Mental Budgeting and Consumer Decisions

CHIP HEATH
JACK B. SOLL *

Consumers often set budgets for categories of expenses (e.g., entertainment) and track expenses against their budget. Because budgets cannot perfectly anticipate consumption opportunities, people may earmark too much or too little money for a particular category. This leads them to overconsume or underconsume goods in that category. The results of three studies suggest that consumers do indeed set budgets and that budgeting may lead to underconsumption. To show that consumers track expenses, the studies demonstrate that budgeting effects are larger for purchases that are highly typical of their category. Such purchases reduce the amount people spend in a category and block the purchase of other typical items. The studies control for satiation and income effects; thus, budgeting adds predictive power to standard economic consumer theory.

The effect of his scientific budget-planning was that he felt at once triumphantly wealthy and perilously poor. (SINCLAIR LEWIS, Babbitt)

Like many of us, George Babbitt, Sinclair Lewis's famous character, engages in budget planning. The result for him, as for us, is to produce artificial experiences of wealth and poverty. Consider the following scenarios. Mr. P recently went shopping for a pair of slacks. When he could not find any slacks he liked, he spent a similar amount of money on a sweater that he normally would not have purchased. Ms. C reluctantly declined a Sunday invitation to dinner because she had "spent too much money" on tickets to the theater two days earlier. On questioning, she admitted that she would have enjoyed the dinner and was well able to afford it, but she felt compelled to decline because of the earlier theater expense. Neither person was entirely content. After her experience of "perilous" poverty, Ms. C found herself declining an enjoyable dinner with friends, and after his experience of "triumphant" wealth, Mr. P found himself the owner of an unnecessary sweater.

We interpret both examples as outcomes of a pervasive process in consumer behavior: mental budgeting. We argue that consumers budget portions of their total resources to separate mental accounts (e.g., entertainment or household expenses) and then track expenses against the budgets. As expenses are incurred, they deplete the funds available in their account, which makes future purchases less likely.

Mental budgeting is consistent with well-known research on mental accounting (Henderson and Peterson 1992; Kahneman and Tversky 1984; Thaler 1980, 1985) that demonstrates that people use resources differently depending on how they are labeled. We extend this literature by showing how budget setting and expense tracking alter consumer choice.

First, we highlight that budget setting leads people to overconsume some goods and underconsume others. Because budgets are set before consumption opportunities arise, they sometimes overestimate or underestimate the money required for a particular account. Mr. P had allocated money to his clothing budget to purchase slacks. After finding no acceptable slacks, he could have reallocated that money elsewhere; instead, he took home a sweater he ordinarily would not have purchased. Ms. C had allocated too little money to her entertainment budget to cover her uncommonly rich entertainment opportunities. Instead of adding money from other accounts, she declined a dinner she would have enjoyed.

Second, we highlight that expense tracking implies that some expenses are more likely to produce over- or underconsumption. As people track their expenses, expenses that are relatively easy to categorize—those that are more typical examples of their categories—will be the most subject to the rigors of budgeting. Theater tickets are a very representative entertainment expense, and we speculate that Ms. C might not have avoided Sunday dinner if she had spent an identical amount

*Chip Heath is associate professor of behavioral science and Jack Soll is a doctoral candidate in behavioral science at the Graduate School of Business, University of Chicago, 1101 East 58th Street, Chicago, IL 60637. For helpful comments on earlier drafts, we thank Colin Camerer, John Deighton, John Gourville, Steve Hoch, Joan Meyers-Levy, Mitch Petersen, Dilip Soman, the associate editor, and two reviewers. Special thanks to Josh Klayman who gave us very thoughtful suggestions at several stages in this research.
early in the weekend on a less representative entertainment expense—for example, consuming an expensive bottle of wine or attending a charity costume ball.

**PREVIOUS RESEARCH ON BUDGETING AND MENTAL ACCOUNTING**

In reviewing mental-accounting research, Thaler (1993) highlighted two fundamental questions. First, how do people group and label resources; and second, how does grouping affect personal satisfaction (e.g., Is it better to aggregate or segregate gains or losses?)? The second question has received substantial theoretical and empirical attention. Research has shown, for example, that people like to separate small gains and that although they might be happier if they integrated losses (Thaler 1985), they typically do not (Linville and Fischer 1991; Thaler and Johnson 1990).

The question of labeling has been addressed less systematically (Thaler 1993). Previous literature has not provided much detail about how people label their resources, but it has shown that labels affect resource use. Shefrin and Thaler (1988), for example, theorized that people distinguish between wealth in categories like “current spendable income” and “current assets” and are more willing to consume an increase in current income (e.g., a raise) than an increase in current assets (e.g., home equity). People are more likely to purchase a vacation when they receive $2,000 as a gift than as a work bonus (Henderson and Peterson 1992), and they are more likely to purchase frivolous goods with winnings from a football pool than with an equivalent amount in overtime pay (O’Curry 1996). These results indicate that labels sometimes lead consumers to violate fungibility, the principle that money is money.

As can be inferred from the examples above, previous work on mental accounting has been primarily designed to show that labels matter. Consequently, many of the results in the literature have focused on relatively rare events (gifts, bonuses, or winnings from a football pool) for which the environment provides a ready-made label for income instead of examining what happens when people generate their own labels. We suggest that people generate two kinds of labels that affect their decisions as consumers. First, people label money as relevant for a certain class of goods, and second, they label the goods as relevant for a certain pool of money. We refer to these processes, respectively, as the budget-setting and the expense-tracking processes.

We believe that consumers have a strong tendency to label money even in situations in which the environment does not provide a label like “bonus,” “home equity,” or “gift.” In fact, people seem to engage in vivid, often physical labeling of the most banal and undifferentiated source of income—wages. For more than 50 years, studies have documented that consumers divide their wages into separate pools and dedicate different pools to different kinds of expenses. For example, in his study of unemployed workers in the 1930s, Bakke (1940, pp. 142–143) documented “all sorts” of budgeting arrangements “from those that make use of small envelopes in which particular sums are put, or the series of china pitchers in the cupboard, to the ones who have checking accounts and carry on a fairly complicated bookkeeping.” In their study of households in the 1950s, Rainwater, Coleman, and Handel (1959) described a process of “tin can accounting” in which households allocated funds into separate accounts symbolized by separate “envelopes, various drawers, or tin cans.” (These references came to our attention in a very interesting paper by Zelizer [1993] on the sociology of money.) Thaler (1985), in a footnote, describes similar behavior among consumers in the early 1980s.

