Market Lett (2010) 21:17–35
DOI 10.1007/s11002-009-9087-0

# I'll have the ice cream soon and the vegetables later: A study of online grocery purchases and order lead time

**Katherine L. Milkman · Todd Rogers · Max H. Bazerman**

Published online: 25 August 2009


**Abstract** How do decisions made for tomorrow or 2 days in the future differ from decisions made for several days in the future? We use data from an online grocer to address this question. In general, we find that as the delay between order completion and delivery increases, grocery customers spend less, order a higher percentage of "should" items (e.g., vegetables), and order a lower percentage of "want" items (e.g., ice cream), controlling for customer fixed effects. These field results replicate previous laboratory findings and are consistent with theories suggesting that people's *should* selves exert more influence over their choices the further in the future outcomes will be experienced. However, orders placed for delivery tomorrow versus 2 days in the future do not show this *want*/*should* pattern, and we discuss a potential explanation.

**Keywords** Lead time · Intertemporal choice · Want/should · E-commerce · Intrapersonal conflict

As internet shopping becomes increasingly ubiquitous, a question of growing importance is whether and how demand for different types of products varies with

K. L. Milkman (✉)
The Wharton School, 3730 Walnut Street, 561 Jon M. Huntsman Hall, Philadelphia, PA 19104, USA
e-mail: kmilkman@wharton.upenn.edu

T. Rogers
Analyst Institute, Washington, DC, USA

M. H. Bazerman
Harvard Business School, Baker Library 453, Soldiers Field Road, Boston, MA 02163, USA

order lead time. In traditional retailing situations, people consider their purchasing options and gain access to their selections in immediate succession. However, e-commerce retailers, which surpassed $100 billion in sales for the first time in 2006 (comScore Press Release 2007), require customers to make choices about products one or more days before those customers will receive their selections. In addition, many e-commerce companies offer multiple shipping options, so customers can order products for delivery with different lead times. In this paper, we investigate how a difference in the time delay separating an online order's completion and its delivery relates to the purchasing decisions made by consumers in the domain of grocery shopping.

Research on intertemporal choice suggests that when the time separating a purchasing decision from the receipt of a purchase is exogenously varied, this delay can significantly alter people's selections. Numerous laboratory studies and a burgeoning number of field studies have shown that people behave more impulsively—spending more and more often choosing items they hedonically *want* to select over items they cognitively believe they *should* select—when outcomes are more immediate (for a review, see Khan et al. 2005 or Milkman et al. 2008). In this study, we extend research from the laboratory documenting systematic differences between choices made for the near future versus the more distant future, presenting field data that replicate this pattern. Past field research on intertemporal choice has focused on examining the differences between choices made for now versus later (Milkman et al. 2009; Ashraf et al. 2006; Malmendier and Della Vigna 2006; Oster and Scott Morton 2005). In this paper, we use data provided by an online grocer to examine how the delay between an order's completion and its delivery relates to a given consumer's overall spending (an impulsive, *want* behavior), purchases of *should* groceries (e.g., healthy foods like vegetables), and purchases of *want* groceries (e.g., unhealthy foods like ice cream).

Our customer-level online grocery shopping data set allows us to control for customer fixed effects in all of our analyses, and we are thus able to examine differences in the choices the same people make about purchases they will receive in the near future, beginning as early as tomorrow, versus the more distant future (Wooldridge 2002). Consistent with the predictions of dual selves theories of impulsive individual decision making (Shefrin and Thaler 1988; Bazerman et al. 1998; Read 2001) and with economic models of consumers as decreasingly impatient (see, for example, Loewenstein and Prelec 1992), we find that on average, the same consumers spend less, order a higher percentage of *should* goods, and order a lower percentage of *want* goods the further in advance of delivery they place a grocery order. It is important to note that another potential explanation for our findings besides decreasing impatience is that the circumstances that lead people to plan further in advance are correlated with the circumstances that lead them to spend less and order relatively healthier foods. Our field data do not afford us the opportunity to disentangle which of these explanations is responsible for our results. However, our findings, which have important implications for online retailers, are consistent with past experimental and theoretical research on intertemporal choice and decreasing impatience, which motivated this study.

## 1 Relevant past research

Past research on intrapersonal conflict, also known as the multiple selves phenomenon (Schelling 1984), has documented a tension between the behaviors people feel they *should* exhibit given their long-term interests (e.g., saving more, going to the gym, starting a diet) and the behaviors they find themselves hedonically *want*ing to exhibit and often choosing to exhibit due to their short-run rewards (e.g., spending more, watching television instead of going to the gym, and eating cake with lunch). Bazerman et al. (1998) describe this tension as stemming from two selves—a *want* self and a *should* self—which have competing preferences. Shefrin and Thaler (1988) also propose that people live in a state of internal conflict between a "doer" self that parallels the *want* self described by Bazerman et al. (1998) and a "planner" self, which parallels the *should* self of Bazerman et al.

The multiple selves framework predicts that in situations where outcomes are more immediate, decision makers will be more likely to make *want* choices that have primarily short-run benefits, such as spending money freely (rather than saving it for the future) and indulging in more unhealthy *want* foods and fewer healthy *should* foods. The closer the reward, the more likely it is that an individual's visceral desires will overwhelm his or her cooler cognitive systems (Loewenstein 1996). Economists have modeled this phenomenon by assuming individuals have a steep short-run discount rate and a relatively flat long-run discount rate, which leads them to overvalue present utility relative to future utility and thus to favor *want* options (such as spending and eating tasty but unhealthy foods) over *should* options (such as saving and eating healthy but less tantalizing foods) at a higher rate the sooner their choices will take effect (see, for example, Ainslie 1975; Loewenstein and Prelec 1992). These theories predict that consumption decisions made for the nearer future will involve heightened overall spending (a *want* behavior) as well as increased spending on goods that would be preferred by the *want* self, or "*want* goods", while spending decisions made for the more distant future will result in less spending overall (a *should* behavior) as well as increased spending on goods preferred by the *should* self, or "*should* goods". These are the predictions we seek to test in our field study.

