3b). Finally, MDB users who set budgets subsequently decrease their spending to a greater degree than similar users who do not (H4).

Empirically, the strengths of Study 1 include a large sample and precise longitudinal measurement of spending, both before and after budgets are set. One limitation of Study 1 is that budgeters may differ from non-budgeters in unobservable ways. To address this concern, we begin by offering the following thought experiment: what if the budgeters in the data set are inherently *more* financially responsible than the non-budgeters? If true, then the budget compliance results become even more interesting, because this implies that even relatively responsible consumers have trouble accurately forecasting their expenses and sticking to their budgets. It is also informative to consider the inverse: what if the budgeters in the data set are somehow *less* financially responsible than the non-budgeters? If true, then the results of the matching analysis become even more interesting, because this implies that budgeting helped relatively less responsible consumers decrease their spending to a greater degree than their relatively more responsible peers. Furthermore, there is no reason to believe that MDB budgeters are any different from the tens of millions of consumers who currently use budgeting apps world-wide, which implies that are our results are likely indicative of a large group of consumers whose budgeting and spending behavior is of immense interest to marketing researchers. Nonetheless, we next test H1, H2, and H4 in a field experiment that controls for unobserved individual differences through random assignment.

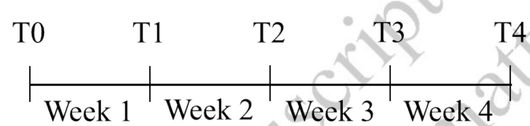
STUDY 2: FIELD EXPERIMENT

The primary purpose of Study 2 was to examine the causal effect of budget setting on spending (H4) in a field experiment. We also used Study 2 to test H1 and H2. Finally, we used this study to examine the association between budgets, spending, and several psychographic variables of theoretical interest. We did this to provide insight on whether the effect of budget setting on spending generalizes across subpopulations (Cook and Campbell 1979; Lynch 1982). This study was preregistered on aspredicted.org (https://aspredicted.org/793m9.pdf).

Method

Participants in this field experiment were members of Vancity, Canada's largest community credit union. Participants were recruited through an online panel of approximately 5,000 members that Vancity uses for market research. Each participant completed five surveys over the course of four weeks, as illustrated by the T0 – T4 markers in Figure 7.

FIGURE 7: DATA COLLECTION SCHEDULE FOR STUDY 2



Our target sample size for this study was 200+ participants. To hit this target, we emailed the T0 survey to 1,250 randomly selected panel members at noon on a Sunday. Two hundred and seventy-one people completed the T0 survey before it was deactivated at 11:59pm the next day. 83.4% of these participants also completed the T1, T2, T3, and T4 surveys, leaving us with a final sample of 226. Compensation for participants who completed the entire study included a personalized spending report (provided at the end of the study) that served as an incentive to budget and report expenses as accurately as possible. Participants who completed the entire study also received a $30 gift certificate to supportlocalbc.com, a website that promotes local businesses in British Columbia, Canada. As illustrated in Table 10, participants who completed the entire study did not differ significantly from participants who failed to complete the entire study in terms of any variable collected at T0.

TABLE 10: PARTICIPANT CHARACTERISTICS IN STUDY 2

Variables Measured at T0	Participants who completed the study	Participants who did not complete the study	*p*-value
N	226	45	n/a
Mean age	55.83 (14.81)	55.87 (13.45)	0.99
% female	50.9%	57.8%	0.40
Education			
High school	5.8%	2.3%	0.34
Some college	19.2%	29.5%	0.12
Associate degree	15.2%	18.2%	0.61
Bachelor's degree	31.7%	27.3%	0.56
Master's degree	17.9%	18.2%	0.96
Professional degree	5.4%	2.3%	0.38
Doctoral degree	4.5%	2.3%	0.50
Employment status			
Full-time	42.0%	44.4%	0.77
Part-time	8.4%	8.9%	0.91
Retired	30.1%	31.1%	0.89
Unemployed	4.9%	6.7%	0.62
Other	12.8%	8.9%	0.47
Median household income	\$80,000-\$89,999 (4.27)	\$70,000-\$79,999 (4.10)	0.17
Median available resources	\$15,000 (149,111.47)	\$10,000 (184,508.04)	0.46
Mean discretionary spending budget created for the study	\$939.11 (886.38)	\$960.42 (776.14)	0.91
Mean discretionary spending in a typical month before the study	\$1006.42 (1226.06)	\$1101.44 (946.86)	0.56
General budgeting behavior			
Budget setting	3.77 (1.74)	3.71 (1.70)	0.85
Expense tracking	4.35 (1.81)	4.47 (1.95)	0.71
Preferred payment method			

Cash	5.3%	0.0%	0.12
Credit card	69.9%	73.3%	0.65
Debit card	20.4%	24.4%	0.55
Other	4.4%	2.2%	0.49
Mean savings goal endorsement	5.68 (1.13)	5.73 (1.29)	0.77
Mean financial anxiety	3.76 (1.73)	4.20 (1.95)	0.13
Mean impulsive buyer behavior	2.97 (1.14)	3.26 (1.19)	0.12

In the T0 survey participants were randomly assigned to either a control condition in which they did not set a budget or a treatment condition in which they set a budget for their discretionary spending over the next month. Specifically, participants in the budgeting condition received the following instructions:

> Please take some time to set a discretionary spending budget for the next month (i.e., the next four weeks). By "discretionary spending" we mean everything you will spend money on in the next month except your regular monthly bills.
>
> **(Line Break)**
>
> Please enter your discretionary spending budget for the next month in the space below.
>
> **(Page Break)**
>
> Thank you. Now please take a moment to record your budget so you can refer back to it in future surveys. Most people make a note in their phone or on their computer, but you can jot it down anywhere that makes it easy for you to find.

We asked participants in the treatment condition to budget for their total discretionary spending rather than their spending in a specific category like dining and drinking so that we did not have to impose any assumptions about which specific categories are most relevant to this sample. We asked these participants to record their budget so they could easily track their spending against it in surveys T1–T4, as Money Dashboard users in Study 1 do when they login to the app.

After setting and recording their budget, participants in the budgeting condition completed the measures described below. Participants in the control condition also completed the measures described below; the only difference between the two conditions is that control condition participants did not set a budget for their discretionary spending.

Savings goal. Following Peetz and Buehler (2009) we measured participants' savings goals by asking them to indicate their level of agreement with the statement "In general, saving money is very important to me (1 = Strongly Disagree; 7 = Strongly Agree)." If the amount of money consumers budget is influenced by savings goals, as Peetz and Buehler suggest, then this measure should be negatively correlated with participants' budgets.