People not only label money; they also label expenses. Previous researchers have typically defined mental accounts as temporal units, assuming, for example, that events on the same day are coded in the same mental account, whereas events on different days are coded in different accounts. Defining accounts in this way makes it easier to study the effects of separating and integrating costs and benefits (Linville and Fischer 1991; Thaler 1985; Thaler and Johnson 1990). However, by studying the crisp categorizations allowed by temporal labels, researchers have neglected the more complex tracking process required when consumers group their purchases by categories that do not have clear boundaries (e.g., “household items” or “entertainment”). Henderson and Peterson (1992) applied cognitive research on categorization to understand the structure of mental accounts. We follow their lead by exploring how the expense-tracking process changes depending on how easy it is for people to assign their expenses to categories.

**A THEORY OF MENTAL BUDGETING FOR EXPENSES**

We propose a model of the consumer similar to Heath’s (1995) model of the investor. According to the model, consumers set budgets for various expense accounts—for example, classes of expenses like household purchases, entertainment, clothing, or food (see Thaler [1985] for a similar point). As they spend money, they assign their expenses to appropriate accounts and periodically recompute the amount of money remaining in their budgets. When a particular budget is depleted, they resist further expenses on items in that category. Thus, there are two major parts of the budgeting process: setting a budget and tracking their ongoing expenses against the budget.

**Setting the Budget**

We assume that consumers typically set budgets in advance of actual consumption. Self-control mecha-
nisms frequently separate planning from doing, because this allows a rational “planner” to constrain the actions of an overly impulsive “doer” (Schelling 1992; Shefrin and Thaler 1988). Setting budgets in advance may also simplify computational costs by reducing the number of alternatives that compete for a particular sum of money (Simon 1947) and by making them more comparable (Johnson 1984).

However, because consumption opportunities change over time, a preset budget is unlikely to allocate exactly the right amount of money for the opportunities that actually arise. We hypothesize that when the predetermined budget earmarks too much or too little money for a particular class of expenses, people stick closely to their budget and resist transferring funds across accounts. For a budget to be an effective self-control mechanism, it must be at least somewhat inflexible. If people were to allow themselves to reallocate at a whim, they might be tempted to engage in inappropriate reallocations (spending the rent money at a nightclub) as well as appropriate ones (spending money on dinner rather than another new sweater). However, this inflexibility comes at a cost. If consumers were to allow themselves to reallocate resources when opportunities change, they could easily maximize their satisfaction on the basis of whatever consumption opportunities actually arise rather than the ones they anticipate before the fact.

An inflexible budget will systematically alter consumer choices. When people budget too little money, they may underconsume goods they desire. When they budget too much, they may overconsume goods that they desire less. These predictions clearly differ from those of economic consumer theory, which assumes that people always consume an optimal quantity of each good. We define under- and overconsumption net of the satiation and income effects described by economic models. For example, we predict that when a budget is too low, consumers will avoid buying goods in that account even though they are not tired of consuming goods in the account (satiation) and even though they do not lack money to consume them (income).

Using our model, we can interpret some demonstrations of mental-accounting effects as evidence of overconsumption. For example, when the environment labels money, people are unusually likely to consume goods that are related to the label (Henderson and Peterson 1992; O’Curry 1996). O’Curry (1996) found that when asked to imagine a cross-category price discount on beer, beer drinkers stated that they would use the extra funds to buy higher quality beer, but they would not do so when they receive an equivalent amount of wealth as a gift. We interpret this pattern as evidence that consumers overconsume beer when the price falls and the beer budget contains a surplus. Thus, overconsumption is a clear implication of budgeting, and several extant demonstrations of income source effects (Henderson and Peterson 1992; O’Curry 1996) can be interpreted as evidence of this effect.

Heath (1995) has shown underconsumption in the domain of investments. In his studies, people adhere to their budgets and invest too little in promising projects, an effect opposite to the traditional concern in the organizational behavior literature with “entrapment” or “escalating commitment” (e.g., Brockner and Rubin 1985; Staw and Ross 1989). In general, there are fewer empirical results to support underconsumption than to support overconsumption. In addition, consumers are probably more likely to react to salient income shocks than surpluses. Thus, in our empirical work below, we concentrate on underconsumption.

Tracking Expenses

In previous experimental studies of how investors budget their expenses (Heath 1995), it was easy for subjects to track expenses because expenses were salient and the context made clear which investments went with which projects. On the other hand, in consumer decisions, people must remember a variety of purchases and assign them to their proper accounts. When expenses are easy to overlook or hard to categorize, budgeting decisions will be strongly affected by the expense-tracking process.

It is useful to divide the tracking process into two stages: (1) expenses must first be noticed and (2) then assigned to their proper accounts. An expense will not affect a budget if either stage fails. To label these stages, we borrow terminology from financial accounting, in which the accounting process is also divided into two stages. Expenses must be booked (i.e., recorded in the accounting system) and posted (i.e., assigned to a specific expense account). Each process depends on a different cognitive system. Booking depends on attention and memory. Posting depends on similarity judgments and categorization.

1When income is fixed, overconsumption of one class of goods necessarily implies that another class of goods is being underconsumed (if it is assumed that future consumption, or savings, is a good). Here we focus separately on each phenomenon, in part because this simplifies our presentation but also because we suspect that psychological experience also separates the effects. If people were to experience simultaneously the effects of penury and profligacy, they might transfer money across accounts and solve their conflicting problems. However, because of limited attention, people are probably aware of each effect at different times.

