Projection Bias in Catalog Orders

By Michael Conlin, Ted O'Donoghue, and Timothy J. Vogelsang*

Evidence suggests that people understand qualitatively how tastes change over time, but underestimate the magnitudes. This evidence is limited, however, to laboratory evidence or surveys of reported happiness. We test for such projection bias *in field data. Using data on catalog orders of cold-weather items, we find evidence of projection bias over the weather—specifically, people's decisions are overinfluenced by the current weather. Our estimates suggest that if the order-date temperature declines by 30°F, the return probability increases by 3.95 percent. We also estimate a structural model to measure the magnitude of the bias.* (*JEL* D12, L81)

People's tastes change over time in systematic ways. A person's taste for cake depends on whether she is hungry or satiated; her taste for coffee depends on whether she has developed that taste; and her (dis)taste for a chronic medical condition depends on whether she has adapted to that condition. The standard economic approach to changing tastes assumes that people accurately predict changes in their tastes—either they know exactly how their tastes will change when taste changes are deterministic, or they have a correct model of how their tastes will change when taste changes are uncertain. Evidence from psychology, however, suggests that people exhibit a systematic bias in such predictions: while people understand qualitatively the direction in which their tastes change—e.g., they understand that eating dinner diminishes one's appetite for dessert—people systematically underestimate the magnitudes of these changes. George Loewenstein, O'Donoghue, and Matthew Rabin (2003) label this tendency *projection bias*.[1]

Although there is a great deal of evidence of projection bias, for the most part this evidence examines either (a) how people's predicted quality of life associated with chronic medical conditions or important life events (e.g., getting or being denied tenure) compare to actual self-reports of people who have experienced those outcomes; or (b) small-scale choices in the laboratory. We are not familiar with any field-data evidence that projection bias influences purchases of goods and services. In this paper, we conduct precisely such a test by analyzing catalog orders for weather-related clothing items and sports equipment. We indeed find evidence of projection bias with respect to the weather, and in particular that people are overinfluenced by the weather at the time they make decisions. In addition, we estimate a structural model to measure the magnitude of the bias, and find that people's predictions for future tastes are roughly one-third to one-half the way between actual future tastes and current tastes.

Our evidence is important for several reasons. First, catalog sales are a large and growing segment of the US economy, with estimated total revenue of over $125 billion in 2006 and growth

* Conlin: Department of Economics, Michigan State University, 110 Marshall-Adams Hall, East Lansing, MI 48824 (e-mail: conlinmi@msu.edu); O'Donoghue: Department of Economics, Cornell University, 404 Uris Hall, Ithaca, NY 14853 (e-mail: edo1@cornell.edu); Vogelsang: Department of Economics, Michigan State University, 110 Marshall-Adams Hall, East Lansing, MI 48824 (e-mail: tjv@msu.edu). For useful comments, we thank two anonymous referees, and also George Loewenstein, Matthew Rabin, and seminar participants at Cornell University, University of California, Berkeley, Harvard University, Michigan State University, Indiana University, Wharton School, Carnegie Mellon University, and SUNY Albany. We are especially grateful to Steve Coate for numerous discussions during the early stages of this project. We thank Chris Cotton for research assistance, and for financial support, O'Donoghue thanks the National Science Foundation (award SES-0214043).

[1] Loewenstein, O'Donoghue, and Rabin (2003) review the evidence from psychology, build a simple model of the bias, and use this model to explore (theoretically) the implications of projection bias for economic environments. For a more detailed discussion of the psychological evidence, see Loewenstein and David Schkade (1999).

estimates of 30 percent over the next two years (*The Directory of Mail Order Catalogs*, 2006 edition). The consumer catalog sector accounts for over half of these total sales. Second, and more importantly, projection bias has implications for many important economic decisions besides catalog orders. It can lead to bad consumer decisions, because it makes people overly prone to buy goods when their tastes for those goods are high—shopping on an empty stomach—and not prone to buy when their tastes for those goods are low—shopping on a full stomach. Moreover, because projection bias leads people to underappreciate their ability to adapt to major life events, it causes people to overemphasize the impact of major economic decisions, such as which job to take, where to live, or which house to buy. Projection bias is also relevant for consumption of addictive goods: people might too often become addicted to products such as cigarettes, illicit drugs, and alcohol because (a) they underappreciate the negative consequences of being an addict; and (b) they underappreciate how hard it will be to quit once addicted. Hence, it is important to measure whether and to what extent projection bias influences real-world economic decisions.[2]

In Section I, we develop a model of catalog orders and returns. We assume that a person's order and return decisions are based on a combination of (a) the expected instrumental value of the item given her local weather; (b) a known individual-specific taste (how much she needs/likes items of that general type); (c) an unknown individual-specific taste that she learns only after she receives and inspects the item; (d) the price of the item; and (e) the cost of returning the item in the event she chooses to do so. For a fully rational person, the weather on the day she makes order and return decisions should be essentially irrelevant. If the person has projection bias, however, her decisions are influenced by the current weather. Specifically, if the weather on the day she decides whether to order an item moves in a direction that would make the item more valuable if used on that day, then she is more prone to order the item, and as a result her likelihood of returning the item (conditional on ordering) increases. Similarly, if the weather on the day she decides whether to keep an item moves in a direction that would make the item more valuable if used on that day, then she is more prone to keep the item, and as a result her likelihood of returning the item (conditional on ordering) decreases.

To test this pair of predictions, we obtained data from a large outdoor-apparel company. The company provided detailed information on over 12 million orders of weather-related items, including the zip code of the buyer, the date of the order, and whether the item was returned. We merged this information with daily weather information for each zip code in the United States. We describe the data in more detail in Section II.

In Section III, we present reduced-form tests of our two main predictions. Specifically, we investigate how the likelihood of returning an item (conditional on ordering) depends on the weather on the order date and the weather in the days after the item is received (which we take to be the day that the return-versus-keep decision is made). For cold-weather items—items that are more valuable the colder is the temperature—our two main predictions imply that the likelihood of returning the item should be declining in the order-date temperature and increasing in the return-date temperature. We find strong support for the order-date prediction: a decline in the order-date temperature of 30°F is associated with an increase in the return rate of 3.95 percent. We find more limited support for the return-date prediction. The likely source for not finding stronger support for the return-date prediction is that we have limited information about when the return decision is actually made. In particular, we observe only when items are received, which we use as our proxy for when the return decision is made. But actual returns are spread out over the weeks following the receiving date. We also test for projection bias over the effects of snowfall and the effects

[2] This paper is part of a recent literature that attempts to use economic field data to test models from behavioral economics—for instance, the empirical tests of loss aversion in Colin Camerer et al. (1997), David Genesove and Christopher Mayer (2001), and Ernst Fehr and Lorenz Götte (2007), and the empirical tests of hyperbolic discounting in George-Marios Angeletos et al. (2001), Dan Ariely and Klaus Wertenbroch (2002), Haiyan Shui and Lawrence M. Ausubel (2004), David Laibson, Andrea Repetto, and Jeremy Tobacman (2005), Stefano DellaVigna and M. Daniele Paserman (2005), and DellaVigna and Ulrike Malmendier (2006).

of rainfall, and we find no evidence of order-date or return-date snowfall or rainfall influencing the likelihood of return. Even so, because the average snowfall and average rainfall have very little effect on behavior, these results might merely reflect that snowfall and rainfall do not significantly affect the utility from the items in our dataset.