General budgeting behavior. We measured participants' general budgeting behavior with two items: "In general, I set a budget for my discretionary spending" and "In general, I carefully track my discretionary spending" (Strongly Disagree = 1, Strongly Agree = 7). The mean response to each of these items did not differ between conditions (p's > .26). However, the mean response to the budget setting item was significantly below the scale midpoint of 4 (M = 3.77, SD = 1.74, $t(225)$ = -2.03, p = .044), and the mean response to the tracking item was significantly above the scale midpoint (M = 4.35, SD = 1.81, $t(225)$ = 2.94, p = .004). So, on average, participants in this study tend to track their discretionary spending without setting a budget. This confirms that if budget setting does influence spending above and beyond expense tracking, as the results of Study 1 indicate, we should observe lower spending in the budgeting condition versus control. (In contrast, if participants in our sample generally did set budgets, we would not necessarily expect to observe an effect of budgeting condition on spending. This is because the budgeting condition would merely be asking participants in that condition to do something most study participants in both conditions naturally do on their own.)

Typical discretionary spending. To measure participants' typical discretionary spending, we asked "Over the past six months, approximately how much was your discretionary spending in a typical month? [Free Response Text Box]." As per our preregistration, we collected this measure to use as a control variable in our analysis of spending between conditions. Assumption checks confirmed that responses to this question did not differ significantly between conditions (M_{control} = \$958.97, SD_{control} = 878.26, $M_{\text{treatment}}$ = \$1061.11, $SD_{\text{treatment}}$ = 1021.76, $t(224) = .81$, $p = .42$), but they were correlated with participants' discretionary spending over the course of the study ($r_s(224) = .53$, $p < .001$).

Preferred payment method. To measure participants' preferred payment method, we asked them to respond to the statement "Typically, I pay for most of my discretionary spending with..." by selecting Cash, Credit Card, Debit Card, or Other. We collected this measure to explore whether consumers who prefer to pay with cash budget and/or spend less than those who prefer to pay with a card, but we were unable to test this possibility in a meaningful way because only 12 participants (5.3% of the sample) indicated a preference to pay with cash.

Financial anxiety. We measured financial anxiety using two items presented in counterbalanced order: "Thinking about my personal finances makes me feel anxious" and "Thinking about my personal finances makes me feel stressed" ($r(224) = .91$, $p < .001$). We measured financial anxiety so we could examine whether more anxious consumers display lower budget compliance in the treatment condition than less anxious consumers, and whether the effect of budgeting condition on spending (vs. control) is moderated by financial anxiety.

Impulsive buying behavior. We measured impulsive buying behavior using the 9-item scale developed by Rook and Fisher (1995) so we could examine whether more impulsive buyers display lower budget compliance in the treatment condition than less impulsive buyers, and whether the effect of budgeting condition on spending (vs. control) is moderated by impulsive buying behavior.

Demographics. At the end of the T0 survey we measured six demographic variables in the following order: 1) available resources (Hardisty, Appelt, and Weber 2013), 2) age, 3) sex (female = 1, male = 0), 4) education, 5) employment status, and 6) household income (in $10K buckets ranging from "less than $10,000" to "more than $150,000). As per our preregistration, we collected these variables to use as controls in our regression analyses. However, we ended up with missing demographic data from 25 participants because the credit union asked that responses to these questions be made optional. We therefore performed all regressions twice: once with a simple model that included the entire sample and controlled only for typical past spending, and again with the full preregistered model that excludes the 25 participants mentioned above so that we could control for all demographic variables in addition to typical past spending. The two models produce substantively identical results, so we report the simple model results across all tests in Study 2 to avoid the issue of missing data, and ultimately, for parsimony.

The T1–T4 surveys were launched at 6pm each Sunday during the study, and they were deactivated at 11:59pm the next day. These surveys began by asking all participants to log into their online bank account and report their discretionary spending for the past week. After reporting their spending participants in the control condition were asked to indicate their level of (dis)agreement with the statement "I am on track to spend less money on discretionary expenses this month than I do in a typical month (1 = Strongly Disagree, 7 = Strongly Agree)."

Participants in the budgeting condition were asked to recall their discretionary spending budget, then indicate their level of (dis)agreement with the statement "I am on track to meet the discretionary spending budget I set for myself this month (1 = Strongly Disagree, 7 = Strongly Agree)" as well as "I am on track to spend less money on discretionary expenses this month than I do in a typical month (1 = Strongly Disagree, 7 = Strongly Agree)." We included these items to mimic the expense tracking process in the Money Dashboard app, where all app users can evaluate their current spending against their past spending, and users who set budgets can also evaluate their current spending against their budget.

We also used the T1–T4 surveys to measure the following individual differences: 1) Temporal discounting for losses and gains (e.g., Chapman 1996), 2) Trait Self-Control (Tangney, Baumeister, and Boone 2004), 3) Spendthrift-Tightwad (Rick, Cryder, and Loewenstein 2008), 4) Trait optimism (Scheier, Carver, and Bridges 1994), 5) Financial Propensity to Plan (Lynch, Netemeyer, Spiller, and Zammit 2010), and 6) Financial literacy (Lusardi and Mitchell 2011). The purpose of measuring these variables was twofold. First, we wanted to examine the relationship between these variables and budget compliance in the treatment condition. Second, we wanted to determine if the effect of budgeting condition on discretionary spending (vs. control) was moderated by these variables.

Analysis

Because the primary goal of this study was to test H4 in a causal fashion, our preregistered dependent variable was each participant's total discretionary spending summed across the T1–T4 surveys. This variable displayed strong positive skew in both conditions ($Skew_{control}$ = 7.36, $Skew_{treatment}$ = 2.56), and it violated the homogeneity of variance (HOV) assumption (F(1, 224) = 4.31, p = .039). As per the standard operating procedure in our preregistration, we compared three common methods for addressing this situation: a simple LN-transformation of total spending, winsorizing total spending at the 2.5th and 97.5th percentile in each condition, and excluding data points +/– 3 standard deviations from the mean in each condition. Each of these approaches produced results consistent with the others, but only the LN-transformation neutralized the HOV violation (F(1, 224) = 1.02, p = .31). Therefore, we performed our budget influence analyses with LN(discretionary spending) as the dependent variable. For consistency, we also LN-transformed all other continuous monetary variables (e.g., budgeted spending and typical discretionary spending). Then, for ease of interpretation, we exponentiated the results to present summary statistics in dollars and cents.

Results

Hypothesis 1: Budget Compliance. Participants in the budget-setting condition spent 36.2% more than they budgeted, as revealed by a paired-samples t-test ($M_{spending}$ = \$856.20, $CI_{95\%}$ = [730.63, 1,003.45]; M_{budget} = \$628.48, $CI_{95\%}$ = [517.60, 763.19]; t(104) = 3.95, p < .001). This replicates the model-free evidence observed for dining and drinking and groceries in Study 1.