2We will concentrate in this article on posting. However, for those who are interested in cognitive resource allocation, the booking process is a potentially interesting subject of study, especially given evidence that consumers are often unable to remember what prices they paid for items (Dickson and Sawyer 1990). Expenses vary in salience, and some expenses are probably more likely to be booked than others. For example, opportunity costs may be less likely to be booked than out-of-pocket costs, small purchases than large purchases, and credit card purchases than cash.
We do not make any strong assumptions about whether people track their expenses constantly or sporadically. Judgments and choices are often constructed on-line (Payne, Bettman, and Johnson 1992; Simonson and Tversky 1992). Regardless of whether records are constantly updated or constructed when the situation demands, we argue that people must book and post expenses in order to track them. For example, if the mental books are reconstructed on-line, people must recall recent expenses (booking) and assign them to their proper accounts (posting). After people establish the current state of their mental accounts, they can then imagine how an additional purchase would affect their budget.

In the current article, we focus on the posting process. Posting requires people to decide how to categorize an expense that may vary in its relevance for existing accounts. Like Henderson and Peterson (1992), we argue that the categorization decision is driven by the same cognitive processes that people use to categorize other objects and events (see, e.g., Rosch 1975). Mental accounts, like most categories that people use, typically cannot be described by a set of necessary and sufficient features, and people may categorize items on the basis of many dimensions. An expense may be assigned to a particular account because it meets similar goals (Barsalou 1991) or because it has similar purchase features such as magnitude (“things you can do for five dollars”), format (“things you pay for with a credit card”), or location (“things you buy at the electronics store”).

Because similarity is multidimensional, it might seem difficult to make any useful predictions about posting. However, for the current studies, we focus on a very general prediction. Categorization research has long emphasized that some category items represent the category more perfectly. For example, attending a play is a very good example of an entertainment expense even though it may be a relatively infrequent one. In general, representative (or typical) members of a category are more easily learned, classified, and remembered. (For a review, see Barsalou [1991]). On the basis of this evidence, we argue that typical expenses will be easier to book and post than less typical expenses.

Combining the expense-tracking process with the budget-setting process, we predict that typical goods will be most subject to budgeting constraints. Because typical goods are particularly likely to be posted to their expense account, they will affect budgets much more than goods that are less typical. Below, three studies document this effect.

Experimental Evidence

In the following studies, we show that people under-consume in an account after making an unexpected purchase in that account. Our major hypothesis is that people will be more likely to under-consume after purchasing an unplanned item that is highly typical of the account. This prediction differs from economic theory, which predicts that underconsumption does not exist. Economic theory also makes no prediction about how typical purchases might differ from less typical purchases.

To show that mental budgeting adds predictive power to the standard economic model, we must control for the satiation and income effects predicted by economic consumer theory. In the studies below, we estimate these effects individually for each subject. For example, in studies 1 and 2, we control for income effects by assessing whether each subject will undertake future expenses after experiencing a set of exogenous, unplanned income shocks (e.g., receiving a parking ticket). If a person spends $50 on entertainment after paying a $20 parking ticket, then income effects do not explain why the individual spends less than $50 after purchasing a $20 sports ticket. We control for satiation by assessing whether each subject will undertake future expenses after being given the same good that was purchased. If a person spends $50 on entertainment after being given a $20 sports ticket, satiation does not explain why that individual spends less than $50 after purchasing the sports ticket. Note that the gift measure also controls for constraints on other resources. For people who budget their time as well as their money, consuming the gift takes the same time as consuming the purchase.

From a psychological viewpoint, note that both budget setting and expense tracking are required to predict that underconsumption will vary with typicality. To see this, consider what would happen if only one process were present. First, suppose that people set an overall budget (e.g., $200 for all expenses this week) but that they do not track expenses by assigning them to categories. In this case, all expenses belong to one large category and would be equally likely to cause underconsumption. Alternatively, suppose that people track their expenses but do not budget. In this case consumers would not underconsume. They might categorize their expenses in different accounts, but if they wanted to purchase an item in an overdrawn account, they would simply replenish the funds in that account.

The gift may contribute a positive income effect, because the subject is “wealthier” after receiving the gift. Mental budgeting predicts no income effect. Because people receive the use of an item rather than the monetary value of the item, their budget should remain unaltered. Economic theory predicts a small income effect that will be spread across categories. Because this effect should be constant across prototypical and nonprototypical items, this predicted effect should not interfere with our main hypothesis.
**STUDY 1: UNDERCONSUMPTION OF A CATEGORY**

In this experiment we ask people how much they would spend on various classes of expenses (food, clothing, entertainment) and then ask them how much they would spend on that class of expenses after purchasing a particular item. We systematically vary the typicality of the purchase item, and we expect, on the basis of the hypothesis above, that more people will underconsume items in the category after purchasing typical items than after purchasing nontypical items.

Our general hypothesis addresses the relationship between underconsumption and typicality, but it does not make predictions about the magnitude of underconsumption. If people underconsume, they may do so by the same amount for all purchases—a $20 costume purchase may cause someone to underconsume clothing as much as a $20 sweatshirt purchase. However, it seems reasonable that the amount of underconsumption may also vary with typicality. People may, for example, divide an expense and post different portions to different accounts. For typical purchases, it is quite likely that people will post the entire dollar amount to its account. For less typical purchases, however, part of the purchase price may be assigned to an account and part may remain unassigned. This study allows us to examine whether the amount of underconsumption increases with the typicality of the purchase.

**Method**

**Subjects.** Subjects were 29 MBA students who were paid $7 for their participation in the experiment. The experiment lasted about 30 minutes.

**Procedure.** Subjects were asked how much they spent per week on entertainment and food and per month on clothing. Pilot work showed that most people had mental accounts with these labels and that they tended to monitor food and entertainment purchases against a weekly budget and clothing against a monthly budget. After subjects estimated the amount they spent in each category, they specified how much they would spend in the category for the remainder of the week (or month) after experiencing a particular event.

Below is the introductory wording and two events for the entertainment category.

**Approximately how much do you spend per week on entertainment?**

Imagine that it is the beginning of the week. Each question below asks how much you would spend on entertainment after a certain event occurs. If the event would have no impact on your spending, feel free to write down the same amount that you wrote down above.