Also in Section III, we attempt to rule out a variety of alternative explanations for these empirical results. These alternatives fall roughly into two categories. First, the current weather might contain information (direct or indirect) about the future value of the item—e.g., the current temperature might help forecast future temperatures and thus contain direct information about the future usefulness; people might be learning about the local weather and use the current temperature as a signal; or people might be learning about their existing clothing, and the current temperature affects the extent of such learning. Second, and perhaps more important, the current weather might alter the types of people who order. An ideal approach would incorporate not only the decision whether to order, but also the decision when to order. Because it is not entirely clear how to model the latter decision, our analysis assumes that the order date is exogenous. Even so, we speculate on some ways that the weather might influence when people order, and attempt to rule these out as alternative explanations for our empirical results.

While our reduced-form analysis supports the existence of projection bias in catalog orders, it does not identify the magnitude of the bias or the importance of the bias. To do so, we impose additional (functional-form and distributional) assumptions on our basic model from Section I and estimate the structural parameters of the model. In Section IV, we provide an overview of this estimation and present the results—a more detailed description of the estimation appears in Appendix B. We conclude in Section V by discussing some limitations of our analysis and some broader implications.

Our empirical analysis suffers from an important limitation. Our goal is to test for projection bias in the sense that people are biased by today's utility from an item when predicting the future utility from that item. Given our data, however, we are unable to distinguish such projection bias from people instead mispredicting the weather—that is, people being biased by today's weather when predicting future weather. This issue is particularly important for the broader relevance of our analysis. If people merely mispredict the weather, then our evidence is relevant only for weather-related decisions. But if, as we believe, people project their current utility onto their predictions for future utility, then our evidence is relevant for a wide range of important economic decisions (as we discuss above). We address this issue in more detail, including some (rather limited) ancillary evidence, at the end of Section I.

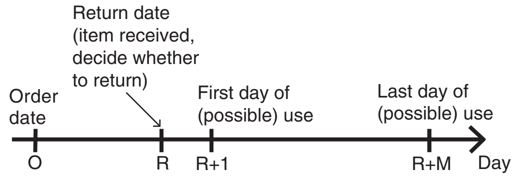

FIGURE 1. TIME LINE FOR DECISIONS

I. A Model of Catalog Orders and Returns

In this section, we develop a model of catalog orders and returns, and we use this model to derive testable implications. The basic environment is described by the time line in Figure 1.

On some exogenous date O—the "order date"—a person must decide whether to order an item of clothing. If the person chooses not to order the item, then her utility is zero (a normalization). If she chooses to order the item, then she receives the item after a (processing and shipping) delay.[3]

If the person chooses to order, then on some exogenous date R she receives the item (the "return date"), inspects the item (tries it on for fit, checks the quality, and so forth), and decides

[3] The assumption that the order date is exogenous is restrictive in that it ignores how the order date might be influenced by financial concerns such as looking for the lowest price, and how it might depend on the weather. We are not overly worried about the former, because it would seem not to provide any systematic bias with regard to our weather results, and because such effects are partially captured when we control for the purchase price. We are more concerned about the latter because it could give rise to alternative explanations for our empirical results; we address some such potential alternatives in Section III.

whether to keep it or to return it. If she decides to return the item, she must incur a return cost c, which reflects both the opportunity cost of her time and any financial shipping cost. If she decides to keep the item, she must pay the price p and she may begin using the item on date $R + 1$.[4] Our theoretical analysis in this section assumes we know the return date; however, the biggest data limitation in our empirical analysis is our inability to identify exactly when the return decision is being made—all we see is whether an item is returned and when it is restocked at the company. We shall return to this issue in Section III.

For simplicity, we assume that the item will be usable for exactly M days, so that the last date of (possible) use is day $R + M$, where M is exogenous to the person and independent of when the order is placed. Also for simplicity, we assume that the price p is paid on the return date. This assumption is not important, because exactly when the price is paid should have very small effects on a person's decision (e.g., if the person is able to delay a \$100 payment by 90 days without paying interest, and even if she can earn 10 percent interest on this money, she will earn only \$2.50 of interest). Since we cannot observe payment dates, we make this assumption as a reasonable approximation.

A. *Behavior by Fully Rational People*

We first analyze fully rational people, by which we mean people who do not exhibit projection bias. A person's decision depends on how much utility she expects to obtain from the item. We assume the utility from the item on day d is

$$v(\mu(\mathbf{x}), \gamma, \varepsilon, \omega_d) = [\mu(\mathbf{x}) + \gamma + \varepsilon]u(\omega_d).$$

The person's utility from the item consists of two components, an instrumental utility for the item, scaled by an individual-specific "slope." The term $u(\omega_d)$ represents the instrumental utility for the item as a function of the weather on day d, which we denote by ω_d. The instrumental utility reflects the item's value in terms of protection against the weather—e.g., keeping one warm in the case of a winter coat. We interpret $u(\omega_d)$ as the marginal utility from owning the item: when she doesn't use the item, $u(\omega_d) = 0$; when she uses the item, $u(\omega_d) > 0$ reflects the marginal utility she experiences relative to wearing her next-best clothing item.

Conditional on the weather, the instrumental utility for the item is the same for everyone (although different people will experience different weather patterns). Even conditional on the weather, however, the utility from the item can differ across individuals due to the individual-specific "slope" $[\mu(\mathbf{x}) + \gamma + \varepsilon]$. The term $\mu(\mathbf{x})$ captures the portion of the individual-specific preference for the item that depends on observable variables, which we denote by the vector $\mathbf{x}$. The variables γ and ε represent unobservable individual-specific preferences for the item. The variable γ is known to the person on the order date, and captures things such as whether the item is the person's style, how much the person needs the item given her existing wardrobe, and whether she tends to be easy or difficult to fit. The variable ε is unknown to the person on the order date but is discovered on the return date, and captures things such as whether the item fits and whether the person likes the attributes of the item that could not be discerned from the catalog or Web site. The population distribution of γ is given by $G(\gamma)$ with a mean of zero, and the population distribution of ε is given by $F(\varepsilon)$ with a mean of zero. We assume that γ and ε are independent, and that F and G are both continuous, differentiable, and strictly increasing on the real line.[5]

While $v(\mu(\mathbf{x}), \gamma, \varepsilon, \omega_d)$ represents the actual utility from the item experienced on date d as a function of the actual weather on date d, there is uncertainty about what the future weather will be. We assume, however, that people have correct expectations about their local weather. Let $H_d(\omega)$ be the distribution of the local weather on date d, and let $E_{H_d}[a(\omega)]$ denote the expectation of function $a(\omega)$ over the distribution H_d. From a prior perspective in which she knows γ and ε,

[4] Our formal analysis assumes $p > c$; for any item with $p \leq c$, the person would never pay to return the item (since she could just throw it away), and would order the item taking this into account. In our empirical analysis, we drop the few observations for which the price is under \$4.