Several laboratory studies have tested the hypothesis that people behave more impulsively when the outcomes of their decisions will be realized in the near future rather than the more distant future.[1] Benzion et al. (1989) conducted such a study, which employed a survey design that allowed the authors to estimate participants' 6-month and 1-, 2-, and 4-year discount rates over different hypothetical sums of money ($40, $200, $1,000, and $5,000). The authors found that participants' inferred discount rates decreased as the time they had to wait for a reward increased, meaning participants exhibited decreasing impatience over monetary gains. In another laboratory study, Zauberman and Lynch (2005) found that, on average, respondents reported being more likely to donate time to charity (a *should* behavior) in 2 weeks than tomorrow, a finding consistent with the idea that people are decreasingly impatient. Finally, Rogers and Bazerman (2008) demonstrated in

[1] Note that a considerable body of work has demonstrated that people behave more impulsively when making choices for *now* rather than for later (see Milkman et al. 2008 for a review).

another series of laboratory studies that people are more likely to support *should* policies when those policies will be implemented in the distant future rather than in the near future.

The current paper provides a field examination of how longer delays between the time of a choice and the time of its realization relate to the same people's preferences for *should* versus *want* options in the near future versus the more distant future using a large field data set from an online grocer. Our results replicate the findings of previous laboratory experiments in a field setting. The rest of this paper is outlined as follows: We begin by describing the details of the data set we obtained from a large, online grocer and the methods we employ to classify the groceries in our data set on a spectrum from extreme *should* items to extreme *want* items. Next, we present the results of a series of panel regressions including customer fixed effects, which examine whether patterns in our field data are consistent with the predictions made by the multiple selves framework and models of decreasing impatience described above. Finally, we conclude with a discussion of our findings and their implications.

## 2 Data

### 2.1 Overview of data

The online grocer we collaborated with on this study operates in North America and serves urban customers. Its customers place orders by browsing the products available on the company's website and adding items to an electronic grocery cart. Customers have the option to schedule a delivery during an available delivery slot for as early as tomorrow or for further in the future. During the period studied, the grocer charged a delivery fee for online orders. In addition, customers were required to spend a minimum dollar amount on each order. To preserve business confidentiality, company-specific information has been withheld from this document.

We obtained a novel panel data set from the aforementioned online grocery retailer containing information about the orders placed by all of the company's customers between January 1, 2005 and December 31, 2005. The online grocery company provided a record of each item in each order placed during the 12-month period in question, as well as the price each customer paid for each item, the date of each order, the date of each order's delivery, and the customer who placed each order. If a customer modified his or her order, we were told how many times order modifications were made, as well as the first and last dates when the customer modified his or her shopping basket. We operationalize order lead time in this paper as the time separating a customer's last visit to the grocer's website to change an order and the date when the customer's groceries were delivered. Note that online grocery customers could modify their selections after placing an initial order up until a cutoff time that allowed the online grocer time to shop, transport, and deliver the customer's order. All customer accounts in our data set are labeled by anonymous, unique ID numbers. Our online grocery collaborator also provided us with category information about each item available for purchase through its website.

We restrict our analysis in this paper to customers who ordered groceries for delivery between 1 and 5 days in advance sometime between January 1, 2005 and December 31, 2005. We exclude all orders involving the redemption of a coupon because discount coupons have been shown to affect online grocery spending as well as the distribution of goods in a customer's shopping basket (Milkman and Beshears 2009).

In total, between January 1, 2005 and December 31, 2005, tens of thousands of customers ordered groceries for delivery between 1 and 5 days in advance without redeeming a discount coupon.[2] We eliminate each customer's first order of the year,[3] spending outliers (top 1%), and outliers in the number of visits made to the grocer's website during an order (top 1%). This leaves us with over a million grocery orders in 2005 (customers in our analyses ordered an average of five to ten times). The average dollar size of an order in this sample is $154.71 and the average grocery order consists of 58 items. For additional summary statistics, see Table 1.

The majority of customers in our data set completed their grocery orders 1 day in advance of delivery. However, many customers completed orders between 2 and 5 days in advance of their scheduled delivery date (see Table 2). There is almost no seasonality in the rate at which customers' order lead times vary with the exception of slight volatility in January and February and one unusual week in February.[4] In all of our analyses of this data, we include week fixed effects, and we also replicate each result without January and February data to ensure that these two somewhat unusual months, when snowstorms may have affected both order lead times and the types of items customers purchased, are not driving our findings.

## 2.2 Classifying groceries

To classify the items in our grocery data set based on their position along the spectrum from *should* to *want*, we conducted an online survey (employing a similar method to that used by Milkman et al. 2009). One hundred fifty-four people were paid to participate in this survey and answered questions about approximately 30 food categories from our database of groceries. Groceries in our data set have all been classified by our online grocer into one of 117 categories (e.g., frozen vegetables, cream, cookies, etc.). We randomly partitioned the grocery categories into four groups of approximately 30 categories each, and every survey participant was randomly assigned to answer questions about one of these four groups. Respondents were only asked about 30 grocery categories to reduce the likelihood of boredom and mechanical responses.

Our survey respondents were anonymous volunteers from all over the USA who signed up over the Internet to participate in online paid polls administered by the

[2] Details about the number of unique customers, total number of grocery orders, and average number of orders per customer in our data set are not provided in order to preserve the anonymity of our data provider.

[3] This allows us to control for how much time has elapsed since a customer's last order in our analyses.

[4] More details on the seasonality of order lead time are available upon request.