**EVENT:** You go out to dinner with friends. You spend $20.

How much would you spend on entertainment for the remainder of the week? ______

**EVENT:** You are given the dinner above.

How much would you spend on entertainment for the remainder of the week? ______

Table 1 lists the nine $20 purchases that were used in the experiment (e.g., spending $20 on dinner with friends or buying a $20 sweatshirt with a college logo). Subjects specified how much they would spend in the entertainment, food, and clothing categories after making each purchase. To assess satiation effects, subjects also specified how much they would spend in each category after being given each item. To assess income effects, subjects specified how much they would spend in each category after experiencing four unexpected $20 expenses that were unrelated to any of the three categories. These four expenses were as follows: replacing a dish that you dropped, getting a flu inoculation, paying a parking ticket, and discovering that today is a friend’s birthday and purchasing a gift. We included four different shocks to average out any idiosyncratic reactions to a particular event.

Thus, for each of three categories (food, entertainment, and clothing) subjects rated the impact of 22 events on their future spending in that category—nine $20 purchases, nine $20 gifts, and four $20 income shocks. To make the task easier, subjects dealt with only one category at a time and answered all 22 questions for that category before moving to the next category. However, the 22 events were presented in a different random order for each category, and the order of the three categories was counterbalanced across subjects. The only constraint on the ordering of items in cate-

<table>
<thead>
<tr>
<th>TABLE 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TYPICALITY RATINGS FOR STUDIES 1 AND 2</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Entertainment</th>
<th>Food</th>
<th>Clothes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Studies 1 and 2:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sports ticket</td>
<td>5.88 (25)</td>
<td>...</td>
</tr>
<tr>
<td>In-line skate rental</td>
<td>4.12 (25)</td>
<td>...</td>
</tr>
<tr>
<td>Party snacks</td>
<td>2.88 (20)</td>
<td>4.08 (26)</td>
</tr>
<tr>
<td>Pizza</td>
<td>2.00 (16)</td>
<td>5.77 (26)</td>
</tr>
<tr>
<td>Dinner out</td>
<td>6.15 (26)</td>
<td>4.46 (25)</td>
</tr>
<tr>
<td>Wine</td>
<td>2.88 (23)</td>
<td>3.08 (26)</td>
</tr>
<tr>
<td>Sweatshirt</td>
<td>.35 (4)</td>
<td>...</td>
</tr>
<tr>
<td>Gloves</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Costume</td>
<td>2.36 (17)</td>
<td>...</td>
</tr>
<tr>
<td><strong>Additional items for study 3:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boat tour</td>
<td>5.20 (26)</td>
<td>...</td>
</tr>
<tr>
<td>Blue jeans</td>
<td>1.16 (11)</td>
<td>...</td>
</tr>
<tr>
<td>Watch</td>
<td>1.12 (9)</td>
<td>...</td>
</tr>
<tr>
<td>Salmon</td>
<td>2.28 (16)</td>
<td>3.08 (26)</td>
</tr>
</tbody>
</table>

**Note.**—Values given are the average prototypicality ratings on a seven-point scale (7 = very typical). Parentheses indicate the number of people (N = 26) who rated the item as at least somewhat typical. Ellipses indicate that no one rated that item as typical of that account.
The nine purchase items were developed by asking a separate group of 26 students to rate 19 purchases on how typical they were of the food, entertainment, and clothing categories. Our instructions stressed that “typical” was meant to refer to a “particularly good example of the category,” not necessarily a frequent one. We selected purchases that had approximately the same cost (around $20) and that could not be considered necessities and thus were less likely to be planned (e.g., groceries would be a planned food purchase). The instructions for the typicality rating task were similar to the standard instructions in the categorization literature (e.g., Mervis and Rosch 1975). Subjects rated the 19 purchases on a 1 (very low typicality) to 7 (very typical) scale but had the option of marking an item with an X if they felt the item was not a member of the category.

We selected three items that were rated as typical of each of the three categories (food, clothing, and entertainment). To ensure some variance in the independent variable, we chose the three items for each category so that their average typicality ratings differed by approximately one scale point. Table 1 reports the typicality ratings for each purchase and category. To compute the overall typicality rating, X responses were replaced with zeros.

**Analysis.** The following example demonstrates the analysis we performed for each subject: Susan normally spends $50 per week on entertainment. She reports that if she purchased a $20 sports ticket, she would spend $42 on entertainment during the remainder of the week. If she were given the sports ticket, she would spend $45. Using these numbers to illustrate, we define the purchase, satiation, income, and mental budgeting effects as follows.

The purchase effect (P) is the difference between what a person would normally spend in the category and what that person would spend after making a particular purchase (P = 50 - 32 = 18). Susan is much less willing to spend on entertainment after the purchase of a typical entertainment item. However, to determine whether this is an effect of mental budgeting, we must separate out satiation and income effects.

We estimate satiation (S) by taking the difference between what Susan would normally spend and the amount she would spend after experiencing the random income shocks (we take the average of the four events; S = 50 - 42 = 8). Similarly, we estimate the income effect (I) by taking the difference between what Susan would normally spend and the amount she would spend after experiencing the random income shocks (we take the average of the four events; I = 50 - 45 = 5).

For each purchase, we calculate underconsumption (U) by subtracting the effects of income and satiation from the purchase effect (U = P - S - I = 18 - 8 - 5 = 5). Note that these estimates are subject to noise. It is possible, for example, to end up with negative underconsumption. This would indicate that satiation and income effects more than account for the purchase effect.

**Results.** Table 2 displays the proportion of subjects who demonstrate underconsumption for each purchase event. Note that the results provide substantial evidence of underconsumption. Entries in the table that are greater than 14 indicate that a significant proportion of subjects underconsume (i.e., the proportion differs significantly from zero, at p < .05—the null hypothesis suggested by the economic model). However, because the experimental procedure might encourage people to state that they would spend less after a purchase, we prefer to test the hypothesis that underconsumption will vary with typicality.