We next regressed budget compliance onto budgeted spending and typical past spending, because when past spending is held constant, a lower budget is a more optimistic budget. The model was significant (F(2, 102) = 40.70, p < .001, Adjusted-R^2 = .433), as were the coefficients on budgeted spending (B = .74, SE = .090, 95% CI = [.559, .915], t(102) = 8.22, p < .001) and typical past spending (B = −.35, SE = .093, 95% CI = [−.533, −.164], t(102) = −3.74, p < .001).

The positive coefficient for budgeted spending indicates that lower, more optimistic budgets are associated with weaker budget compliance, which replicates the budget compliance regression results in Study 1.

To explore consumer heterogeneity in budget compliance we began by generating the correlation matrix in Table 11, which presents the Pearson correlation coefficient for budget compliance, budgeted spending, actual spending, and each of the psychographic and demographic variables that were measured in Study 2. Higher buyer impulsiveness (Rook and Fisher 1995) is associated with weaker compliance. This is also true of temporal discount rates for monetary gains and losses (Chapman 1996), which serve as an indirect measure of impulsiveness (e.g., Dittmar and Bond 2010). Interestingly, this appears to be in part because higher impulsiveness and discount rates are associated with (directionally) lower budgets more than with (directionally) higher spending. This suggests that impulsive consumers may display weaker budget compliance because they budget less than non-impulsive consumers, not because they spend more. A similar pattern is observed for financial anxiety, which is also associated with lower budgets and weaker compliance, but not with actual spending. In contrast, trait optimism is associated with (marginally) stronger compliance because optimism is associated with higher budgets more strongly than with higher spending. This is notable, because it suggests that the tendency to set an optimistically low budget does not stem from an optimistic disposition (see also Buehler, Griffin, and Peetz 2010; Howard et al. 2022). Finally, trait self-control, spendthrift-tightwad tendencies, savings goals, financial propensity to plan, and financial literacy are not significantly correlated with budget compliance because they have a roughly equal effect on budgets and spending. For example, being a tightwad is not associated with stronger budget compliance than being a spendthrift because tightwads set lower budgets *and* spend less than spendthrifts.

TABLE 11: STUDY 2 BUDGET COMPLIANCE CORRELATION MATRIX

Values represent Pearson correlation coefficients. Budgeted spending, actual spending, available resources, and typical monthly discretionary spending were LN-transformed for analysis. Budget compliance is calculated as LN(budgeted spending) minus LN(actual spending) so that lower values represent weaker compliance and higher values represent stronger compliance. Temporal discounting for gains is coded so that a higher score represents a preference to receive less money today than more money in the future. Temporal discounting for losses is coded so that a higher score represents a preference to pay more money in the future than less money today. Spendthrift-tightwad is coded so that lower values represent spendthrift tendencies and higher values represent tightwad tendencies.

	Variables	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
1	Budget compliance	-																			
2	Budgeted spending	.61**	-																		
3	Actual spending	-.24*	.63**	-																	
4	Buyer impulsiveness	-.22*	-.16	.02	-																
5	Temporal discounting (gains)	-.18†	-.13	.02	.23*	-															
6	Temporal discounting (losses)	-.23*	-.15	.04	.15	.19*	-														
7	Trait self-control	.09	.13	.08	-.42**	-.09	-.15	-													
8	Spendthrift-Tightwad	.00	-.11	-.14	-.37**	.07	.01	.17†	-												
9	Trait optimism	.16†	.17†	.04	-.23*	-.15	-.15	.45**	.11	-											
10	Savings goal	.01	.05	.03	-.19†	-.25*	-.12	.16†	-.06	.06	-										
11	Financial propensity to plan	.08	-.07	-.15	-.17†	.10	.11	.25**	.08	.05	.17†	-									
12	Financial literacy	.00	.12	.14	-.09	-.19†	-.16	.13	-.10	.01	.08	-.14	-								
13	Financial anxiety at T0	-.25**	-.20*	-.01	.29**	.18†	.22*	-.23*	-.04	-.35**	.05	.06	-.19†	-							
14	Age	.10	-.03	-.11	-.15	-.05	.08	.23*	.12	.11	-.07	.23*	-.12	-.07	-						
15	Sex (female = 1, male = 0)	.13	.00	-.11	.00	-.04	-.06	.03	.11	.09	.07	.11	-.31**	.16†	-.07	-					
16	Annual household income	.05	.23*	.22*	-.07	-.07	-.11	.09	-.10	.15	.06	-.25*	.23*	-.30**	-.11	-.10	-				
17	Available resources	.29**	.44**	.25*	-.27**	-.31**	-.33**	.07	-.07	.11	.19*	-.25**	.32**	-.36**	-.11	-.10	.45**	-			
18	Employed (yes =1 , no = 0)	-.06	.09	.15	.18†	-.07	.23*	-.05	-.14	-.11	.08	-.29**	.23*	.23	-.45**	.02	.22*	.09	-		
19	Education	-.06	-.07	-.02	-.36**	-.20*	-.08	.28**	.13	.08	.13	.04	.13	.13	.01	.00	.12	.05	.16	-	
20	Typical discretionary spending	.28**	.76**	.65**	-.01	-.06	-.11	.19†	-.16	.18†	.00	-.07	.25**	.25**	-.01	.00	.24*	.33**	.04	-.11	-

NOTES: $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$

To test the robustness of the relationship between budget compliance and impulsiveness, temporal discount rates, anxiety, and optimism, we next performed a set of OLS regressions with budget compliance as the dependent variable, each psychographic as the focal independent variable, and typical past spending as a control variable. So, for example, in model one we regressed budget compliance onto impulsiveness and typical past spending, in model two we regressed budget compliance onto temporal discount rates for gains and typical past spending, and in model three we regressed budget compliance onto temporal discount rates for losses and typical past spending. The omnibus test for all models was significant (p's < .01). Model one revealed a significant relationship between budget compliance and impulsiveness (B = −.161, SE = .070, 95% CI = [−.299, −.022], $t(102) = -2.31$, $p = .023$), controlling for typical past spending. Model two produced similar results using temporal discounting for gains in place of impulsiveness (B = −.478, SE = .266, 95% CI = [−1.101, .049], $t(102) = -1.80$, $p = .075$), as did model three, which used temporal discounting for losses in place of impulsiveness (B = −.972, SE = .462, 95% CI = [−1.888, −.055], $t(102) = -2.10$, $p = .038$). It warrants emphasis that these analyses are exploratory rather than confirmatory, but the consistent pattern of results across three different measures of impulsiveness collected during different weeks of the study is notable. Model four revealed a significant relationship between budget compliance and anxiety (B = −.097, SE = .045, 95% CI = [−.187, −.008], $t(102) = -2.15$, $p = .034$). However, model five revealed that the relationship between budget compliance and trait optimism does not reach significance when typical past spending is included as a control (B = .16, SE = .124, 95% CI = [−.091, .400], $t(102) = 1.25$, $p = .214$).