**Table 1** Summary Statistics

| | Mean | Standard deviation |
|---|---|---|
| Spending | \$154.71 | \$65.83 |
| Number of groceries | 58.38 | 25.95 |
| Number of web visits for order | 3.27 | 2.59 |
| Days between first and last web visits for order | 1.37 | 0.73 |
| Days since last delivery | 21.84 | 29.48 |

research lab of a large university. After being provided with concept definitions,[5] respondents were asked to rate grocery categories along a 1–7 Likert scale anchored on *not a "want" grocery category* and *a strong "want" grocery category* and a 1–7 Likert scale anchored on *not a "should" grocery category* and *a strong "should" grocery category*. Respondents saw the name of a grocery category and the names of its associated subcategories when completing our survey (e.g., candy and gum: candy chocolate, candy nonchocolate, gum and mints), and the order in which they were asked to rate grocery categories along *should* and *want* scales was randomized. No significant order effects were present in our survey data.[6]

We gave participants an incentive to provide accurate ratings of grocery categories by paying them for performance. For each grocery category, a survey participant classified within one point of the average rating across respondents, her "accuracy score" was increased by one. The 20% of participants who received the highest accuracy scores were paid a bonus of \$5 on top of their \$5 participation fee.

To generate a single variable quantifying where on the spectrum from an extreme *should* to an extreme *want* each grocery category falls, we subtract each grocery category's *want* score from its *should* score. We average our raters' *should minus want* scores to create an overall *should minus want* index for each grocery category (again following Milkman et al. 2009). If our survey ratings contain a meaningful signal, we should find that the *should minus want* scores assigned by different survey participants to the same grocery category are more tightly clustered than the *should minus want* scores assigned by different survey participants to different grocery categories. We run a one-way analysis of variance to compare ratings variation between grocery categories to ratings variation within grocery categories (Shrout and Fleiss 1979). An intraclass correlation of 0.34 and an estimated reliability of a grocery category mean of 0.95 confirm that our survey averages are reliable—survey ratings vary significantly more between grocery categories than within grocery categories. For a catalog of the grocery categories in our sample and an ordered list of their associated average *should minus want* ratings, see "Appendix".

[5] Lengthy concept definitions were provided to participants and they were also quizzed on their understanding of these concepts. Full materials are available upon request. The final summary of a "want" grocery read: "The 'want' score is intended to reflect the extent to which someone's decision to consume this type of grocery would be indulgent and pleasure-based." The final summary of a "should" grocery read: "The 'should' score ought to reflect the extent to which someone's choice to consume the grocery would be made for virtuous, self-improving reasons, regardless of other potential factors."

[6] Wilks' lambdas from multivariate analysis of variances run to examine potential ordering effects were all insignificant at the 5% level.

**Table 2** Delivery lead time summary statistics

| | |
|---|---|
| % of orders completed 1 day in advance of delivery | 74.40 |
| % of orders completed 2 days in advance of delivery | 18.17 |
| % of orders completed 3 days in advance of delivery | 4.76 |
| % of orders completed 4 days in advance of delivery | 1.85 |
| % of orders completed 5 days in advance of delivery | 0.82 |

Summary statistics describing the percentage of orders completed varying numbers of days in advance of delivery, excluding each customer's first order of 2005

In order to validate our *should minus want* metric, we examined the correlation between the average *should minus want score* for each of the grocery categories rated including foods and the average healthfulness rating (on a scale from −5 = very unhealthy to +5 = very healthy) given to the two most popular items in each of these grocery categories by a panel of 13 nutrition experts (see Martin et al. 2009 for more on these expert ratings). A correlation of 0.49 ($p$ value<0.0001) indicates that our *should minus want score* is closely related to experts' perceptions of a food's healthfulness, increasing our confidence in this measure.

In addition to developing *should minus want* scores for each of the grocery categories in our data set, we created two other means of classifying *should* and *want* items so we would have multiple, imperfectly correlated measures of *should* and *want* groceries to use in our analyses. Heilman et al. (2002) developed a method for classifying groceries as "treats", or hedonically attractive, *want* items, and we adopt the authors' classifications as a second, independent method for identifying extreme *want* groceries. These authors created a list of treats based on the items that 57 grocery shoppers said they would buy if they "wanted to treat themselves or their families to something special" (Heilman et al. 2002, p. 246). Of the groceries that were listed, the 50% that were listed most often by these survey respondents were labeled "treats", as were goods found in the checkout aisle of a grocery store. We match grocery categories in our database to the groceries in the Heilman et al. "treats" list, as shown in Table 3. To develop another means of classifying extreme *should* groceries, we look to the nutrition literature for a class of items that people are advised to consume in order to improve long-term health outcomes. A class of groceries fitting this description includes fresh fruits, vegetables, seafood, and meats (Willet 1994; Van Duyn and Pivonka 2000; Drewnowski and Barratt-Fornell 2004). The grocery category designations used by the online grocery company allow us to classify a subset of foods as "fresh foods", or fresh fruits, vegetables, seafood, and meats (see Table 3), which constitute a set of foods people *should* consume in greater quantities.