Consistent with our main hypothesis, the number of subjects who show an effect of underconsumption does increase with the typicality of the previous purchase. Note, for example that 69 percent of subjects underconsume entertainment after purchasing a $20 sports ticket (typicality = 5.88), while only 48 percent of subjects underconsume entertainment after purchasing $20 worth of party snacks (typicality = 2.88). Overall, the proportion of subjects who underconsume is highly correlated with typicality (r(25) = .912, p < .001).

This test provides strong evidence supporting the hypothesis that underconsumption will be greater for typical goods. The budgeting model explains this pattern by assuming that typical goods are more likely to be posted. However, there is a stronger and a weaker version of our hypothesis. The weak version of the hypothesis, the threshold model, argues that purchases are equally likely to be posted as long as they are at least somewhat typical of a category. Note in Table 1 that some items in each category are completely nontypical.

**TABLE 2**

<table>
<thead>
<tr>
<th>娱乐</th>
<th>食物</th>
<th>衣服</th>
</tr>
</thead>
<tbody>
<tr>
<td>体育票</td>
<td>69</td>
<td>17</td>
</tr>
<tr>
<td>旱冰鞋租赁</td>
<td>69</td>
<td>10</td>
</tr>
<tr>
<td>派对零食</td>
<td>48</td>
<td>55</td>
</tr>
<tr>
<td>比萨</td>
<td>36</td>
<td>59</td>
</tr>
<tr>
<td>外出用餐</td>
<td>48</td>
<td>52</td>
</tr>
<tr>
<td>葡萄酒</td>
<td>34</td>
<td>45</td>
</tr>
<tr>
<td>运动衫</td>
<td>21</td>
<td>3</td>
</tr>
<tr>
<td>手套</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>服装</td>
<td>52</td>
<td>7</td>
</tr>
</tbody>
</table>
TABLE 3
STUDY 1: MAGNITUDE OF UNDERCONSUMPTION (IN DOLLARS)

<table>
<thead>
<tr>
<th></th>
<th>Entertainment</th>
<th></th>
<th>Food</th>
<th></th>
<th>Clothes</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
<td>U &gt; 0</td>
<td>All</td>
<td>U &gt; 0</td>
<td>All</td>
<td>U &gt; 0</td>
</tr>
<tr>
<td>Sports ticket</td>
<td>5.94</td>
<td>10.66</td>
<td>.52</td>
<td>5.25</td>
<td>-1.51</td>
<td>6.75</td>
</tr>
<tr>
<td>In-line skate rental</td>
<td>6.01</td>
<td>9.76</td>
<td>.52</td>
<td>8.33</td>
<td>-.30</td>
<td>5.00</td>
</tr>
<tr>
<td>Party snacks</td>
<td>3.18</td>
<td>9.75</td>
<td>6.38</td>
<td>12.50</td>
<td>-.65</td>
<td>7.08</td>
</tr>
<tr>
<td>Pizza</td>
<td>.68</td>
<td>7.64</td>
<td>7.07</td>
<td>12.87</td>
<td>-.20</td>
<td>6.25</td>
</tr>
<tr>
<td>Dinner out</td>
<td>1.94</td>
<td>7.91</td>
<td>6.21</td>
<td>12.42</td>
<td>-.47</td>
<td>10.00</td>
</tr>
<tr>
<td>Wine</td>
<td>.34</td>
<td>7.15</td>
<td>5.86</td>
<td>13.37</td>
<td>-.23</td>
<td>1.75</td>
</tr>
<tr>
<td>Sweatshirt</td>
<td>-2.33</td>
<td>6.08</td>
<td>-.34</td>
<td>1.25</td>
<td>4.87</td>
<td>13.75</td>
</tr>
<tr>
<td>Gloves</td>
<td>-3.56</td>
<td>3.13</td>
<td>-.34</td>
<td>1.25</td>
<td>4.87</td>
<td>12.27</td>
</tr>
<tr>
<td>Costume</td>
<td>3.18</td>
<td>8.35</td>
<td>.17</td>
<td>10.00</td>
<td>4.87</td>
<td>14.17</td>
</tr>
</tbody>
</table>

Note.—The first column in each category gives the averages across all subjects; the second column gives the averages across only those subjects who showed positive underconsumption.

The threshold model predicts that the effect of typicality is driven by the contrast between these goods and all others. A continuous-effect model predicts that underconsumption increases with each increment in typicality, for example, that moving from moderate to high typicality produces an additional increment in underconsumption. The test above does not distinguish between these two models.

One simple way to test whether there is a continuous effect of typicality is to delete the items with typicalities of less than one. Here, the correlation remains significant ($r(12) = .63, p < .02$), providing evidence for a continuous effect of typicality. Alternatively, we can test for both threshold and continuous effects simultaneously: Here, we regress underconsumption on typicality and a nontypicality dummy variable (that equals one if typicality is less than one). This regression explains a large proportion of the variance (adjusted $R^2 = .72$) and produces a significant coefficient for the nontypicality dummy variable ($B = -.48, p < .05$) and marginally significant coefficient on typicality ($B = .40, p = .08$). By reducing the variance of the typicality coefficient, this analysis provides somewhat stronger evidence for a continuous effect of typicality. However, typicality seems to have a more continuous effect on the number of people demonstrating underconsumption than on the magnitude of underconsumption.

Figure 1 illustrates the effects in Table 3 by decomposing the purchase effect into satiation and income effects. The figure plots two regressions based on the data: satiation on typicality (with the constant effect of income added in) and purchase effect on typicality. It also shows the income effect, which by construction does not vary with typicality. Satiation increases with typicality, but the purchase effect increases faster, leading to an effect of mental budgeting that increases with typicality. Note that people report that they are less likely to consume additional items even after purchasing untypical items. However, for these items the purchase effect is almost completely due to an income effect.

Discussion

This study shows that when purchases are highly typical, underconsumption is more common—more people underconsume items in the target category after they imagine making a highly typical purchase. In addition, there is weaker evidence that when purchases are highly typical, underconsumption is of larger magnitude.
Economic consumer theory predicts that there will be no underconsumption and makes no prediction about the relationship between underconsumption and typicality. The economic predictions are correct when typicality is low (see the low typicality endpoint of Fig. 1), but the results diverge quickly from this prediction as typicality increases.