[5] Given the presence of an intercept term in $\mu(\mathbf{x})$, the assumptions that γ and ε have mean zero are not restrictive.

the person's expected utility from the item on future date d is

$$E_{H_d}[v(\mu(\mathbf{x}),\gamma,\varepsilon,\omega_d)]=[\mu(\mathbf{x})+\gamma+\varepsilon]E_{H_d}[u(\omega_d)].$$

Consider the person's decision whether to return the item on day R. At this time, she knows $\mu(\mathbf{x})$, γ, and ε, and forms expectations about the weather. She can use the item for exactly M days, and we assume that all people have a common daily discount factor δ. Hence, if she decides to keep the item, her (gross) expected discounted utility (from the perspective of date R) is

$$U_R \equiv \sum_{d=R+1}^{R+M} \delta^{d-R} E_{H_d}[v(\mu(\mathbf{x}),\gamma,\varepsilon,\omega_d)]$$

$$= [\mu(\mathbf{x}) + \gamma + \varepsilon]\Psi_R,$$

where $\Psi_R \equiv \sum_{d=R+1}^{R+M}\delta^{d-R}E_{H_d}[u(\omega_d)]$. The term Ψ_R represents the person's expected discounted instrumental utility for the item. While our formal analysis assumes that the person formulates Ψ_R in this specific way, our interpretation is that people have a relatively accurate assessment of the instrumental value of the item for their locale. Indeed, our comparative statics below describe the implications of changing Ψ_R, and not the implications of changing the components of Ψ_R. (For our structural estimation, however, we are forced to take this formulation more literally.)

On the return date, the person compares paying the price p and then experiencing this expected utility to incurring cost c to return the item. Assuming risk neutrality, she will keep the item if

$$U_R - p \geq -c,$$

which we can rewrite as

$$\gamma + \varepsilon \geq [(p - c)/\Psi_R] - \mu(\mathbf{x}) \equiv \bar{\Lambda}.$$

Next, consider the person's decision whether to order the item on date O. At this time, she knows $\mu(\mathbf{x})$ and γ, and forms expectations about the weather and about ε. Her expectations about the weather are exactly as above. Moreover, a rational person has correct beliefs about the distribution of ε and correctly predicts her behavior on the return date. Hence, given γ, the person

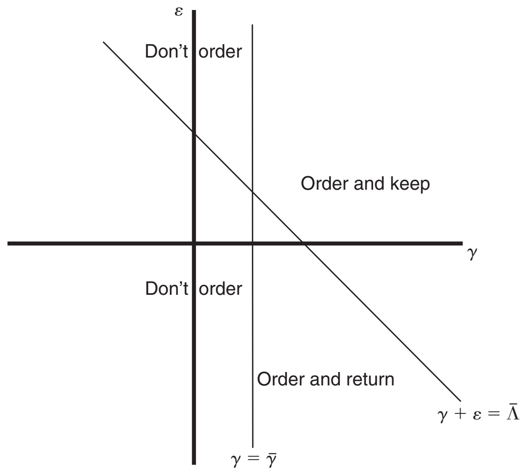

FIGURE 2. WHETHER TO ORDER AND RETURN AS A FUNCTION OF γ AND ε

believes that, if she orders, she will end up incurring cost c to return the item when $\varepsilon < \bar{\Lambda} - \gamma$, and she will end up paying price p to keep the item when $\varepsilon \geq \bar{\Lambda} - \gamma$, in which case her net expected utility will be $[\mu(\mathbf{x}) + \gamma + E(\varepsilon|\varepsilon \geq \bar{\Lambda} - \gamma)]\Psi_R - p$. The person will order when

$$\Pr[\varepsilon \geq \bar{\Lambda} - \gamma][(\mu(\mathbf{x}) + \gamma + E(\varepsilon|\varepsilon \geq \bar{\Lambda} - \gamma))\Psi_R - p]$$

$$+ \Pr[\varepsilon < \bar{\Lambda} - \gamma][-c] \geq 0.$$

It is straightforward to derive—see Appendix A for the details—that the person will follow a cutoff rule wherein she orders the item when

$$\gamma \geq \bar{\Lambda} - \bar{\varepsilon}(c/\Psi_R) \equiv \bar{\gamma},$$

where $\bar{\varepsilon}(a)$ is the $\bar{\varepsilon}$ such that $\Pr(\varepsilon \geq \bar{\varepsilon})[\mathrm{E}(\varepsilon|\varepsilon \geq \bar{\varepsilon}) - \bar{\varepsilon}] = a$.

Hence, a person's behavior can be described by two cutoffs, the order-date cutoff $\bar{\gamma}$ and the return-date cutoff $\bar{\Lambda}$. Figure 2 depicts an individual's decision of whether to order an item and whether to return an item, and how these decisions depend on ε and γ.

B. *Projection Bias*

We adopt the model of simple projection bias from Loewenstein, O'Donoghue, and Rabin (2003). In abstract terms, suppose a person's true tastes are given by $w(c,s)$, where c is the person's consumption and s is the person's "state," which parameterizes her tastes. Suppose further that the person is trying to predict her future utility—specifically, suppose she is currently in state s' and is attempting to predict what her future utility would be from consuming c in state s. Let $\tilde{w}(c,s|s')$ denote her prediction. For a fully rational person, predicted utility equals true utility, so $\tilde{w}(c,s|s') = w(c,s)$. But for a person who exhibits *simple projection bias*,

$$\tilde{w}(c,s|s') = (1-\alpha)w(c,s) + \alpha w(c,s')$$

for some $\alpha \in (0,1)$.

With this formulation, $\alpha = 0$ implies no projection bias, while any $\alpha \in (0,1)$ implies projection bias—that is, implies the person understands qualitatively the direction in which her tastes change, but underestimates the magnitudes of those changes. The bigger is α, the stronger is the bias.[6]

When a person with projection bias faces an intertemporal choice, she makes her choice in the same way that a fully rational person would, except that she uses her biased predictions of future utility in place of her true future utility (as we shall illustrate below within our model). Note, however, that because a person's tastes might change over time in ways she does not predict, she may exhibit *dynamic inconsistency*—she may plan to behave a certain way in the future, but later, in the absence of new information, revise this plan. In the present context, she may plan to return the item according to one criterion, but later, when the current weather has changed, actually return the item according to another criterion. We assume, as in Loewenstein, O'Donoghue, and Rabin (2003), that a person is completely unaware of such effects.[7]

C. *Behavior of People with Projection Bias*

Our hypothesis is that people exhibit projection bias with regard to the weather—when a person predicts the future value of a clothing item, she overestimates its value for days where it is less valuable than today, and underestimates its value for days where it is more valuable than today. To introduce such effects into our model, we assume that, whereas the true utility of the item on future date d is $v(\mu(\mathbf{x}),\gamma,\varepsilon,\omega_d)$, if the current weather is ω_t, the person perceives the utility on future date d to be

$$\begin{aligned}&\tilde{v}(\mu(\mathbf{x}),\gamma,\varepsilon,\omega_d|\omega_t)\\ &\qquad \equiv (1-\alpha)v(\mu(\mathbf{x}),\gamma,\varepsilon,\omega_d) + \alpha v(\mu(\mathbf{x}),\gamma,\varepsilon,\omega_t)\\ &\qquad = [\mu(\mathbf{x}) + \gamma + \varepsilon][(1-\alpha)u(\omega_d) + \alpha u(\omega_t)].\end{aligned}$$