Hypothesis 2: Budget Influence. To examine the association between budgets and spending we regressed LN(discretionary spending) onto LN(budgeted spending) and LN(typical spending). The model was significant ($F(2, 102) = 45.16$, $p < .001$, Adjusted-$R^2 = .459$), as were the coefficients on budgeted spending (B = .263, SE = .090, 95% CI = [.085, .441], $t(102) = 2.94$, $p = .004$) and typical past spending (B = .348, SE = .093, 95% CI = [.164, .533], $t(102) = 3.74$, $p < .001$). The positive coefficient for budgeted spending indicates that lower budgets are associated with lower spending, even when typical past spending is held constant. These findings provide further support for H2 by replicating the budget influence regression results in Study 1.

Hypothesis 4: Budgeters vs. Non-Budgeters. To test the causal effect of budget setting on spending we regressed LN(discretionary spending) onto a dummy variable for experimental condition (control = 0, treatment = 1) and LN(typical discretionary spending). The model was significant ($F(2, 223) = 57.05$, $p < .001$, Adjusted-$R^2 = .333$), as were the coefficients on condition (B = −.219, SE = .094, 95% CI = [−.405, −.033], $t(223) = -2.32$, $p = .021$) and typical past spending (B = .516, SE = .049, 95% CI = [.420, .613], $t(223) = 10.52$, $p < .001$). Specifically, the negative coefficient on the condition variable indicates that participants in the budgeting condition spent 21.9% less than participants in the control condition, holding constant typical past spending. This replicates and extends the matching results in Study 1 by providing causal support for H4.

To examine the possibility that budget setting has heterogenous effects on spending across psychographics, we next performed a series of OLS regressions in which LN(spend) was regressed onto condition, a mean centered psychographic variable, a condition × psychographic interaction term, and typical past spending included as a control variable. So, for example, in model one we regressed LN(spend) onto condition, mean centered buyer impulsiveness, a condition × impulsiveness interaction term, and typical past spending. The results of these regressions are presented in Table 12. The coefficient for Condition is negative and significant in all models, indicating that budgeters spent less than non-budgeters even after controlling for psychographics. None of the coefficients for the psychographic variables or interaction terms reach significance. Taken together, one interpretation of the results in Table 12 is that budget setting is an effective way to reduce spending for a diverse array of consumers. More plainly, it may be the case that budget setting is an effective way for most people to reduce their spending.

TABLE 12: CONDITION × PSYCHOGRAPHIC REGRESSION RESULTS IN STUDY 2

	(1)	(2)	(3)	(4)	(5)
Condition (control = 0, budget = 1)	−0.210*	−0.225*	−0.220*	−0.205*	−0.213*
	(0.095)	(0.094)	(0.094)	(0.096)	(0.095)
Psychographic	0.085	0.291	0.679	−0.107	−0.065
	(0.054)	(0.226)	(0.413)	(0.103)	(0.053)
Interaction	−0.066	−0.121	−0.242	0.055	0.032
	(0.085)	(0.332)	(0.595)	(0.156)	(0.081)
TypSpend	0.513*	0.525*	0.528*	0.516*	0.504*
	(0.050)	(0.049)	(0.049)	(0.050)	(0.050)
Adjusted R-Squared	0.334	0.332	0.340	0.330	0.332
	(6)	(7)	(8)	(9)	(10)
Condition (control = 0, budget = 1)	−0.213*	−0.218*	−0.215*	−0.217*	−0.215*
	(0.095)	(0.095)	(0.094)	(0.094)	(0.094)
Psychographic	−0.057	−0.037	0.020	−0.358	0.044
	(0.092)	(0.056)	(0.055)	(0.300)	(0.038)
Interaction	−0.032	0.068	−0.091	0.323	0.014
	(0.145)	(0.085)	(0.076)	(0.441)	(0.056)
TypSpend	0.522*	0.512*	0.516*	0.520*	0.529*
	(0.050)	(0.050)	(0.049)	(0.050)	(0.049)
Adjusted R-Squared	0.330	0.329	0.332	0.331	0.336

NOTES: The dependent variable in all models is LN(spend). The psychographic variable in each model is: 1) buyer impulsiveness, 2) temporal discounting for gains, 3) temporal discounting for losses, 4) trait self-control, 5) spendthrift-tightwad tendencies, 6) trait optimism, 7) savings goal, 8) financial propensity to plan, 9) financial literacy, and 10) financial anxiety (measure at T0). Standard errors in parentheses. $*p < .05$, $**p < .01$, $***p < .001$.

Discussion

Study 2 complements and extends Study 1 in three primary ways. First and foremost, Study 2 provides causal evidence showing that budget setting helps consumers spend less versus control (H4). Moreover, the results of Study 2 suggest that the effect of budget setting on spending may be fairly uniform across different psychographics. Second, Study 2 replicates the finding that budget optimism leads to weaker budget compliance (H1b), but that budget optimism is also associated with lower spending (H2). Third, Study 2 provides evidence that, counterintuitively, more impulsive consumers display worse budget compliance than less impulsive consumers because they under-budget rather than overspend.

One limitation of Studies 1 and 2 is that the evidence they provide for H2 is correlational. Thus, the primary goal of Study 3 is to test H2 in a causal fashion.

STUDY 3: THE FINANCIAL DIARY STUDY

The primary purpose of Study 3 is to provide a causal test of H2. To accomplish this, we randomly assigned consumers to one of two conditions that were designed to produce relatively high or low budget estimates. Participants then reported their actual spending in a series of online financial diary entries. Our principal expectation is that consumers who make lower budget estimates will subsequently spend less than those who make higher budget estimates (H2).

Method

We recruited four hundred and fifty US residents via Amazon Mechanical Turk to participate in a week-long financial diary study that required completing eight surveys (one survey per day from Sunday to Sunday). Payment for completing the first seven surveys was \$0.50, and payment for completing the eighth survey was \$7.00. Three hundred and forty participants completed the full study (47.9% female; M_{age} = 39.34).

In the first survey participants were randomly assigned to one of two conditions: an ‘outside-view’ budgeting condition or a control budgeting condition. In the outside-view condition, participants were asked to estimate their total spending for the past week then the next week. We expected budget estimates in this condition to be relatively high, because taking an outside-view (i.e., considering relevant past behavior) increases estimates of future behavior by prompting people to think of outcomes that would not come to mind otherwise (Buehler, Griffin and Ross 1994; Peetz and Buehler 2012). In the context of expenses, taking an outside-view also anchors budget estimates on recalled past spending, which should lead to higher budget estimates versus control because recalled past spending tends to be substantially higher than estimated future spending (Howard et al. 2022). In the control budgeting condition, participants were asked to estimate their total spending for the next week then the past week. Consistent with Studies 1 and 2, our expectation was that these participants would make optimistically low budget estimates for the next week. Thus, our overarching hypothesis for this study was that spending would be lower in the control budgeting condition than in the outside-view budgeting condition because lower budgets are associated with lower spending (H2).