We employ multiple outcome variables in our analyses of whether people purchase a lower proportion of impulse goods and higher proportion of healthy goods when order lead times are longer. To capture the relative dominance of *should* goods compared to *want* goods in a given customer's basket, we calculate the average *should minus want* score of all of the groceries in that basket. Two of our other outcome variables capture the proportion of extreme *want* groceries purchased: the percentage of an order's dollar value composed of groceries receiving one of the ten lowest *should minus want* scores and the percentage of an

Table 3 Classification of groceries

| Fresh foods | Treats | |
|---|---|---|
| | In Heilman et al. (2002) | Corresponding groceries in our data |
| Produce—vegetables | Ice cream | Ice cream (category) |
| Meat—fresh | Bakery goods | Bakery—fresh (category) |
| Seafood—fresh | Steak | All other fresh meat (subcategory) |
| Produce—fruits | | Meat (subcategory) |
| Deli—fresh | Wine | Wine/wine coolers (subcategory) |
| Bakery—fresh | Candy | Candy and gum (category) |
| | Cheese | Cheese (category) |
| | Cookies | Cookies (category) |
| | Magazine | Mags/newspapers/books (subcategory) |
| | Chocolate | Candy and gum (category) |
| | | Hot chocolate mix (subcategory) |
| | Flowers | Floral (category) |
| | Cake | Cake mixes (subcategory) |
| | | Cakes (fresh; subcategory) |
| | Seafood | Seafood—fresh (category) |
| | | Seafood—frozen (category) |
| | Baby toy | NA |
| | Chips | Potato chips (subcategory) |
| | | Tortilla chips (subcategory) |
| | | Corn chips/snacks (subcategory) |
| | Cosmetics | Cosmetics (category) |
| | Movie rental | Music/movies (subcategory) |
| | Pie | Pies (fresh; subcategory) |
| | Gum/Mints | Candy and gum (category) |

order's dollar value composed of treats.[7] Our final two outcome variables capture the proportion of extreme *should* groceries purchased: the percentage of an order's dollar value composed of fresh foods and the percentage of an order's dollar value composed of groceries receiving one of the ten highest *should minus want* scores.[8] Table 4 presents the correlations between these different outcome variables as well as summary statistics about dollar spending per order on each category of groceries.

## 3 Results

We begin by evaluating the relationship between the time separating an order's completion from its delivery and customer spending. Table 5 presents the results of

[7] Because the choice to look at ten categories rather than some other number is somewhat arbitrary, we replicate all results examining the top five categories of *should* and *want* groceries as a robustness check.
[8] Ibid.

**Table 4** Correlations between outcome measures and summaries of spending on each category of groceries

| | Basket's average SMW score | % of order's dollar value composed of | | | |
|---|---|---|---|---|---|
| | | Fresh foods | 10 highest SMW scores | 10 lowest SMW scores | Treats |
| Fresh foods | 0.3722[a] | | | | |
| 10 highest SMW scores | 0.5524[a] | 0.1865[a] | | | |
| 10 lowest SMW scores | −0.4551[a] | −0.1860[a] | −0.1510[a] | | |
| Treats | −0.3098[a] | 0.0006 | −0.1572[a] | 0.5485[a] | |
| Average spending/(score) on category | −0.0646 | \$39.00 | \$21.84 | \$7.21 | \$14.91 |
| Standard deviation of spending/(score) on category | 0.6678 | \$29.20 | \$16.36 | \$10.95 | \$14.28 |

[a] Denotes significance at the 1% level

ordinary least squares (OLS) regressions estimating the relationship between the amount a given customer spends on groceries and how far in advance of delivery she completes her grocery order. In these regressions and in subsequent regressions, the explanatory variables include the number of days in advance of delivery a customer completed her order, the number of times the customer visited the online grocer's website in the course of placing an order, the number of days between the first and last visit the customer made to the grocer's website in the course of placing an order, the number of days since the customer last received a grocery delivery, a dummy indicating if 60 or more days have passed since the customer's last grocery order, the number of orders placed by the customer year to date, dummies for the day of the week when the order was placed, dummies for the day of the week when the order was delivered, dummies for each week in 2005, and customer fixed effects. Standard errors are clustered at the customer level.

By including customer fixed effects, we are able to identify off of within-customer variation in our analyses of the effects of lead time on consumer choice (Wooldridge 2002). In other words, the results of our regressions provide insights into how customers' orders differ when the delay between order and delivery varies, controlling for the average decisions made by a given customer.

Consistent with the hypothesis that people spend money more freely when they make decisions for the more immediate future, we find that holding all else constant, the dollar size of a grocery order decreases by approximately 2.0% for each additional day that separates a customer's last visit to the online grocer's website and the date when her groceries are delivered (see Table 5, regression 2). Regression 1 in Table 5 indicates that this effect corresponds to approximately \$2.70 less in spending on groceries per day of additional order lead time. It is important to note that although this result is consistent with our first hypothesis, which is based on the theory that people's *should* selves exert more influence over their decisions the further in the future their decisions will take effect, there are many plausible