Again assuming that people have correct expectations about their local weather, the person's perception on date t of her expected utility from the item on future date d is

$$\begin{aligned}&E_{H_d}[\tilde{v}(\mu(\mathbf{x}),\gamma,\varepsilon,\omega_d|\omega_t)]\\ &\qquad = [\mu(\mathbf{x}) + \gamma + \varepsilon]\\ &\qquad\qquad \times \left[(1-\alpha)E_{H_d}[u(\omega_d)] + \alpha u(\omega_t)\right].\end{aligned}$$

Next, consider a person's perception on the return date R of her (gross) expected discounted utility of the item. On date R, she knows $\mu(\mathbf{x})$, γ, and ε, and therefore her perception is

$$\begin{aligned}\tilde{U}_R &\equiv \sum_{d=R+1}^{R+M} \delta^{d-R} E_{H_d}[\tilde{v}(\mu(\mathbf{x}),\gamma,\varepsilon,\omega_d|\omega_R)]\\ &= [\mu(\mathbf{x}) + \gamma + \varepsilon][(1-\alpha)\Psi_R + \alpha m u(\omega_R)],\end{aligned}$$

[6] With this formulation, the evidence reviewed by Loewenstein, O'Donoghue, and Rabin (2003) suggests that people usually exhibit $\alpha \in (0,1)$. But any α generates a coherent model of behavior: $\alpha < 0$ would mean that the person understands qualitative taste changes but overestimates magnitudes, $\alpha = 1$ would mean the person mistakenly believes that her tastes don't change, and $\alpha > 1$ would mean that the person believes taste changes go in the opposite direction from what they actually do. Hence, our empirical analysis does not merely test whether we can reject the rational model; it also tests whether the direction of mispredictions is consistent with projection bias. In other words, our testable implications derived in this section will permit us to distinguish $\alpha > 0$ from $\alpha = 0$ (and from $\alpha < 0$).

[7] As discussed by Loewenstein, O'Donoghue, and Rabin (footnote 13), the person must be unaware of her current misprediction, because otherwise there would be no error. The question is whether a person could be aware of her tendency to mispredict in some future situation, while at the same time be unaware that she is mispredicting when that future situation arrives; our analysis here assumes the answer is no.

where $m \equiv \sum_{d=R+1}^{R+M} \delta^{d-R} = (\delta - \delta^{M+1})/(1-\delta)$ and Ψ_R is the same expected discounted instrumental utility as for fully rational people. Hence, the perceived expected discounted utility $\tilde{U}_R$ is identical to the actual expected discounted utility U_R, except that the actual expected discounted instrumental utility Ψ_R is replaced by the perceived expected discounted instrumental utility $\tilde{\Psi}_R(\omega_R) \equiv (1-\alpha)\Psi_R + \alpha m u(\omega_R)$. On the return date R, a person with projection bias will keep the item if

$$[\mu(\mathbf{x}) + \gamma + \varepsilon]\tilde{\Psi}_R(\omega_R) - p \geq -c$$

or

$$\gamma + \varepsilon \geq [(p-c)/\tilde{\Psi}_R(\omega_R)] - \mu(\mathbf{x}) \equiv \tilde{\Lambda}(\omega_R).$$

Next, consider how a person with projection bias behaves on the order date. Again, her predictions of the future usefulness of the item are biased by her utility given the current weather, which is now the weather on the order date ω_O. Because the person is unaware of how her prediction is being biased by the current weather, she believes on the order date that she will keep the item on the return date when $\gamma + \varepsilon \geq \tilde{\Lambda}(\omega_O)$. A logic identical to that for rational people—again, see Appendix A for the details—implies that a person with projection bias will follow a cutoff rule wherein she will order the item when

$$\gamma \geq \tilde{\Lambda}(\omega_O) - \bar{\varepsilon}(c/\tilde{\Psi}_R(\omega_O)) \equiv \tilde{\gamma}(\omega_O),$$

where $\bar{\varepsilon}(a)$ is defined as before.

Hence, the behavior for a person with projection bias, like the behavior for rational people, can be described by two cutoffs, the order-date cutoff $\tilde{\gamma}(\omega_O)$ and the return-date cutoff $\tilde{\Lambda}(\omega_R)$. This behavior can be depicted in a figure exactly analogous to Figure 2.

D. *Testable Predictions*

We now develop the testable implications of this theoretical framework. Because fully rational behavior is incorporated as a special case of behavior by people with projection bias (when $\alpha = 0$), we can limit attention to the projection-bias formulation. In our data—which we discuss in more detail in the next section—the dependent variable that we observe is whether a person returns an item conditional on ordering that item. Hence, we are interested in the predictions of our model for the probability of return conditional on ordering. Given that a person will order the item when $\gamma \geq \tilde{\gamma}(\omega_O)$ and, after ordering, return the item when $\gamma + \varepsilon < \tilde{\Lambda}(\omega_R)$, the probability of return conditional on ordering is

$$\Pr[\text{return} \mid \text{order}] = \frac{\Pr[\gamma \geq \tilde{\gamma}(\omega_O) \text{ and } \gamma + \varepsilon < \tilde{\Lambda}(\omega_R)]}{\Pr[\gamma \geq \tilde{\gamma}(\omega_O)]}.$$

As a first step toward deriving comparative statics, Lemma 1 describes how $\Pr[\text{return} \mid \text{order}]$ depends on the order-date and return-date cutoffs. (All proofs are collected in Appendix A.)

LEMMA 1: $\Pr[\text{return} \mid \text{order}]$ *is increasing in the return-date cutoff* $\tilde{\Lambda}(\omega_R)$ *and decreasing in the order-date cutoff* $\tilde{\gamma}(\omega_O)$.

Lemma 1 reflects that $\Pr[\text{return} \mid \text{order}]$ can be influenced either directly via return-date decisions or indirectly via order-date decisions. The direct *return-date effect* is straightforward: the more prone people are to return an item on the return date—that is, the higher the return-date cutoff—the higher is $\Pr[\text{return} \mid \text{order}]$. The indirect *order-date effect* is (slightly) more subtle, but the key intuition is that the individuals who are on the margin between ordering and not ordering are more prone to return the item than the individuals who are above the margin. Hence, the more prone people are to order an item—that is, the lower the order-date cutoff—the higher is $\Pr[\text{return} \mid \text{order}]$.

Proposition 1 uses Lemma 1 to establish our main predictions with regard to projection bias.