Consistent with Peetz and Buehler (2009) and Ülkümen, Thomas, and Morwitz (2008), we operationalized budgets in this study as an estimate of future spending. We did this for three reasons. First, this standardizes the budgeting approach across participants. In contrast, some participants in Study 2 who were asked to budget may have tried to accurately estimate future spending, whereas others may have purposefully set an ambitious goal. Second, estimates of future behavior typically elicit a higher level of accuracy motivation than other judgments (Peetz and Buehler 2009). This makes our budget compliance and intervention tests in this study conservative. Finally, asking participants for an "estimate" rather than a "budget" minimizes the possibility that self-reported expenses in the daily expense diaries (described below) are underreported due to a demand effect.

We used a one-week time frame in Study 3 rather than one month because a shorter time frame reduces the probability of encountering atypically high expenses (e.g., an emergency medical bill) and/or impulsive purchases (e.g., a big night out to celebrate a new publication). Therefore, the one-week time frame also helps make our budget compliance test in Study 3 conservative. We used total spending as our dependent variable in this study (rather than total discretionary spending, as in Study 2) to eliminate the possibility that participants misunderstood which expenses they should be estimating.

The seven "diary" surveys were fielded at 9pm EST Monday – Sunday during the target week. Participants who did not complete a diary survey within 24 hours of it being launched were excluded from the remaining diary surveys. Each of these surveys asked participants to describe each expense they had incurred that day (e.g., "groceries") and report the amount of each expense (e.g., "$57.39"). This allowed us to sum each participant's total spending over the course of the week and compare it to their budget estimate at the start of the week. Of the two hundred and twenty-seven participants who were randomly assigned to the control budgeting condition, one hundred and seventy-six (77.5%) completed the full study; of the two hundred and twenty-three participants who were randomly assigned to the outside-view budgeting condition, one hundred and sixty-four (73.5%) completed the full study ($\chi^2 = .97$, $p = .33$). Participants who completed the full study did not differ significantly from participants who failed to complete the full study in terms of their recalled spending for the past week ($M_{complete}$ = \$229.27, 95% CI = [207.45, 253.38]; $M_{incomplete}$ = \$205.16, 95% CI = [172.52, 243.98]; $F(1, 443) = 1.17$, $p = .28$), estimated spending for the next week ($M_{complete}$ = \$219.07, 95% CI = [197.95, 242.48]; $M_{incomplete}$ = \$180.86, 95% CI = [151.08, 216.50]; $F(1, 444) = 3.36$, $p = .068$), or gender (completes = 47.9% female, incompletes = 42.7% female, $\chi^2 = .91$, $p = .34$), but they were somewhat older ($M_{complete} = 39.34$, $SD = 11.67$; $M_{incomplete} = 36.67$, $SD = 13.18$; $F(1, 448) = 4.05$, $p = .045$). As in Study 2, all spending variables were LN-transformed for inferential analysis then exponentiated so we can present summary statistics in dollars and cents.

Results and Discussion

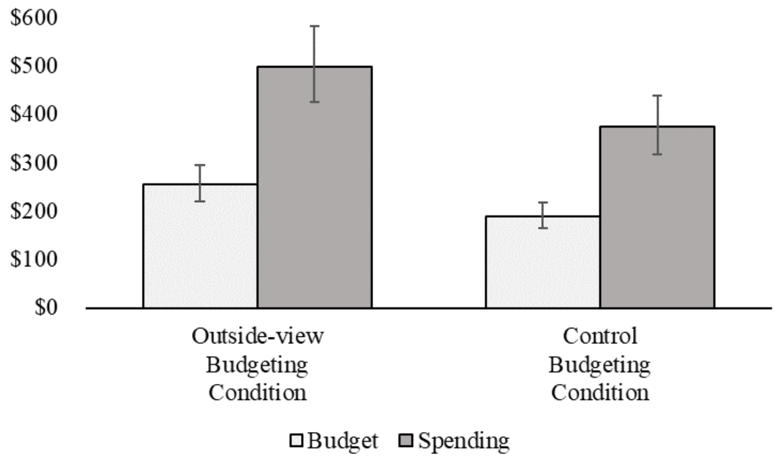
Spending estimates. As expected, an ANOVA revealed that participants in the control budgeting condition made significantly lower budget estimates for the next week than participants in the outside-view condition ($M_{control}$ = \$189.88, $CI_{95\%}$ = [165.64, 217.65]) $M_{outside\text{-}view}$ = \$255.44, $CI_{95\%}$ = [220.15, 296.43]; $F(1, 336) = 8.45$, $p = .004$, $\eta_p^2 = .03$). An ANOVA also revealed that participants in the control budgeting condition recalled significantly lower spending for the past week than participants in the outside-view condition ($M_{control}$ = \$196.21, $CI_{95\%}$ = [169.71, 226.85]; $M_{outside\text{-}view}$ = \$271.29, $CI_{95\%}$ = [237.37, 310.04]; $F(1, 335) = 10.42$, $p = .001$,

$\eta_p^2 = .03$). We did not make a prediction regarding the effect of condition on recalled past spending, because this study is concerned with the effect of budget estimates on future spending. However, this result is consistent with an anchor-based estimation process: in the same way that participants in the outside-view condition produced higher budget estimates for the next week by anchoring on recalled spending for the past week, participants in the control likely recalled lower spending for the past week by anchoring on their optimistic budget estimate for the next week. The Web Appendix presents a preregistered replication of our budgeting manipulation that supports this hypothesis.

Reported spending during the week of the study. Figure 8 plots mean reported spending in each condition next to the mean budget estimate in each condition. Supporting H2, an ANOVA revealed participants in the control budgeting condition reported significantly lower spending during the week of the study than participants in the outside-view condition ($M_{control}$ = \$374.32, $CI_{95\%}$ = [318.68, 439.70]; $M_{outside\text{-}view}$ = \$498.35, $CI_{95\%}$ = [426.92, 581.73]; $F(1, 337) = 6.37$, $p = .012$, $\eta_p^2 = .02$). Additionally, participants in both conditions spent significantly more than their budget estimate (p's $< .001$), and the difference between budget estimates and reported spending did not differ between conditions ($p = .87$). This once again indicates that budgets influence spending even when compliance is weak. It is also noteworthy that in both conditions reported spending during the week of the study is significantly higher than recalled spending for the week preceding the study (p's $< .001$). This is consistent with the observation that recollections of past behavior tend to be somewhat optimistic (Buehler, Griffin, and Peetz 2010), and that this study was conducted at the end of the month, so participants may have been more likely to pay recurring bills during the week of the study than during the week before the study.