**Table 5** The effects of order lead time on spending and purchases of *want* and *should* groceries

| | Spending | Log(1+ spending) | Basket's average SMW score | % of order's dollar value composed of | | | |
|---|---|---|---|---|---|---|---|
| | | | | Fresh foods | 10 highest SMW scores | 10 lowest SMW scores | Treats |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| One day between order completion and delivery | | | 0.0070[a] (0.0026) | 0.0028[a] (0.0005) | 0.0009[b] (0.0004) | −0.0010[a] (0.0002) | −0.0005 (0.0004) |
| Days between order completion and delivery | −2.6994[a] (0.0860) | −0.0195[a] (0.0005) | 0.0053[a] (0.0019) | 0.0024[a] (0.0004) | 0.0012[a] (0.0003) | −0.0006[a] (0.0002) | −0.0004[c] (0.0003) |
| Number of web visits for order | 3.1705[a] (0.0316) | 0.0209[a] (0.0002) | −0.0023[a] (0.0003) | −0.0015[a] (0.0001) | −0.0002[a] (0.0000) | 0.0000 (0.0000) | 0.0002[a] (0.0000) |
| Days between first and last web visits for order | −0.1359[a] (0.0050) | −0.0009[a] (0.0000) | 0.0000 (0.0000) | 0.0001[a] (0.0000) | 0.0000 (0.0000) | 0.0000[b] (0.0000) | 0.0000 (0.0000) |
| Days since last delivery | 0.2517[a] (0.0049) | 0.0016[a] (0.0000) | 0.0004[a] (0.0000) | −0.0002[a] (0.0000) | −0.0000[a] (0.0000) | −0.0000[a] (0.0000) | −0.0001[a] (0.0000) |
| 60 or more days since last order | −11.8873[a] (0.4033) | −0.0762[a] (0.0025) | −0.0046 (0.0030) | 0.0055[a] (0.0009) | 0.0021[a] (0.0006) | 0.0006 (0.0004) | 0.0019[a] (0.0005) |
| Days since first order with grocer×$10^3$ | 70.7254[a] (0.0077) | 0.0005[a] (0.0000) | −0.0416 (0.0898) | −0.0276 (0.0187) | −0.0151 (0.0153) | −0.0142 (0.0106) | −0.0299[b] (0.0129) |
| Orders year to date | −0.0186 (0.0201) | −0.0002 (0.0001) | 0.0007[a] (0.0002) | −0.0001[c] (0.0000) | 0.0001[b] (0.0000) | 0.0000 (0.0000) | −0.0001[b] (0.0000) |
| Day of the week order placed fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Day of the week order delivered fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Week of the year fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Customer fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1 million+ | 1 million+ | 1 million+ | 1 million+ | 1 million+ | 1 million+ | 1 million+ |
| Customers | 100,000+ | 100,000+ | 100,000+ | 100,000+ | 100,000+ | 100,000+ | 100,000+ |
| $R^2$ | 0.6740 | 0.6812 | 0.6423 | 0.6605 | 0.5472 | 0.5294 | 0.4713 |

Columns 1 and 2 report OLS coefficients from regressions of customer spending on a continuous variable indicating how far in advance of delivery an order was completed controlling for the other variables listed. Columns 3 through 7 report OLS coefficients from regressions of customer spending on categories of groceries on a dummy indicating whether an order was completed 1 day in advance of delivery and a continuous variable indicating how far in advance of delivery an order was completed, controlling for the other variables listed. Robust standard errors clustered at the customer level are in parentheses

[a] Denotes significance at the 1% level

[b] Denotes significance at the 5% level

[c] Denotes significance at the 10% level

alternative explanations for the observed decrease in spending associated with orders placed for the more distant future. For example, this result may be driven by the fact that people know more about exactly what their needs will be when ordering groceries for the more immediate future and thus purchase more groceries the sooner their groceries will be delivered.

In the following analyses, we investigate the impact of delivery lead time on the percentage of a customer's spending that is concentrated on different types of goods and the average *should minus want* score of goods in a customer's basket. By looking at the percentage composition and average *should minus want* score of groceries in customers' baskets, we control for the overall decrease in spending across categories of goods that is associated with orders placed for the more distant future.

In the regressions that follow, rather than simply including a linear effect for the number of days in advance of delivery a customer places an order, we also include a dummy variable indicating whether an order was completed 1 day in advance of delivery. We include this dummy variable because exploratory data analyses revealed that this regression specification was most appropriate given the patterns in our data. In order to determine the appropriate specification for our regressions, we began by running each analysis with dummy variables indicating the number of days in advance of delivery an order had been completed. These regressions demonstrated a consistent pattern—a linear trend was apparent in the *should* and *want* contents of orders completed between 2 and 5 days in advance of delivery, as predicted, but orders completed 1 day in advance of delivery did not follow this monotonic pattern.

In Table 5, we present the results of a series of OLS regressions estimating the relationship between the percentage of a customer's grocery spending concentrated on different types of *should* and *want* groceries, the average *should minus want* score of items in a customer's basket, and how many days in advance of delivery a customer completes her order. The results presented in Table 5 indicate that for orders completed between 2 and 5 days in advance of delivery, the further in advance of delivery a customer completes an order, the relatively more *should* goods and fewer *want* goods she will purchase, consistent with the hypothesis that people are more likely to favor *should* options over *want* options the further in advance of consumption they make decisions. However, contrary to our prediction, orders completed 1 day in advance of delivery contain about the same percentage of *should* and *want* goods as orders completed 2 days in advance of delivery. This apparent nonlinearity in customers' patterns of choice is persistent across different measures of *should* and *want* goods, although the nonlinearity lies within one standard error of the linear trend detected across our analyses. We will discuss this unexpected pattern in our data in more detail and offer a potential explanation for it in Section 4 of this paper. The remainder of this section, however, will focus on our findings with respect to the differences between orders completed between 2 and 5 days in advance of delivery.

Regression 3 in Table 5 demonstrates the effect of an increase in the time between an order's completion and its delivery on the average *should minus want* score of a grocery basket. It shows that for orders placed between 2 and 5 days in advance of delivery, for each additional day in advance of delivery an order is completed, the average *should minus want* score of an entire grocery basket increases by 0.0053 (or approximately 0.008 standard deviations). Regressions 4 and 5 in Table 5 provide information about the change in the percentage of an order composed of *should* items that is associated with a change in how far in advance of delivery the order is completed. These regressions show that for orders placed between 2 and 5 days in advance of delivery, for each additional day in advance of delivery an order is completed, the percent of the order composed of fresh foods increases by 0.24 (or an average of \$0.37), and the percent composed of groceries with the ten highest *should minus want* scores increases by 0.12 (or an average of \$0.19). Regressions 6 and 7 in Table 5 focus on the change in the percentage of an order composed of *want* items that is associated with a change in how far in advance of delivery an order is completed. They show that for orders placed between 2 and 5 days in advance of delivery, for each additional day in advance of delivery an order is completed, the percent of the order composed of groceries with the ten lowest *should minus want* scores decreases by 0.06 (or an average of \$0.09). In addition, the percent of an order composed of *treats* decreases by 0.04 (or an average of \$0.07). Each of these results is consistent with our general prediction that people will have a stronger preference for *should* goods and a weaker preference for *want* goods the further in advance of consumption they make their grocery selections. Each of these regressions also contains a nonlinearity of the type described above, which we did not predict.[9]

To ensure that our results are not driven by any unusual events in January and February that may have caused more orders to be completed further in advance of delivery than usual (see Section 2.1), we rerun all of the above analyses without including orders placed in these months. The results of our regressions remain meaningfully and statistically unchanged when orders placed in these months are eliminated. We also rerun all of the above analyses excluding orders made by customers who did not place orders with each of the five possible lead times examined in this paper, and the magnitude of the effects we observe do not differ meaningfully with this restricted sample, although their statistical significance is weakened somewhat. These additional analyses are all available upon request.