STUDY 2: UNDERCONSUMPTION OF A TYPICAL TARGET

One possible concern with study 1 is that it assumes the presence of mental accounts with labels like “entertainment,” “food,” and “clothing.” Although pilot work supported this assumption, we might want to assess the effects of mental budgeting without directly invoking the labels. A second possible concern with study 1 is that the task may have required an unnatural amount of precision. For example, people may not typically recalculate the amount remaining in an account after each purchase they make.

Mindful of these potential concerns, we eliminated the explicit use of category labels in study 2 and used a different dependent measure. In study 1, we asked people how much they would spend on entertainment after purchasing a $20 sports ticket. In study 2, we ask people whether they would purchase a $20 theater ticket after purchasing a $20 sports ticket.

On the basis of the results of the typicality ratings discussed in study 1, we selected a target purchase for each of the three categories that was rated highly typical of that category. Targets were Chinese take-out (food), a theater ticket (entertainment), and a sweatshirt with a college logo (clothing). Each target received one of the highest typicality ratings for its category: Chinese take-out received 5.4, theater ticket 6.2, and sweatshirt 5.0.

Because targets are highly typical of their categories, mental budgeters should always check how much money remains in the appropriate budget before purchasing the target. We predict that people will be less likely to purchase a target after a typical purchase—for example, people will be less likely to purchase a theater ticket after purchasing a sports ticket than after buying a party costume. Because the sports ticket is a more representative entertainment purchase, it is more likely to be posted and reduce the funds for the entertainment account. In turn, it is more likely to block the purchase of the theater ticket.

We also add a second manipulation to provide additional evidence of budgeting. The budgeting process
should be more likely to constrain behavior when people have recently made a large purchase. To examine this prediction we manipulate the magnitude of previous purchases in a between-subjects design. For some subjects, the previous purchase involves a $20 expense; for other subjects, the previous purchase involves a $50 expense. We predict more underconsumption in the $50 condition because a $50 expense is more likely than a $20 expense to exhaust the budget allotment for the expense category. Again, the economic model predicts no difference between the conditions once we control for income effects.

Thus, we have two main hypotheses in this study. First, the budget-setting process should promote greater underconsumption in the $50 condition than the $20 condition. Second, the expense-tracking process should promote greater underconsumption for more typical purchases.

Method

Materials. Because we needed a set of items that could plausibly cost either $50 or $20, we modified the set of events used in the previous study. We substituted the following four items for the stimuli in study 1 (see Table 1 for typicality ratings): a boat tour in the entertainment category, salmon in the food category, and blue jeans and a watch in the clothing category.

As in the previous study, for each expense category, subjects answered 18 questions about whether they would purchase the target item after purchasing another item that cost $50 (or $20) or after receiving that item as a gift. An additional four questions asked whether they would purchase the target item after an unanticipated $50 (or $20) expense. Counterbalancing was handled as in study 1. We also asked a series of dual-event questions to determine whether people respond the same when events are assessed as a conjunction. These dual-event questions read as follows: “The following two things happen to you early in the week: (1) You are given a ticket to a sporting event that normally costs $50. (2) You find a parking ticket on your car. The fine is $50.” The dual-event questions produced results that were substantively identical to the separate items, so we only report the analysis on separate items below.

Subjects. Subjects were 66 MBA students, split evenly between conditions, who were provided with pizza and beer in exchange for their participation. The experiment lasted about 30 minutes.

Procedure. Below is the introductory wording and one event for the entertainment category.

Imagine that it is the beginning of the week. Each question below asks you whether you would make a certain purchase later in the week after a certain event occurred earlier in the week.

EVENT: You go out to dinner with friends. You spend $50 ($20).
Would you buy a $25 theater ticket later in the week? Yes No

EVENT: You are given the dinner above.
Would you purchase a $25 theater ticket later in the week? Yes No

Analysis. Our hypothesis is the same as in study 1, and our analysis of underconsumption is conceptually equivalent. However, because people responded in yes/no form, the analysis is necessarily coarse. To show underconsumption while controlling for income and satiation effects, we look for a pattern of responses like the following.

EVENT: You spend $50 on a sports ticket.
Would you purchase a $25 theater ticket later in the week? No
EVENT: You are given the sports ticket above.
Would you purchase a $25 theater ticket later in the week? Yes

Would you purchase a $25 theater ticket later in the week? Yes

This hypothetical subject’s decision not to purchase the theater ticket after purchasing the $50 sports ticket is evidence of underconsumption because it cannot be explained by satiation or income effects. Satiation does not explain the decision, because the subject would still purchase the theater ticket even if given the sports ticket (thus, attending the sporting event does not make the play unappealing). Income constraints do not explain the decision, because the subject would still purchase the theater ticket even after an unexpected $50 expense.

This test is very stringent. If a person’s budget for entertainment is sufficiently large, no $50 expense will prevent the purchase of a desirable target item. Likewise, if the target purchase is sufficiently undesirable, it might never be purchased in any circumstance. To minimize this issue, the cover page allowed subjects to specify a substitute if they would never consider purchasing a particular target. Substitutions were relatively rare.

Results. Table 4 reports the results from this study. First, we test budget setting by examining whether people are more likely to underconsume in the high-cost than in the low-cost condition. A larger proportion do so ($26 = 2.17, p < .05 by paired t-test for each stimulus item).