PROPOSITION 1:

(i) If $\alpha = 0$, *then the return-date cutoff* $\tilde{\Lambda}(\omega_R)$ *and the order-date cutoff* $\tilde{\gamma}(\omega_O)$ *are both independent of* $u(\omega_R)$ *and* $u(\omega_O)$, *and therefore* $\Pr[\text{return} \mid \text{order}]$ *is independent of both* $u(\omega_R)$ *and* $u(\omega_O)$;

(ii) If $\alpha > 0$, *then* $\tilde{\Lambda}(\omega_R)$ *is decreasing in* $u(\omega_R)$ *while* $\tilde{\gamma}(\omega_O)$ *is independent of* $u(\omega_R)$, *and*

therefore Pr[return | order] *is decreasing in* $u(\omega_R)$*; and*

(iii) *If* $\alpha > 0$*, then* $\tilde{\gamma}(\omega_O)$ *is decreasing in* $u(\omega_O)$ *while* $\tilde{\Lambda}(\omega_R)$ *is independent of* $u(\omega_O)$*, and therefore* Pr[return | order] *is increasing in* $u(\omega_O)$.

If the person does not have projection bias ($\alpha = 0$), then neither the weather on the order date nor the weather on the return date should affect Pr[return | order]. If the person has projection bias ($\alpha > 0$), in contrast, then the weather on the order date and the weather on the return date both influence Pr[return | order]. More precisely, if the return-date weather changes in a way that makes the item more valuable (if used at that time), then Pr[return | order] decreases. Intuitively, if the weather changes in this way, the person's return-date perception of the expected instrumental value increases, and she is therefore more prone to keep the item ($\tilde{\Lambda}(\omega_R)$ decreases). At the same time, the order decision is unaffected by the weather on the return date. It follows that Pr[return | order] must decrease.

The order-date weather has the opposite effect. Controlling for the weather on the return date, the return decision (conditional on ordering) is unaffected by the weather on the order date. But much as for the return-date weather, if the order-date weather changes in a way that makes the item more valuable (if used at that time), the person's order-date perception of the expected instrumental value increases, and she is therefore more prone to order ($\tilde{\gamma}(\omega_O)$ decreases). As discussed above, because the marginal people who order are the individuals most prone to return, Pr[return | order] increases.[8]

In order to identify control factors that we need to include in our reduced-form regressions, and to provide additional testable implications to check the validity of our model, it is useful to consider comparative statics with respect to the other variables Ψ_R, $\mu(\mathbf{x})$, c, and p. Proposition 2 describes these comparative statics.

PROPOSITION 2:

(i) *For any* α*, an increase in* c *leads to a decrease in* $\tilde{\Lambda}(\omega_R)$ *and an increase in* $\tilde{\gamma}(\omega_O)$*, and therefore* Pr[return | order] *is decreasing in* c*;*

(ii) *For any* α*, an increase in* $\mu(\mathbf{x})$ *or a decrease in* p *leads to a decrease in both* $\tilde{\Lambda}(\omega_R)$ *and* $\tilde{\gamma}(\omega_O)$*; and*

(iii) *For any* $\alpha < 1$*, an increase in* Ψ_R *leads to a decrease in both* $\tilde{\Lambda}(\omega_R)$ *and* $\tilde{\gamma}(\omega_O)$*; for any* $\alpha > 1$*, an increase in* Ψ_R *leads to an increase in both* $\tilde{\Lambda}(\omega_R)$ *and* $\tilde{\gamma}(\omega_O)$.

Each of these variables influences Pr[return | order] via both the direct and indirect avenues discussed after Lemma 1. For instance, an increase in the return cost c makes people less prone to return the item, which tends to decrease Pr[return | order] (return-date effect). It also makes people less prone to order the item, and since the marginal individuals are the most prone to return, this selection effect also tends to decrease Pr[return | order] (order-date effect). It follows that an increase in the return cost unambiguously reduces Pr[return | order].

For the three other variables, the return-date and order-date effects go in opposite directions, so it is ambiguous how each variable affects Pr[return | order]. If, however, the distribution of γ satisfies the usual hazard-rate property—$G'(\gamma)/(1 - G(\gamma))$ increasing in γ, which holds for many common continuous distributions including normal and uniform distributions—then the direct return-date effect seems to dominate for most parameter values. If the return-date effect does dominate, we would expect an increase in Ψ_R (if $\alpha < 1$) or $\mu(\mathbf{x})$, or a decrease in p to decrease Pr[return | order]. Our empirical analysis suggests that this is indeed the case.[9]

[8] If $\alpha < 0$, the conclusions in parts (ii) and (iii) of Proposition 1 are reversed: Pr[return | order] is increasing in $u(\omega_R)$ and decreasing in $u(\omega_O)$. Hence, to the extent that we find support for our two main predictions, we reject not only $\alpha = 0$ but also $\alpha < 0$. Note further that, if $u(\omega)$ is actually independent of ω, then projection bias over the weather would of course not influence behavior. We mention this point because, as a kind of placebo, our empirical analysis will check whether we see our two main predictions for non-winter items, and indeed we do not.

[9] Our main predictions from Proposition 1 permit reduced-form tests for $\alpha > 0$ versus $\alpha \leq 0$. Part (iii) of Proposition 2 suggests a reduced-form test for $\alpha < 1$ versus $\alpha > 1$. If we are examining items that become more

E. *Discussion*

Before turning to our empirical analysis, it is worth discussing a few issues with regard to our model. Our theoretical analysis assumes a specific functional form for the daily utility function—specifically, $v(\mu(\mathbf{x}),\gamma,\varepsilon,\omega_d) = [\mu(\mathbf{x}) + \gamma + \varepsilon]u(\omega_d)$. It is worth highlighting that our main qualitative comparative statics—as reflected in Propositions 1 and 2—do not rely on this functional form. In particular, these results hold for any utility function $v(\mu(\mathbf{x}),\gamma,\varepsilon,\omega_d)$ that is increasing in the first three arguments and monotonic in the fourth argument. We use the specific functional form above because it is what we will use for the structural estimation in Section IV. A formal analysis of the more general model is available from the authors.

Our theoretical analysis also assumes that people have a correct understanding of their local weather, and that the bias arises from an underappreciation of how the weather affects their utility from winter-clothing items—that is, they have projection bias with regard to the effects of the weather on their utility. Similar predictions might arise, however, in a model where people have a correct understanding of their utility from winter-clothing items, but their predictions for the future weather are biased toward the current weather. Hence, our empirical analysis cannot distinguish between people having projection bias versus people mispredicting the weather. This issue is particularly important for the broader relevance of our analysis—whether it's relevant only for weather-related decisions or for a wider range of economic decisions. It is also important from a policy perspective—e.g., whether we should help people achieve a better understanding of local weather patterns, or help them better appreciate changes in their utility.

An indirect way to assess whether it's projection bias versus mispredicting the weather is to examine the evidence of projection bias in other domains. In particular, for much of the evidence reviewed by Loewenstein, O'Donoghue, and Rabin (2003), there is no uncertainty about objective states, and thus the bias must come from mispredicting utilities. A more direct way is to examine evidence on people's predictions for their local weather. The one relevant study that we found (Joachim Krueger and Russell W. Clement 1994) asked Brown University undergraduates to estimate the average high and low temperatures in Providence for various days of the year. In general, the students were relatively accurate in their predictions. Particularly since undergraduates should be less familiar with the local weather than full-time residents, this study casts a little doubt on the notion that people mispredict the local weather. Even so, it is only one study, and moreover we could not find any evidence on the more relevant question of how individuals' expectations of future weather conditions might depend on the current weather.