## 4 Discussion

The results presented above demonstrate systematic differences in the choices a given customer makes when she completes a grocery order between 2 and 5 days in advance of delivery, which are consistent with the predictions of theories of multiple selves conflict and decreasing impatience. First, we find that customers engage in less spending (an impulsive, *want* behavior) the further in advance of delivery they complete an online grocery order. Second, we find that for orders placed between 2 and 5 days in advance of delivery, for each additional day in advance of delivery an order is completed, the percent of an order composed of *want* groceries decreases and the percent composed of *should* groceries increases.

In addition to providing evidence that is consistent with our predictions about the impact of order lead time on online purchasing decisions, the regression analyses

[9] Regressions examining the percent spending on the five grocery categories receiving the highest and lowest *should minus want* scores reveal the same patterns and are available upon request. These results also hold if grocery categories containing alcohol and/or cigarettes are removed.

discussed above also expose one unexpected but persistent feature of our data. The results of our analyses indicate that orders completed 1 day in advance of delivery include a slightly lower proportion of *want* goods and a slightly higher proportion of *should* goods than orders placed 2 days in advance of delivery (although the difference is never significant). This pattern in our data is not consistent with our ex ante predictions or with previously discussed trends in the composition of orders completed between 2 and 5 days in advance of delivery.

While theories of multiple selves conflict and decreasing impatience did not lead us to predict this pattern of results ex ante, we explored potential ex post explanations for our findings informed by the literature on construal level theory (CLT) and discussions with online grocery shoppers. CLT suggests that when making choices for the more distant future, people tend to focus on more abstract and less concrete features of their options than when making choices for the more immediate future (Trope and Liberman 2003). We developed two hypotheses. First, we hypothesized that people are more likely to order groceries for specific, planned meals (as opposed to general pantry stocking) when ordering for tomorrow than for the more distant future. This is because ordering foods for a planned meal involves ordering for a concrete, specific purpose, while ordering for general pantry stocking involves ordering for an abstract purpose. Second, we hypothesized that groceries ordered for specific, planned meals are more likely to be *should* items and less likely to be *want* items than groceries ordered for general pantry stocking. This is because planned meals typically include fresh ingredients, which are likely to be *should* items, while pantry stockers typically involve unhealthy packaged and processed foods (Willet 1994; Van Duyn and Pivonka 2000; Drewnowski and Barratt-Fornell 2004).

In order to test these ex post hypotheses to account for the unexpected pattern we detected in consumers' online grocery purchases, we ran a survey with 230 participants. Survey respondents were randomly assigned to a condition in which they were instructed to imagine ordering groceries for tomorrow, 2 days in the future, or 5 days in the future and to create a shopping list. Consistent with our first hypothesis, participants in the "tomorrow" condition created hypothetical lists that contained significantly more groceries that were self-reported to be intended for specific meals ($t(227)=-3.49$, $p=0.001$) and fewer groceries intended for pantry stocking ($t(227)=-1.92$, $p=0.056$) than participants in the other two conditions. Consistent with our second hypothesis, respondents also reported that, in general, groceries they order for specific meals are significantly more likely to be *should* foods (binomial test of proportions, $N=168$, $p=0.053$) and less likely to be *want* foods (binomial test of proportions, $N=182$, $p<0.001$) than groceries they order for general pantry stocking. Before responding to these questions, participants were provided with detailed descriptions of *want* and *should* following Milkman et al. (2009). More details on this survey are available upon request.

Although the survey results described above do not provide the only plausible explanation for the unexpected pattern in our field data, they provide data consistent with one potential explanation. Together, our field data and survey data suggest that increasing the lead time between a grocery order's completion and its delivery may give rise to two separate psychological effects. First, we present evidence from our field data set that is consistent with past research showing that people generally behave more impulsively the sooner their decisions will take effect. However, the field data we examine suggest that this pattern is not apparent when the types of groceries in orders placed 1 and 2 days in advance of delivery are compared with one another. Our survey data offers a potential explanation for this: People order groceries for delivery tomorrow with more specific purposes in mind than when they order groceries for delivery in the more distant future, and this leads them to order more *should* and fewer *want* groceries for tomorrow than for the more distant future. We propose that these two effects may combine to produce the purchasing patterns we observe.

It is important to note that while the findings presented in this paper are generally consistent with our predictions, with theories of decreasing impatience, and with past laboratory studies, we cannot draw causal conclusions from our analyses. Our findings may result from multiple selves conflict, a correlation between the situational factors that lead people to order further in advance and lead to less impulsive behavior, or some other phenomenon altogether. In spite of this, combined with consistent evidence from past laboratory studies in which the time separating a choice from its realization was exogenously varied (dispelling causality concerns), we believe the findings we present in this paper may have a number of potentially important implications. Testing each of these potential implications presents a promising opportunity for future research.

First, our findings may have implications for online and catalog retailers that offer a range of goods for sale and also offer different delivery options. Such companies might be able to improve their demand forecasting by taking into account the fact that their customers may spend more and order a higher percentage of *want* goods and a lower percentage of *should* goods for delivery in the near future than in the more distant future. They might also be able to increase their customers' spending by persuading them to place orders for the more immediate future.