FIGURE 8: BUDGETED VS REPORTED SPENDING IN STUDY 3

Error bars represent 95% Confidence Intervals.



Mediation analysis. We next tested a mediation model with budgeting condition as the independent variable (outside-view = 0, control = 1), participants' spending during the study as the dependent variable, and budget estimates as the mediator. The indirect effect of condition on spending via budget estimates was significant (indirect effect = –.17, SE = .06, 95% CI = [.06, .30]). Specifically, the model confirms that budget estimates were significantly lower in the control budgeting condition than in the outside-view budgeting condition (b = –.30, 95% CI = [.10, .50]; t(336) = 2.91, *p* = .004), and it demonstrates that lower budget estimates are associated with lower reported spending, even while controlling for condition (b = .59, 95% CI = [.50, .70]; t(335) = 11.56, *p* < .001). Comparing the correlation between budgets and spending in each condition shows that the two correlation coefficients do not differ significantly ($r_{outside}(161)$ = 0.51, $r_{control}(173)$ = 0.56, z = .56, *p* = .29).

The results of Study 3 complement and extend Studies 1 and 2 by providing causal evidence in support of H2, and by showing that the effect of condition on spending is mediated by budget estimates. It is interesting that although the outside-view manipulation increased budget estimates versus control, it did not increase estimate accuracy, as other versions of it typically do (Buehler, Griffin, and Peetz 2010). We believe this is because our version of the outside-view manipulation uses a very light touch compared to those that have people deeply consider their past behavior, then explicitly encourage them to base their estimate on it (e.g., Buehler, Griffin, and Ross 1994). In contrast, our participants considered only one week of past behavior, which then served as an implicit anchor. Thus, it is perhaps unsurprising that budget estimates in our outside-view condition were not closer to actual spending.

GENERAL DISCUSSION

The present research was motivated by a question of substantive importance to consumers, the firms that serve them, and the scholars who study their behavior: do budgets actually work? Our findings suggest that they do, although imperfectly. Consumers tend to set wildly optimistic budgets, which makes budget compliance difficult. However, we find that lower, more optimistic budgets also lead to lower spending, even when budget compliance is weak. Moreover, our longitudinal analysis shows that the influence of budgets on spending is surprisingly persistent, and our psychographic analysis indicates that budget influence may be broadly generalizable across different types of consumers.

One practical, actionable recommendation based on our findings is that consumers who want to reduce their spending should set an optimistic budget and regularly track their expenses against it. A second recommendation is that consumers who fail to comply with their budget should not abandon it. In Study 1, app users were still decreasing their spending six months after setting a budget, even though their budget compliance remained imperfect. Finally, our results suggest that budgeting works for different types of consumers, different types of expenses, and different time frames, which suggests these recommendations may be broadly applicable. We next discuss the theoretical implications of our findings and directions for future research.

Consumer Budgeting Theory

Our findings have several implications for consumer budgeting theory. First, we demonstrate that budgets are optimistic as compared to past spending. However, we also show that relative budget optimism varies between budget categories. One reason for this variability could be that spending on dining and drinking is entirely discretionary, whereas spending on groceries and fuel is less so. Investigating the relationship between budget optimism and the nature of spending across categories (e.g., purely discretionary vs. entirely necessary; Putnam-Farr and Ghosh 2018) is an important direction for future research.

Second, we show that budgets are not as inflexible as past research suggests (Heath and Soll 1996). More specifically, we find that budget compliance is weak even when expenses are easy to track and categorize, and that this occurs at least in part because of budget optimism. However, importantly, we also find that lower, more optimistic budgets do lead to lower spending, even when budget compliance is weak. This is notable because prior work implies that a high degree of budget compliance is necessary for budgets to influence spending (Heath and Soll 1996). Prior work also suggests that budget violations can cause spending to rebound back to (or above) pre-budget levels (Soman and Cheema 2004). Although we do find evidence that higher budget compliance in one month is associated with lower spending in the next month, we do not find evidence of a rebound effect. An interesting direction for future research is to better understand when and why budget violations do or do not cause such an effect.

Third, our data indicate that the influence of budgets on spending is surprisingly persistent. This is notable because prior work implies that budgets may exert less influence on spending over time (Choe and Kan 2021; Soman and Cheema 2004). Potential explanations for the divergence between our results and prior work include using an app to budget versus setting a mental budget, and budgeting for categories versus individual expenses. Investigating these factors as potential moderators of the longitudinal effect of budgets on spending is an important direction for future research.

Fourth, our exploratory analysis of the relationship between psychographics and budget compliance reveals three results of particular interest. The first is that impulsiveness – measured in three different ways – is associated with lower budget compliance. However, surprisingly, this is not because more impulsive consumers display higher spending than less impulsive consumers, it is because they budget less, an effect which holds even when we control for pre-budget spending. Our interpretation of this result is that impulsive consumers are trying – but failing – to use budgets as a tool to reduce spending. We observe a very similar pattern of results for financially anxious consumers, which suggests that systematically examining the relationship between impulsiveness, financial anxiety, and budget compliance may be a fruitful direction for future research. Finally, we find that budget optimism is *not* associated with trait optimism, which is consistent with past research on plans and predictions (Buehler, Griffin, and Peetz 2010, Howard et al. 2022).

Fifth, we find that the effect of budgets on spending is consistent across several psychographics of theoretical interest. Of course, null results need to be interpreted with caution, but related work has also found that the effect of financial judgments on subsequent spending is not moderated by seemingly relevant individual differences (Howard et al. 2022). This implies that the influence of budgets on spending may be broadly generalizable, although we emphasize that more research is required regarding the relationship between psychographics and budget influence before this conclusion can be made with certainty.

More broadly, our review of the consumer budgeting literature reveals a great deal of variability in how budgets are defined. This accurately reflects the reality that budgets can represent different things to different people, but future research should examine the degree to which, for example, a planning versus prediction mindset affects budget optimism, compliance, and influence. Budget setting is almost certainly a multiply determined psychological process, and more work is required to understand its nuances.

Planning Fallacies and Prediction Biases

Research on planning and forecasting is often focused on reducing optimism (e.g., Buehler, Griffin, and Ross 1994; Howard et al. 2022; Peetz and Buehler 2012; Peetz et al. 2015). One reason for this is undoubtedly that de-biasing techniques can inform theory. However, this focus is also driven in part by the assumption that optimistic plans and predictions are detrimental, and that more realistic ones are beneficial. This assumption is undoubtedly true in some circumstances: for example, consumers who take out a mortgage or finance a car based on an optimistic budget estimate may find themselves in a financial bind that could have been avoided if they had made a more realistic estimate and decided to borrow less. However, the present research provides evidence that budget optimism is generally a good thing, contingent on expense tracking. Future work can and should look deeper into the boundary conditions of this effect, identify when or for whom budget optimism backfires, and more clearly delineate the types of circumstances in which optimism and realism are preferred.