A related issue is whether the degree of projection bias depends on how far into the future one is predicting. Our theoretical analysis assumes that the answer is no. In other words, we assume, for instance, that if the current temperature is 20°F, and a person is predicting the utility of a winter jacket on a 40°F day, her prediction is the same whether that 40°F day is the following week, month, or year. This assumption reflects that projection bias is caused by an "empathy gap" (Loewenstein 1996) wherein a person fails to grasp how she'll feel on a future 40°F day, and there is no reason to think that temporal distance will affect this empathy gap. It is worth noting, however, that even if the degree of projection bias is independent of temporal distance, the implications of projection bias can still depend on temporal distance because of discounting. In particular, for a fixed degree of projection bias, our model predicts that changes in the order-date temperature will have a larger effect if most of the usage occurs in the near future as opposed to a more distant future.

II. Data

The data were obtained from three sources: a large US company that sells outdoor apparel and gear, the National Climatic Data Center, and the US Census.

The apparel and gear company provided detailed information on orders of weather-related items. Specifically, the company provided a list of all its item categories, and we selected

valuable in colder temperatures, and if we believe the return-date effect dominates, then finding that colder average temperatures are associated with fewer returns supports $\alpha < 1$, while the reverse finding would support $\alpha > 1$. Our empirical results support the former.

from this list the set of item categories that could plausibly contain weather-related items. The company then provided information on every item ordered from one of these categories between January 1995 and December 1999 (over 12 million items). Our primary empirical analysis will test for projection bias with regard to temperature and snowfall using items from the following seven categories: gloves/mittens, winter boots, hats, sports equipment, parkas/coats, vests, and jackets. In addition, we tested for projection bias with regard to rainfall using items from the rain-wear category; we briefly discuss these results in Section III.[10]

For each item ordered, information was provided on the date the item was ordered, the date the item was shipped, whether the item was returned, and, if so, the date on which the company restocked the item.[11] In addition, we have information on the five-digit zip code associated with the billing address, whether the shipping address is the same as the billing address, the price of the item, whether the order was placed over the Internet, by phone, or through the mail, and whether the buyer used a credit card to purchase the item. The company also provided us with information that allows us to construct, for each item ordered, the two-day window during which the buyer is most likely to have received the item.[12]

In addition, the company provided some general information on the buyer (household identification number, gender, the number of items in the particular order, the number of items ordered from the company in the past, and the number of these items returned) and on the item (item identification number, item category, and whether the item was designed for a woman, man, girl, boy, child, or infant). We do not have more detailed information about the characteristics of each item; for example, we do not know the size or color of an item. However, for items in the parkas/coats and jackets categories, we know whether each item has a temperature rating and if so what it is, and, for the boots category, we know whether each item is designated a "winter" or "non-winter" boot.

The National Climatic Data Center collects daily observations of maximum and minimum temperature, snowfall amount, and rainfall amount from 1,062 weather stations across the contiguous United States. Throughout our analysis, we construct the daily temperature as the midpoint of the maximum and minimum temperatures.[13] Almost all of these 1,062 weather stations have daily records from the early 1960s. The Carbon Dioxide Information Analysis Center (CDIAC) assembles this information in a database. In addition, the CDIAC maintains a dataset containing the longitude and latitude of each weather station. We were also able to obtain from the US Census the longitude and latitude for the centroid of each five-digit zip code in the contiguous United States. Using these longitudes and latitudes, for each zip code we identified the three closest weather stations. For each zip code, if the closest weather station has weather information, then we assign those daily weather conditions to that zip code. If there is missing weather information from the closest weather station, then we consider the second-closest weather station; and if there is missing weather information from the second-closest weather station, then we consider the third-closest weather station.[14]

After merging the weather information with the order information using the five-digit zip code, we drop observations for items that are

[10] Of the over 12 million items, 8.84 million belonged to one of the seven winter-related categories, and 546,756 belonged to the rain-wear category. The remainder were from the following categories: (a) pants, shorts, and shirts (knit and woven); (b) fleece; (c) outerwear pants; and (d) sweaters. We do not use these categories because company representatives indicated that these categories contain many non-cold-weather items that could not be differentiated from the cold-weather items.

[11] At this company, items can be returned at any time after purchase and for any reason. Representatives at the company told us that several days could pass between when the company received a returned item and when that item was restocked.

[12] We construct this window by combining the date the item was shipped with an estimate for the number of days that the item was in the mail, where the latter depends on the region of the country in which the buyer is located and the day of the week on which the item was shipped. In the empirical estimation, we take the average weather conditions in this two-day window and refer to this average as the return-date weather conditions.

[13] The average difference between the daily maximum and minimum temperatures is slightly more than 20°F .

[14] The average distance between a zip code and the closest weather station is less than 21 miles (for those observations used in the estimation).

less likely to be cold-weather related. In the parkas/coats category and the jackets category, we drop observations where the item either did not have a temperature rating or had a temperature rating greater than 0°F (3,301,210 observations). In the boots category, we drop observations where the item was not designated as "winter" boots (597,579 observations). In the five remaining categories, we drop observations involving items for which more than 50 percent of the orders occurred when the average temperature in the month the order was placed and in the month after the order was placed was greater than 40°F (360,790 observations). In Section III we address the robustness of our results to variations in this criterion.

To avoid some confounds, we further reduce the dataset in a variety of ways. First, in an effort to focus on orders placed by individuals (and not corporations), we drop observations where the order consisted of more than 9 items (538,136 observations). Second, we drop observations where the order was placed through the mail because for such orders we cannot precisely identify the order date (584,617 observations). Third, we drop observations for which the shipping address differed from the billing address because we cannot identify at which address the buyer placed the order (641,635 observations). Fourth, because some households order multiple units of the same item (sometimes as part of a single order, sometimes as part of multiple orders) and it is unclear what is driving such orders, we drop all such observations (527,184 observations).[15] Fifth, we drop observations for which we do not have weather information, which means all orders from Alaska and Hawaii and all orders where the three closest weather stations all had missing weather information (89,065 observations). Finally, we drop observations for which the price of the item is less than $4 (78 observations).

After reducing the dataset, there are 2,200,073 observations. Summary statistics for each of the seven winter-related categories are presented in Table 1. Table 1 indicates that there are many different items within each category. For example, there are 233 different items in the sports equipment category, and 133 different items in the parkas/coats category. Table 1 also indicates that return rates vary across categories from 6.6 percent for sports equipment to 22.2 percent for parkas/coats. With the exception of sports equipment, categories with higher-priced items have higher return rates. In all categories other than sports equipment, the return rate is higher than the buyers' return rate on prior purchases from the company. This feature is mostly due to the fact that we code first-time buyers as having a return rate on prior purchases of zero. In addition, it could also be that the items in our dataset have higher return rates than the other items that the company sells, or that first-time buyers have higher return rates than repeat buyers (our estimation results provide some support for the latter).