Our finding that people select healthier foods for themselves the further in the future their groceries will be delivered also has potential policy implications. Motivated by past research on intertemporal choice and intrapersonal conflict, Rogers and Bazerman (2008) conducted a series of studies demonstrating that people are more likely to select *should* policies (e.g., increased taxes on fossil fuels, increased charitable spending, etc.) when they will be implemented in the distant future rather than the near future. Offering people *should* choices that will take effect in the future is a strategy that they termed "future lock-in". Our finding that people are more likely to buy a higher proportion of *should* items and a lower proportion of *want* items the further in advance of delivery they order groceries raises the possibility that "future lock-in" could be more effective the further in advance of implementation people are asked to vote on *should* policies.

Finally, combining the specific domain in which our research was conducted with past work on future lock-in, our findings may have implications for nutrition policy. Our findings raise the question of whether encouraging people to order their groceries many days in advance of consumption could influence the healthfulness of the foods people consume.[10] Perhaps asking students in schools to select their lunches up to a week in advance could increase the healthfulness of the foods they

[10] Although it is possible that people only buy a healthier bundle of groceries when they order further in the future and do not actually eat healthier groceries, it seems likely that purchases are highly correlated with consumption.

elect to eat. If these predictions could be confirmed, an attractive aspect of implementing policies that encourage advance planning is that they would preserve the decision maker's choice set and autonomy by changing only the context in which decisions are made. By changing the decision context, policy-makers might be able to increase the likelihood that people would make "better" choices without infringing upon their freedom (Sunstein and Thaler 2003).

**Acknowledgments** The authors thank John Beshears, George Loewenstein, Kathleen McGinn, Nava Ashraf, David Parkes, Carey Morewedge, Bill Simpson, Sarah Woolverton, and a very helpful set of reviewers for their assistance with this project. We are also grateful to the employees of the online grocer who generously shared their time, data and ideas with us.

## Appendix

**Table 6** Average should minus want scores for grocery categories in our data set

| Grocery category | Average *should minus want* score |
|---|---|
| Cookies | −5.098 |
| Wine/wine coolers | −4.976 |
| Ice cream | −4.976 |
| Candy and gum | −4.420 |
| Cigars and tobacco | −4.300 |
| Mixers/bar needs | −4.140 |
| Frozen pizza | −4.073 |
| Cigarettes | −4.000 |
| Spirits | −4.000 |
| Prepared cocktails | −3.963 |
| Cosmetics | −3.951 |
| Floral | −3.927 |
| Baking mixes | −3.659 |
| Frozen snacks/appetizers | −3.600 |
| Beverages—soda | −3.600 |
| Cream | −3.439 |
| Frozen potatoes/onion rings | −3.360 |
| Toys/cards | −3.185 |
| Bakery—commercial | −3.049 |
| Party favors/balloons | −3.000 |
| Bakery—fresh | −2.951 |
| Baking supplies/ingredients | −2.902 |
| Spreads | −2.854 |
| Beverages—creamers | −2.640 |
| Dips (refrigerated) | −2.481 |
| Syrup flavoring (no-breakfast) | −2.407 |
| Beverages—coffee | −2.320 |

**Table 6** (continued)

| Grocery category | Average *should minus want* score |
|---|---|
| Prepared food | −2.260 |
| Beverages—juice/drinks | −2.244 |
| Fruit snacks | −2.220 |
| Gravy/marinade/sauces | −2.140 |
| Sauces (refrigerated) | −2.049 |
| Frozen dinners/entrees | −1.926 |
| Sour cream | −1.880 |
| Seasonal | −1.880 |
| Breakfast (frozen) | −1.778 |
| Salad dressing/toppings | −1.732 |
| Beverages—isotonics | −1.560 |
| Deli—packaged | −1.520 |
| Butter/margarine/spreads | −1.512 |
| Salty snacks | −1.455 |
| Beer and cider | −1.303 |
| Dough (refrigerated) | −1.259 |
| Bread/dough (frozen) | −1.222 |
| All other general merchandise | −1.200 |
| Ice cream toppings/cones | −1.182 |
| Frozen dessert/pie/pastries | −1.182 |
| Gelatin/pudding snacks (refrigerated) | −1.152 |
| Nonalcoholic beer/wine | −1.148 |
| Olive/pickle/peppers (refrigerated) | −1.000 |
| Entertainment | −0.909 |
| Spices/extracts | −0.900 |
| Beverages—hot chocolate | −0.848 |
| Gelatin/pudding | −0.788 |
| Crackers | −0.727 |
| Pasta (refrigerated) | −0.704 |
| Soft goods | −0.606 |
| beverages (frozen) | −0.576 |
| Fruits (frozen) | −0.545 |
| Breakfast | −0.481 |
| Dried bread | −0.481 |
| Condiments | −0.455 |
| Ice | −0.444 |
| Beverages (refrigerated) | −0.364 |
| Diet care | −0.280 |
| Fruits | −0.242 |
| Film/batteries | −0.212 |
| Beverages—tea | −0.185 |
| Air care | −0.182 |