We now examine whether underconsumption is affected by expense tracking. The proportion of subjects who underconsume the target is highly correlated with typicality for both high-cost ($r(25) = .80, p < .01) and low-cost conditions ($r(25) = .67, p < .01). The correlation is also positive for each category (low-cost condition: clothing $r(7) = .58, NS, food $r(7) = .74, p < .05,
### TABLE 4

STUDY 2: PERCENTAGE OF SUBJECTS SHOWING UNDERCONSUMPTION

<table>
<thead>
<tr>
<th>Entertainment</th>
<th>Food</th>
<th>Clothes</th>
</tr>
</thead>
<tbody>
<tr>
<td>$50</td>
<td>$20</td>
<td>$50</td>
</tr>
<tr>
<td>$20</td>
<td></td>
<td>$20</td>
</tr>
<tr>
<td>$50</td>
<td>$20</td>
<td>$50</td>
</tr>
<tr>
<td>Sports ticket</td>
<td>38</td>
<td>33</td>
</tr>
<tr>
<td>Boat tour</td>
<td>41</td>
<td>33</td>
</tr>
<tr>
<td>Party snacks</td>
<td>21</td>
<td>9</td>
</tr>
<tr>
<td>Salmon</td>
<td>21</td>
<td>9</td>
</tr>
<tr>
<td>Dinner out</td>
<td>32</td>
<td>21</td>
</tr>
<tr>
<td>Wine</td>
<td>15</td>
<td>3</td>
</tr>
<tr>
<td>Jeans</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Watch</td>
<td>15</td>
<td>9</td>
</tr>
<tr>
<td>Costume</td>
<td>38</td>
<td>18</td>
</tr>
</tbody>
</table>

and entertainment $r(7) = .75, p < .05$; high-cost condition: clothing $r(7) = .74, p < .05$, food $r(7) = .80, p < .01$, and entertainment $r(7) = .86, p < .01$).

As in the previous study, we can test whether there is any evidence of a continuous effect of typicality by deleting the items with typicalities of less than one. The correlation is significant and positive for the high-cost condition ($r(14) = .61, p < .05$) and marginally significant for the low-cost condition ($r(14) = .42, p = .10$). We can simultaneously test for both threshold and continuous effects by regressing underconsumption on typicality and a nontypicality dummy variable (that equals one if typicality is less than one). The regression for the high-cost condition (adjusted $R^2 = .62$) produces a significant coefficient for typicality ($B = .64, p < .02$) but not for the dummy variable ($B = -.19, p > .40$). Similar results are found in the low-cost condition (adjusted $R^2 = .41$), with a significant coefficient for typicality ($B = .54, p < .05$) but not for the dummy variable ($B = -.16, p > .50$). In contrast with the results of study 1, this analysis reveals a continuous effect of typicality but no threshold effect.

### Discussion

Despite the stringent nature of the analysis we perform, this study successfully replicates the major finding of study 1. Controlling for income and satiation effects, we find that more people show an effect of mental budgeting for highly typical expenses. The two studies thus provide converging results using different procedures and dependent measures.

#### STUDY 3: UNDERCONSUMPTION WITHIN A CATEGORY

In studies 1 and 2 we measured satiation by assessing how people change their consumption after receiving a gift. The budgeting model suggests a potential problem with this procedure. When people consume a gift, their marginal utility for consuming additional items in the category is likely to decrease and they should respond by spending less in that category. However, if budgets are sufficiently sticky (i.e., difficult to change), people may not alter their budgets even though they receive less satisfaction from consuming items in a category (i.e., they overconsume items in the category in which the gift was received). If so, the gift measure of satiation we use in studies 1 and 2 may underestimate true satiation.

Note that the pattern documented in studies 1 and 2 still provides evidence of budget setting and expense tracking. However, we do not know whether the pattern occurs because people underconsume after purchasing a typical item or because people overconsume after receiving that item as a gift. This potential flaw does not affect the validity of our tests against economic theory. Because economics does not recognize either overconsumption or underconsumption, studies 1 and 2 still demonstrate consumption patterns that cannot be explained by standard economic consumer theory. However, because mental budgeting distinguishes between underconsumption and overconsumption, we would like to separate the two effects. Therefore, in this experiment we use a scale measure of satiation that avoids the potential confound.

A second reason for this study is to revisit whether the magnitude of underconsumption increases continuously with prototypicality. Studies 1 and 2 show that typicality has a continuous effect on the number of people who underconsume, but they provide only limited evidence that it has a continuous effect on the magnitude of underconsumption. The test of this effect in study 1 reached only marginal significance. Study 3 includes a wider range of stimuli and limits them to a single category; therefore, it provides a more sensitive test of this effect.

### Method

#### Subjects

Subjects were 114 MBA students who completed the survey in exchange for a ticket for a lottery with a $100 prize.

#### Procedure

Subjects saw a list of 12 items that varied in price and in their typicality as members of the entertainment category, for example, a $60 ticket to the opera and a $27 hardcover novel. An equal number of typical and nontypical items were assigned to the $30 range or the $60 range.

There were three sections of the survey, each of which required responses on seven-point Likert scales. In the first section of the survey, subjects rated how much purchasing the item would reduce their spending on entertainment for the remainder of the week (1 = a little, 7 = a lot). In the second section of the survey, subjects rated the extent to which each item satisfied
(i.e., satiated) their need for entertainment (1 = does not satisfy my need for entertainment, 7 = very much satisfies my need for entertainment). In the final section, subjects rated the degree to which each item was typical of the entertainment category (1 = not typical, 7 = highly typical).

Results. To control for idiosyncratic use of the scales, we converted each subject's responses on the three scales into standardized z-scores. These standardized scores were used in a regression across all 1,368 observations (114 subjects × 12 items each). Reduction in spending was regressed on price, satisfaction, and typicality. The correlations between the dependent variables did not exceed .50. The regression equation is reported below (standard errors are in parentheses).

\[
\text{Spending reduction} = -1.23 + .028 \times \text{Price} + .073 \times \text{Satisfaction} + .250 \times \text{Typicality}
\]

Overall, the regression explained a reasonable amount of the variance \(R^2 = .25\). Furthermore, the coefficient for the typicality rating was both positive and highly significant \(t = 9.1, p < .001\); partial \(R^2 = .06\). In fact, the coefficient for typicality was significantly larger than the coefficient for satiation \(t = 3.7, p < .001\).

As mentioned above, one reason to do this study was to see if the magnitude of underconsumption was continuously related to typicality for purchases that are at least somewhat typical. Although the average typicality ratings were all above 2.51 on the seven-point typicality scale, we performed a more rigorous analysis to ensure that the analysis above was not skewed by the presence of low-typicality items. In the second analysis, we deleted for each subject any item the subject rated as a one or a two on the seven-point typicality scale. This analysis ensures that we only analyze items that each subject considers at least somewhat typical. The effect of typicality remains highly significant \(t = 6.5, p < .001\), partial \(R^2 = .04\).