For most categories, the average time between the order date and the shipping date is less than a day, reflecting that items are often shipped on the day of the order. Across all categories, the average time between the order date and the receiving date is 4.84 days, and 85 percent of orders are received within six days.[16] This indicates that most items are delivered approximately three to four days after being shipped. Around 3 percent of the items were ordered through the Internet, about 70 percent were ordered by a female, and over 95 percent were purchased using a credit card. The average number of items in an order was three.

For the parkas/coats category and the jackets category, Table 1 notes that the average temperature rating (from among the items with temperature ratings of 0°F or below) is about −10°F

[15] One possible source of such orders is that a household is ordering different sizes or colors with the intent of keeping only their favorite. Evidence supporting this explanation comes from the fact that such orders were less prevalent for items where one would expect fit to be less of a concern (e.g., gloves/mittens, hats, sports equipment, and vests) and more prevalent for items where fit is likely more important (i.e., boots, parkas/coats, and jackets). There are also instances in our dataset where a household ordered an item, returned it, and then ordered the same item again. A natural interpretation of such instances is that they are exchanges for a different size or color that somehow got coded as multiple orders. A second source of such orders is that a household wants to purchase and keep multiple units of the same item, as either multiple units for one household member or one unit for multiple household members.

[16] For our measure of days between order and receipt, we use days between the order date and the first day of the two-day window during which the buyer is most likely to have received the item.

Table 1—Summary Statistics

	Gloves/ mittens	Winter boots	Hats	Sports equipment	Parkas/ coats	Vests	Jackets	All seven categories
Observations	484,084	262,610	484,086	146,594	524,831	151,958	145,910	2,200,073
Number of items	106	93	88	233	133	20	37	710
Percent returned	10.9	15.6	10.8	6.6	22.2	12.8	18.0	14.4
Price of item (dollars)	29.26	68.33	23.74	74.10	148.58	40.90	106.70	70.10
First-time buyer	0.15	0.18	0.17	0.14	0.23	0.17	0.18	0.18
Number of prior purchases	27.3	22.2	23.9	27.7	20.5	21.71	25.3	23.83
Percent of prior purchases returned	7.2	6.6	6.9	7.2	7.3	6.8	8.2	7.14
Days between order and shipment	0.42	0.97	0.72	0.94	2.17	1.24	1.13	1.11
Days between order and receipt	4.13	4.66	4.46	4.58	5.92	5.04	4.89	4.84
Ordered through Internet	0.04	0.03	0.03	0.02	0.04	0.02	0.05	0.03
Ordered by female	0.71	0.66	0.71	0.70	0.66	0.72	0.66	0.69
Purchased with credit card	0.97	0.98	0.98	0.97	0.98	0.98	0.97	0.98
Items in order	3.5	2.5	3.4	2.9	2.2	2.8	2.3	2.9
Temperature rating					−10.11		−5.64	
Weather conditions								
Order-date temperature (°F)	40.60	39.74	41.48	37.81	43.29	44.76	46.88	41.85
Receiving-date temperature (°F)	39.90	38.97	40.72	36.70	42.29	43.20	45.70	40.94
Order-date snowfall (0.1″)[a]	1.79	2.69	1.69	2.65	1.30	1.26	0.63	1.70
Receiving-date snowfall (0.1″)[a]	1.58	2.32	1.51	2.35	1.33	1.43	0.66	1.57

[a] Snowfall is based on those observations where there exists snowfall information.

and −5°F, respectively. The temperature ratings range from 0°F to −40°F for parkas/coats and from 0°F to −20°F for jackets. Finally, Table 1 also provides information on the average weather conditions at the order date and at the receiving date. Because most of these items are ordered at the end of fall or the start of winter and received a few days later, the average temperature at the order date is slightly higher than at the receiving date.

In addition, we investigated the extent of actual serial correlation in temperatures. First, we calculated the day-to-day correlations of de-meaned temperatures (aggregated across stations).[17] The day-to-day correlation of de-meaned temperatures is 0.70, and this correlation declines to 0.40, 0.25, 0.19, 0.15, 0.11, and 0.07 for dates two through seven days apart, respectively. For dates that are two weeks apart, the correlation is 0.03. Second, for the observations in our dataset, we calculated the correlation between the de-meaned order-date temperature and the de-meaned return-date temperature, which is 0.21.

III. Reduced-Form Estimation

In this section we use estimates from reduced-form models to test the two main predictions of projection bias from Proposition 1. Our initial focus will be projection bias with regard to temperature. Specifically, for cold-weather items, we expect the instrumental utility to be larger the lower is the temperature. If so, then, applying Proposition 1, projection bias predicts that the likelihood of returning an item (conditional on ordering) will be decreasing in the order-date temperature and increasing in the return-date temperature. In other words, a lower order-date temperature or a higher return-date temperature is associated with a higher probability of returning the item.[18]

[17] To do so, we first calculated an average temperature for each calendar-date/weather-station pair by using a seven-day window around the calendar date (three days before to three days after) for the years 1990 to 1999. We then used these averages to de-mean the daily temperature data, and computed correlation coefficients for the de-meaned data. These measures of serial correlation aggregate station-specific serial correlation.

[18] Another prediction of projection bias is that lower temperatures will increase sales. To test this prediction, we regressed total daily sales in a zip code on temperature in that zip code and on average temperature between 1990 and

To test these predictions, we first estimate a probit model that permits the likelihood of returning an item to depend on the local temperature on the order date ($T_{O,i}$), the local temperature on the receiving date ($T_{R,i}$), and the expected local temperature (ET_i). We include the expected local temperature to control for the expected instrumental value of the item (i.e., the variable Ψ_R from our theoretical model). For the order-date temperature, $T_{O,i}$, we use the temperature in that zip code for that day. For the return-date temperature, $T_{R,i}$, we use the average daily temperature over the two-day window during which the buyer is most likely to have received the item. For the expected weather, ET_i, we use the average winter temperature in that zip code from 1990 through 1994.[19] Letting $y_i = 1$ if order-item i is returned and $y_i = 0$ otherwise, we estimate the following probit model:

$$y_i = 1 \text{ if } y_i^* > 0;$$

$$y_i = 0 \text{ if } y_i^* \leq 0,$$

where

$$y_i^* \equiv \mathbf{D}_i\mathbf{a} + \mathbf{X}_i\mathbf{b} + \beta_1(T_{O,i}) + \beta_2(T_{R,i}) + \beta_3(ET_i) + \upsilon_i,$$

and υ_i is standard normally distributed and is assumed to be uncorrelated with the regressors. The υ_i's are allowed to be correlated across observations with the same household identification number, but otherwise the υ_i's are assumed to be uncorrelated with each other. The vector $\mathbf{D}_i$ incorporates fixed effects for clothing type (whether the item was designed for a woman, man, girl, boy, child, or infant), item identification number, month-region, and year-region.[20] The vector $\mathbf{X}_i$ includes the price of the item, the number of days between the order date and the shipment date, the number of days between the order date and the receiving date, the number of items the buyer had previously purchased, the percentage of those items that were returned, the number of items in the order, and indicator variables for whether the item was ordered through the Internet, whether the buyer was female, whether the buyer was a first-time buyer, and whether the item was purchased using a credit card.