**Table 6** (continued)

| Grocery category | Average *should minus want* score |
|---|---|
| Seafood—frozen | −0.148 |
| Soap | −0.061 |
| Cheese | 0.024 |
| Septic system/softener salt | 0.030 |
| Baby health | 0.061 |
| Deli—fresh | 0.061 |
| Automotive | 0.122 |
| Meat—frozen | 0.140 |
| Pesticides/bug repellents | 0.240 |
| Housewares | 0.364 |
| Meat/seafood | 0.364 |
| Pasta/grains | 0.488 |
| Medications | 0.515 |
| Office/school supplies | 0.545 |
| Skin care | 0.556 |
| Baby food | 0.576 |
| Oil/vinegar/cooking wine | 0.593 |
| Beverages—water | 0.606 |
| Soup | 0.704 |
| All other dairy | 0.732 |
| Bags/wraps/disposable containers | 0.758 |
| Pet care | 0.780 |
| Hair care | 0.815 |
| Produce—vegetables | 0.939 |
| Meat—fresh | 0.940 |
| Yogurt | 0.980 |
| Seafood—fresh | 1.000 |
| Family planning | 1.200 |
| Pet care—cat food | 1.300 |
| Incontinence | 1.370 |
| Shaving needs | 1.407 |
| Paper | 1.740 |
| Dish care | 1.880 |
| Pet care—dog food | 1.976 |
| Deodorants/antiperspirant | 2.037 |
| Eggs/egg substitutes | 2.146 |
| Eye/ear/foot care | 2.268 |
| Beverages—soy/rice | 2.296 |
| Laundry care | 2.512 |
| Household cleaners | 2.556 |
| Milk | 2.593 |
| Feminine care | 2.700 |

**Table 6** (continued)

| Grocery category | Average *should minus want* score |
|---|---|
| Vegetables | 2.704 |
| Produce—fruits | 2.732 |
| Vegetables (frozen) | 2.829 |
| Vitamins | 2.852 |
| First aid | 2.900 |
| Oral hygiene | 3.390 |

## References


Ainslie, G. (1975). Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychological Bulletin, 82*, 463–496.

Ashraf, N., Karlan, D., & Yin, W. (2006). Tying Odysseus to the mast: Evidence from a commitment savings product in the Philippines. *Quarterly Journal of Economics, 121*(2), 635–672.

Bazerman, M. H., Tenbrunsel, A. E., & Wade-Benzoni, K. (1998). Negotiating with yourself and losing: Making decisions with competing internal preferences. *Academy of Management Review, 23*(2), 225–241.

Benzion, U., Rapopport, A., & Yagil, Y. (1989). Discount rates inferred from decisions: An experimental study. *Management Science, 35*(3), 270–284.

comScore Press Release (2007) comScore Networks reports total non-travel e-commerce spending reaches $102 billion in 2006; up 24 percent versus 2005. comScore Press Release. January 3, 2007. Accessed April 30, 2008: http://www.comscore.com/press/release.asp?press=1166.

Drewnowski, A., & Barratt-Fornell, A. (2004). Do healthier diets cost more? *Nutrition Today, 39*(4), 161–168.

Heilman, C. M., Nakamoto, K., & Rao, A. G. (2002). Pleasant surprises: Consumer response to unexpected in-store coupons. *Journal of Marketing Research, 39*(2), 242–252.

Khan, U., Dhar, R., & Wertenbroch, K. (2005). A behavioral decision theory perspective on hedonic and utilitarian choice. In S. Ratneshwar & D. G. Mick (Eds.), *Inside consumption: Frontiers of research on consumer motives, goals, and desires* (pp. 144–165). London: Routledge.

Loewenstein, G. F. (1996). Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes, 65*(3), 272–292.

Loewenstein, G. F., & Prelec, D. (1992). Anomalies in intertemporal choice: Evidence and an interpretation. *Quarterly Journal of Economics, 107*, 573–597.

Malmendier, U., & Della Vigna, S. (2006). Paying not to go to the gym. *American Economic Review, 96*(3), 694–719.

Martin, J. M., Beshears, J., Milkman, K. L., Bazerman, M. H., & Sutherland, L. (2009). Modeling expert opinions on food healthiness: A nutrition metric. *Journal of the American Dietetic Association, 109*(6), 1088–1091.

Milkman, K. L., & Beshears, J. (2009). Mental accounting and small windfalls: Evidence from an online grocer. *Journal of Economic Behavior and Organization, 71*(2), 384–394.

Milkman, K. L., Rogers, T., & Bazerman, M. H. (2008). Harnessing our inner angels and demons: What we have learned about want/should conflicts and how that knowledge can help us reduce short-sighted decision making. *Perspectives on Psychological Science, 3*, 324–338.

Milkman, K. L., Rogers, T., & Bazerman, M. H. (2009). Highbrow films gather dust: Time-inconsistent preferences and online DVD rentals. *Management Science, 55*(6), 1047–1059.

Oster, S., & Scott Morton, F. M. (2005). Behavioral biases meet the market: The case of magazine subscription prices. *BE Journals Economic Analysis and Policy—Advances, 5*(1), 1323.

Read, D. (2001). Intrapersonal dilemmas. *Human Relations, 54*(8), 1093–1117.

Rogers, T., & Bazerman, M. H. (2008). Future lock-in: Future implementation increases selection of should choices. *Organizational Behavior and Human Decision Processes, 106*(1), 1–20.

Schelling, T. C. (1984). *Choice and consequence: Perspectives of an errant economist*. Cambridge: Harvard University Press.

Shefrin, H., & Thaler, R. H. (1988). The behavioral life-cycle hypothesis. *Economic Inquiry, 26*(4), 609–643.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420–428.

Sunstein, C. R., & Thaler, R. H. (2003). Libertarian paternalism is not an oxymoron. *University of Chicago Law Review, 70*, 1159–1199.

Trope, Y., & Liberman, N. (2003). Temporal construal. *Psychological Review, 110*, 403–421.

Van Duyn, M. A. S., & Pivonka, E. (2000). Overview of the health benefits of fruit and vegetable consumption for the dietetics profession: Selected literature. *Journal of the American Dietetic Association, 100*(12), 1511–1521.

Willet, W. C. (1994). Diet and health: What should we eat? *Science, 264*(5158), 532–537.

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. Cambridge: MIT.

Zauberman, G., & Lynch, J. G. (2005). Resource slack and propensity to discount delayed investments of time versus money. *Journal of Experimental Psychology: General, 134*(1), 23–37.