Discussion

Again, this experiment demonstrates the phenomenon of underconsumption, without the possible confusion produced by the gift measure of satiation. This provides additional evidence of underconsumption with a very different experimental paradigm than the one used in studies 1 and 2. Furthermore, it shows that the magnitude of underconsumption increases continuously with typicality. The effects of typicality are particularly strong when compared with the effects of satiation. Typicality exerts more than three times the impact of satiation.

GENERAL DISCUSSION

In this article we have described the psychological process of mental budgeting and have explored some important implications of budgeting. Like previous literature in mental accounting, we highlight that labels affect people's behavior. However, we focus on a broader pattern of labeling that has been previously considered. We assume that people set budgets by labeling money as relevant for a certain class of goods and that they track expenses by labeling goods as relevant for a certain pool of money.

Although the current article concentrates on consumer decisions, similar kinds of labeling will play an important role in any theory of mental accounting. The way that people label their resources will be important anywhere that resources are relatively unidimensional, for example, in important domains like money, time, and effort. By dividing and labeling otherwise uniform resources, people simplify cognitive calculations and allow themselves to monitor their self-control efforts.

They decide how much time to spend on hobbies and work and how much mental energy to spend maintaining old relationships or acquiring new ones. The second kind of labeling, expenditure labeling, will be important in any domain in which people pursue multiple goals or in which activities have multiple dimensions. When asked to go to dinner with a visiting speaker, people may respond differently depending on whether they classify the time they spend at dinner as work or entertainment and whether they subtract the calories they consume from their minimum daily allowance or their allotment for special-occasion gluttony.

The current article provides evidence for both budget setting and expense tracking in the context of consumer decisions. Although underconsumption has previously been demonstrated in investment decisions (Heath 1995), the consumer decisions explored in the current studies exhibit a much richer context for demonstrating the phenomenon of expense tracking. We used the logic of the expense-tracking process to predict that typical expenses are the most likely to be underconsumed. Highly typical purchases are more likely to reduce the amount people are willing to spend on a class of expenses (studies 1 and 3) and to block the purchases of other typical items (study 2). Mental budgeting adds predictive power to the standard economic theory of consumer behavior—each of the three experiments controls for satiation and income effects. Thus, mental budgeting has an impact above and beyond the standard effects predicted by the economic analysis.

Our findings show that budgeting affects consumption decisions. However, there are still many open questions. For example, are budgets vague or specific? How often are budgets refreshed? Conceptually, the precision of the budgeting process can be pictured as a continuum; on one end is the bean-counting precision.
of “tin can accounting,” in which people know precisely what amount remains in each canister, and on the other end are informal methods in which people have only a vague notion about what they should spend. The effects in this article would not have been documented if the average level of precision were sufficiently low; however, individuals are likely to differ on this dimension.

Budgeting and Self-Control

The traditional literature on self-control has focused primarily on situations in which people are weak willed or undercontrolled (Ainslie and Haslam 1992; Elster 1979; Schelling 1992). Our studies of budgeting demonstrate that, in order to avoid weakness, people sometimes adopt self-control strategies that are too strong. In Freud’s classic description of intrapersonal conflict, an ego mediates between the internalized social norms of the superego and the base demands of the id. The traditional literature has focused on problems that arise when the id is unrestrained. The current phenomenon suggests problems that may arise when the superego is too controlling.

Other research on self-control has also noted that people may underconsume to maintain control: for instance, smokers often choose to quit cold turkey, even though they might prefer a few cigarettes to none, because they fear that one cigarette will lead to 20 (Thaler 1985). However, in contrast with traditional examples that have focused on tempting goods like cigarettes or alcohol that inherently provoke mixed feelings, we find that budgets cause people to underconsume unobjectionable goods like sports tickets and blue jeans.

When self-control is an issue, people should adopt some amount of inflexibility in their budgets. However, it is an open question whether people adopt an appropriate amount. The evidence in this article raises the question of whether people would benefit by allowing themselves to reallocate more freely.

Budgeting and Purchase Rationalizations

In addition to explaining features of consumer choice like over- and underconsumption, the budgeting model also has the potential to explain how consumers rationalize their purchases. Any system of rules can be circumvented, and mental budgets are no exception. However, by understanding the system of rules, we are in a better position to identify attempts at rule evasion. Because budgets are inflexible, people may be especially likely to justify their expenses through various ingenious methods of posting expenses. When a given expense could be assigned to multiple categories, people may have some leeway for hedonic posting—that is, posting items in a way that satisfies short-term interests and skirts the budget. For example, people may prefer to post dinner to the food account when the entertain-

ment budget is low, and they may accomplish this feat of hedonic posting by focusing on the nutritional value of their pad thai rather than the atmosphere of the restaurant in which it is consumed.

Although people might prefer to engage in this type of hedonic posting, they may find it difficult to generate alternative ways of looking at their desired purchases. Friends (and advertisers) often attempt to aid potential consumers by suggesting alternative ways to post an expense: “Don’t think of the new Oriental rug as a household expenditure, think of it as an investment,” or “Don’t think of the lobster as an expensive grocery item, think of it as an inexpensive form of entertainment.”

Budgeting is worth understanding because descriptions of consumers over the past 50 years indicate that it is a pervasive part of consumer behavior. It is also worth understanding because it alters consumer choice in interesting ways. However, even when people successfully evade their budgets—as in the case of the hedonically posted pad thai—we might still want to understand mental budgeting because it defines the very rules that guide the grammar of rationalization.

[Received February 1994. Revised January 1996. Brian Sternthal served as editor and Joseph W. Alba served as associate editor for this article.]

REFERENCES


Bakke, E. Wright (1940), The Unemployed Worker, New Haven, CT: Yale University Press, 142–143.


O’Curry, Sue (1996), “Income Source Effects,” unpublished working paper, Department of Marketing, DePaul University, Chicago, IL 60604.