Past research has shown that plans and forecasts either do not influence behavior (Buehler, Griffin, and Ross 1994; Peetz and Buehler 2009; Ülkümen, Thomas, and Morwitz 2008), or that their influence diminishes quickly (Buehler, Griffin, and Peetz 2010). However, we found in Study 1 that actual spending continues to move toward budgeted spending even six months after a budget is set. Moreover, this occurs regardless of whether the initial outcome is spending more than budget (as is the case for dining and drinking and groceries) or less than budget (as is the case for fuel). We believe one explanation for the difference between our results in this regard and past research is that Money Dashboard users engage in a relatively high degree of expense tracking (i.e., behavior monitoring). Thus, their budget becomes a very salient reference point against which they can compare their spending behavior. Moreover, because the app makes past spending behavior easily accessible, it is likely that consumers are motivated by the observation that they are spending less than they used to, rather than becoming demotivated by the observation that they have spent more than they budgeted (Soman and Cheema 2004). Future research on the temporally extended influence of plans and predictions on behavior should test this conjecture. It is also well worth investigating the extent to which behavior-monitoring can ameliorate related phenomena such as the planning fallacy.

DATA COLLECTION INFORMATION

Data for Study 1 was collected in the UK by Money Dashboard from 2014 to 2016. It was analyzed by Marcel Lukas, and it is currently stored at the University of St. Andrews. Data for Studies 2, 3, and the Web Appendix was collected online and analyzed by Ray Howard in 2020, 2018, and 2022, respectively. These data sets are currently stored at Texas A&M University. The data sets for Studies 1 and 2 are proprietary and cannot be archived at a third-party repository. The data for Study 3 and the replication study in the web appendix have been archived in a project folder on the Open Science Framework.

REFERENCES

Abadie, Alberto, and Guido W. Imbens (2011), "Bias-corrected matching estimators for average treatment effects," *Journal of Business & Economic Statistics*, 29 (1), 1-11.

Abeler, Johannes, Armin Falk, Lorenz Goette, and David Huffman (2011), "Reference points and effort provision," *American Economic Review*, 101 (2), 470-92.

Allen, Eric J., Patricia M. Dechow, Devin G. Pope, and George Wu (2017), "Reference-dependent preferences: Evidence from marathon runners," *Management Science,* 63 (6), 1657-1672.

Ariely, Dan, George Loewenstein, and Drazen Prelec (2003), "“Coherent arbitrariness”: Stable demand curves without stable preferences," *The Quarterly Journal of Economics*, 118 (1), 73-106.

Barber, Brad M., and Terrance Odean (2002), "Online investors: do the slow die first?," *The Review of Financial Studies*, 15 (2), 455-488.

Bell, Amy (2019), “Six Reasons Why You Need a Budget,” https://www.investopedia.com/financial-edge/1109/6-reasons-why-you-need-a-budget.aspx

Browne, Melissa (2020), “*Budgets Don't Work, But This Does…*,” Allen and Unwin.

Buehler, Roger, Dale Griffin, and Johanna Peetz (2010), "The planning fallacy: Cognitive, motivational, and social origins," In *Advances in Experimental Social Psychology*, 43, pp. 1-62. Academic Press, 2010.

Buehler, Roger, Dale Griffin, and Michael Ross (1994), "Exploring the" planning fallacy": Why people underestimate their task completion times," *Journal of Personality and Social Psychology*, 67 (3), 366.

Caldwell, Miriam (2019), “Reasons Why You Should Budget Your Money,” https://www.thebalance.com/reasons-to-budget-money-2385699

Chapman, Gretchen B (1996), "Temporal discounting and utility for health and money," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22 (3), 771.

Cheema, Amar, and Dilip Soman (2006), "Malleable mental accounting: The effect of flexibility on the justification of attractive spending and consumption decisions," *Journal of Consumer Psychology*, 16 (1), 33-44.

Choe, Yuna, and Christina Kan (2021), "Budget depreciation: when budgeting early increases spending," *Journal of Consumer Research*, 47 (6), 937-958

Choe, Yuna, and Christina Kan (2018), "Gift Budget Adherence and Price Discounts," *ACR North American Advances,* 46, 902-902.

Cook, Thomas K. and Donald T. Campbell (1979), *Quasi-Experimentation: Design and Analysis Issues for Field Settings,* Chicago: Rand McNally.

Credit Counselling Society (2019), “What is Budgeting? What is a Budget?,” https://www.mymoneycoach.ca/budgeting/what-is-a-budget-planning-forecasting

Datta, Hannes, George Knox, and Bart J. Bronnenberg (2018), "Changing their tune: How consumers’ adoption of online streaming affects music consumption and discovery," *Marketing Science*, 37 (1), 5-21.

Dittmar, Helga, and Rod Bond (2010), "I want it and I want it now: Using a temporal discounting paradigm to examine predictors of consumer impulsivity," *British Journal of Psychology*, 101 (4), 751-776.

Elkins, Kathleen (2018), “39-year-old retired millionaire: ‘Budgets don’t work’—do this instead," https://www.cnbc.com/2018/12/26/39-year-old-retired-millionaire-budgets-dont-work-do-this-instead.html

Fernbach, Philip M., Christina Kan, and John G. Lynch Jr (2015), "Squeezed: Coping with constraint through efficiency and prioritization," *Journal of Consumer Research*, 41 (5), 1204-1227.

Ganti, Ahkilesh (2021), "What Is a Budget?," https://www.investopedia.com/terms/b/budget.asp

Hardisty, David J., Kirstin C. Appelt, and Elke U. Weber (2013), "Good or bad, we want it now: Fixed-cost present bias for gains and losses explains magnitude asymmetries in intertemporal choice," *Journal of Behavioral Decision Making*, 26 (4), 348-361.

Heath, Chip (1995), "Escalation and De-escalation of Commitment in Response to Sunk Costs: The Role of Budgeting in Mental Accounting," *Organizational Behavior and Human Decision Processes*, 62, 38-54.

Heath, Chip., and Jack B. Soll (1996), "Mental budgeting and consumer decisions," *Journal of Consumer Research*, 23 (1), 40-52.

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

Howard, Chuck, David J. Hardisty, Abigail B. Sussman, and Marcel F. Lukas (2022), "Understanding and Neutralizing the Expense Prediction Bias: The Role of Accessibility, Typicality, and Skewness," *Journal of Marketing Research* 59 (2), 435-452.

Imbens, Guido W., and Donald B. Rubin (2015), *Causal inference in statistics, social, and biomedical sciences*, Cambridge University Press.

InCharge.org (2021), "How To Create a Successful Budget," https://www.incharge.org/financial-literacy/budgeting-saving/how-to-make-a-budget/

Kahneman, Daniel, and Amos Tversky (1979), “Intuitive prediction: Biases and corrective procedures,” *TIMS Studies in Management Science, 12,* 313-327.

Kőszegi, Botond, and Matthew Rabin (2006), "A model of reference-dependent preferences," *The Quarterly Journal of Economics*, 121 (4), 1133-1165.

Leuven, Edwin, and Barbara Sianesi (2003), "PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing," (2003).

Lusardi, Annamaria, and Olivia S. Mitchell, (2011), "Financial literacy around the world: an overview," *Journal of Pension Economics & Finance*, 10 (4), 497-508.

Lynch Jr, John G. (1982), "On the external validity of experiments in consumer research," *Journal of Consumer Research,* 9 (3), 225-239.

Lynch Jr, John G., Richard G. Netemeyer, Stephen A. Spiller, and Alessandra Zammit (2010), "A generalizable scale of propensity to plan: The long and the short of planning for time and for money," *Journal of Consumer Research*, 37 (1), 108-128.

Mathera, Richard, and Lindsay Juarez (2021), “Do Budgets Actually Help You Save Money? Results from Our Large-Scale Budgeting Experiment,” Common Cents Lab, Duke University, https://irrationallabs.com/blog/money-budgeting-experiment/

Novemsky, Nathan, and Daniel Kahneman (2005), "The boundaries of loss aversion," *Journal of Marketing research*, 42 (2), 119-128.

Olen, Helaine (2015), “Toss Your Budget - Why a pillar of personal finance isn’t nearly as essential as we think,” http://www.slate.com/articles/business/the_bills/2015/06/personal_budgets_are_overrated_seriously.single.html

Oxford English Dictionary (2021), https://languages.oup.com/google-dictionary-en/

Peetz, Joanna, and Roger Buehler (2009), “Is there a budget fallacy? The role of savings goals in the prediction of personal spending,” *Personality and Social Psychology Bulletin*, 35 (12), 1579-1591.

Peetz, Johanna, and Roger Buehler (2012), "When distance pays off: The role of construal level in spending predictions," *Journal of Experimental Social Psychology*, 48 (1), 395-398.

Peetz, Johanna, Roger Buehler, Derek J. Koehler, and Ester Moher (2015), "Bigger not better: Unpacking future expenses inflates spending predictions," *Basic and Applied Social Psychology*, 37 (1), 19-30.

Peetz, Johanna, Melanie Simmons, Jingwen Chen, and Roger Buehler (2016), "Predictions on the go: Prevalence of spontaneous spending predictions," *Judgment and Decision Making*, 11 (1), 48-61.

Putnam-Farr, Eleanor, and Anastasiya Pocheptsova Ghosh (2018), "Turning “Expenses” Into “Bills”: How Spending Categorization Impacts Budget Optimism and Likelihood of Success," *ACR North American Advances.*

Rick, Scott I., Cynthia E. Cryder, and George Loewenstein (2008), "Tightwads and spendthrifts." *Journal of Consumer Research*, 34 (6), 767-782.

Rook, Dennis W., and Robert J. Fisher (1995), "Normative influences on impulsive buying behavior," *Journal of Consumer Research*, 22 (3), 305-313.

Rosenbaum, Paul R., and Donald B. Rubin (1983), "The central role of the propensity score in observational studies for causal effects," *Biometrika*, 70 (1), 41-55.

Scheier, Michael F., Charles S. Carver, and Michael W. Bridges (1994), "Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): a reevaluation of the Life Orientation Test," *Journal of Personality and Social Psychology*, 67 (6), 1063.

Sharot, Tali (2011), "The Optimism Bias," *Time Magazine*, Saturday, May 28th, 2011. http://content.time.com/time/health/article/0,8599,2074067,00.html

Soman, Dilip, and Amar Cheema (2004), "When goals are counterproductive: The effects of violation of a behavioral goal on subsequent performance," *Journal of Consumer Research*, 31 (1), 52-62.

Sussman, Abigail B., and Adam L. Alter (2012), "The exception is the rule: Underestimating and overspending on exceptional expenses," *Journal of Consumer Research*, 39 (4), 800-814.

Tangney, June P., Roy F. Baumeister, and Angie Luzio Boone (2004), "High self-control predicts good adjustment, less pathology, better grades, and interpersonal success," *Journal of Personality*, 72 (2), 271-324.

Thaler, Richard H. (1985), "Mental Accounting and Consumer Choice," *Marketing Science*, 4, 119-214.

Thaler, Richard H. (1999). "Mental accounting matters," *Journal of Behavioral decision making*, 12(3), 183-206.

Thaler, Richard H., & Hal M. Shefrin (1981), "An economic theory of self-control," *Journal of Political Economy*, 89 (2), 392-406.

Ülkümen, Gülden, Manoj Thomas, and Vicki G. Morwitz (2008), "Will I spend more in 12 months or a year? The effect of ease of estimation and confidence on budget estimates," *Journal of Consumer Research*, 35 (2), 245-256.

Zhang, C. Yiwei, and Abigail B. Sussman (2018), "Perspectives on mental accounting: An exploration of budgeting and investing," *Financial Planning Review*, 1, e1011.

HEADINGS LIST

1) **THE BUDGETING PROCESS**
1) **THE PSYCHOLOGY OF BUDGETING**
1) **BUDGETS AS REFERENCE POINTS**
1) **BUDGET OPTIMISM**
1) **BUDGET COMPLIANCE**
1) **BUDGET INFLUENCE**
1) **STUDY 1: PERSONAL FINANCE APP DATA**
2) Spending Data
2) Budgeting Data
2) Analysis and Results
3) *Budget optimism*
3) *Hypothesis 1: Budget Compliance*
3) *Hypothesis 2: Budget Influence*
3) *Hypothesis 3: Longitudinal Budget Influence*
3) *Hypothesis 4: Budgeters vs. Non-budgeters*
2) Discussion
1) **STUDY 2: FIELD EXPERIMENT**
2) Method
3) *Savings goal*
3) *General budgeting behavior*
3) *Typical discretionary spending*
3) *Preferred payment method*
3) *Financial anxiety*
3) *Impulsive buying behavior*
3) *Demographics*
2) Analysis
2) Results
3) *Hypothesis 1: Budget Compliance*
3) *Hypothesis 2: Budget Influence*
3) *Hypothesis 4: Budgeters vs. Non-Budgeters*
2) Discussion
1) **STUDY 3: THE FINANCIAL DIARY STUDY**
2) Method
2) Results and Discussion
3) *Spending estimates*
3) *Reported spending during the week of the study*
3) *Mediation analysis*
1) **GENERAL DISCUSSION**
2) Consumer Budgeting Theory
2) Planning Fallacies and Prediction Biases