For each category and for the entire dataset, Table 2 presents the estimated marginal effects (not the coefficient estimates) associated with a change in each independent variable. The marginal effects are calculated at the sample means of the regressors. First, note that the estimated marginal effect for average winter temperature is positive and statistically significant for six of the seven categories and for the entire dataset. This result provides some confirmation that these are indeed cold-weather items for which colder temperatures imply higher utility. In particular, if a lower average winter temperature yields a larger expected instrumental utility (Ψ_R), and if the direct return-date effect dominates the indirect order-date effect (recall our discussion of these effects in Section I), Proposition 2 implies that a lower average winter temperature would indeed yield a lower probability of return.

The estimates provide strong support for the first implication of projection bias that the likelihood of returning an item should be greater the lower is the order-date temperature. For all eight specifications, the marginal effect associated with temperature on the order day is indeed negative and this marginal effect is statistically significant for six of the eight specifications. In terms of the magnitude of the effect, these estimates indicate that a reduction in the order-date temperature by 30°F—e.g., a person orders on a 10°F day as opposed to a 40°F day—will increase the probability of a return by 0.39 percentage points for gloves/mittens, 0.78 percentage points for boots, 0.60 percentage points for hats, 0.33 percentage points for sports equipment, 0.27

1994 in that zip code on that day, and then incrementally include zip code, month, and year fixed effects. Our point estimates from these regressions (which are all statistically different from zero at the 1 percent confidence level) suggest that a decrease in temperature of 10°F increases daily sales in a zip code between 11 and 21 percent. These empirical results are available from the authors.

[19] Specifically, we use the average daily temperature over the months of December, January, February, and March. We discuss robustness to alternative measures later in this section.

[20] We divide the contiguous United States into six regions: Pacific, Southwest, Great Lakes, Southeast, Northeast, and Mountain/Prairie.

Table 2—Baseline Probit Regressions

	Gloves/ mittens	Winter boots	Hats	Sports equipment	Parkas/ coats	Vests	Jackets	All seven categories
Order-date temperature	−0.00013**	−0.00026**	−0.00020**	−0.00011*	−0.00009	−0.00048**	−0.00014	−0.00019**
	(0.00005)	(0.00009)	(0.00005)	(0.00006)	(0.00007)	(0.00011)	(0.00013)	(0.00003)
Receiving-date temperature	0.00005	0.00018*	−0.00005	−0.00008	0.00007	−0.00010	0.00010	0.00003
	(0.00006)	(0.00009)	(0.00006)	(0.00007)	(0.00008)	(0.00011)	(0.00014)	(0.00003)
Average winter temperature, 1990–1994	0.00029**	0.00055**	0.00038**	0.00042**	0.00056**	0.00098**	0.00035	0.00049**
	(0.00010)	(0.00016)	(0.00010)	(0.00012)	(0.00013)	(0.00018)	(0.00022)	(0.00005)
Days between order and shipment	−0.00189**	−0.00075	−0.00136**	−0.00032	−0.00179**	0.00141*	−0.00173*	−0.00105**
	(0.00048)	(0.00072)	(0.00044)	(0.00052)	(0.00060)	(0.00086)	(0.00101)	(0.00023)
Days between order and receipt	0.00065	−0.00008	0.00029	0.00035	0.00069	−0.00213**	0.00082	0.00029
	(0.00043)	(0.00069)	(0.00041)	(0.00050)	(0.00058)	(0.00082)	(0.00096)	(0.00022)
Ordered through Internet	−0.01083**	−0.01357**	−0.00965**	−0.00796**	−0.01556**	−0.00466	−0.01391**	−0.01153**
	(0.00246)	(0.00440)	(0.00262)	(0.00296)	(0.00311)	(0.00572)	(0.00478)	(0.00129)
Ordered by female	0.00435**	0.01197**	0.00823**	0.00590**	0.01259**	0.00146	0.00180	0.00781**
	(0.00101)	(0.00155)	(0.00095)	(0.00116)	(0.00126)	(0.00193)	(0.00216)	(0.00051)
First-time buyer	0.01570**	0.01531**	0.01065**	0.00202	0.01535**	0.01587**	0.02448**	0.01394**
	(0.00149)	(0.00213)	(0.00144)	(0.00177)	(0.00159)	(0.00261)	(0.00312)	(0.00070)
Number of prior purchases	0.00013**	0.00026**	0.00017**	0.00005**	0.00020**	0.00014**	0.00013**	0.00016**
	(0.00001)	(0.00002)	(0.00001)	(0.00001)	(0.00002)	(0.00003)	(0.00003)	(0.00001)
Percent of prior purchases returned	0.19922**	0.24204**	0.19078**	0.06806**	0.30153**	0.20275**	0.30637**	0.22252**
	(0.00364)	(0.00646)	(0.00558)	(0.00498)	(0.00446)	(0.00679)	(0.01016)	(0.00216)
Price of item	0.00075**	0.00005	0.00145**	0.00033**	0.00019**	0.00166**	0.00016	0.00023**
	(0.00024)	(0.00013)	(0.00025)	(0.00008)	(0.00004)	(0.00024)	(0.00018)	(0.00003)
Purchased with credit card	0.02042**	0.04337**	0.02876**	0.02395**	0.05893**	0.02294**	0.05312**	0.03531**
	(0.00250)	(0.00418)	(0.00244)	(0.00191)	(0.00405)	(0.00535)	(0.00568)	(0.00137)
Items in order	−0.00157**	0.00012	−0.00035	−0.00078**	0.00196**	−0.00177**	0.00141**	−0.00028**
	(0.00022)	(0.00039)	(0.00022)	(0.00028)	(0.00033)	(0.00045)	(0.00058)	(0.00012)
Clothing-type fixed effects	YES	YES	YES	NO[a]	YES	YES	YES	YES
Item fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
Month-region fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
Year-region fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
Observations	484,067	262,610	484,085	146,403	524,831	151,958	145,910	2,199,950
R-squared	0.04	0.05	0.07	0.13	0.03	0.03	0.04	0.07

Notes: For each column, the dependent variable is whether an item is returned (= 1 if item returned, and 0 otherwise), and the table presents the estimated marginal effects calculated at the sample means of the regressors. Standard errors are in parentheses—the standard errors are robust to arbitrary heteroskedasticity and correlation within a household.

* Significant at the 10 percent level.

** Significant at the 5 percent level.

[a] Clothing-type information was not provided for sports equipment items.

percentage points for parkas/coats, 1.44 percentage points for vests, and 0.42 percentage points for jackets. Estimating the specification using orders from all seven categories indicates that an order-date temperature decrease of 30°F will increase the probability of a return by 0.57 percentage points, which corresponds to a 3.95 percent increase in return rate.[21]

In contrast, these estimates do not provide strong support for the second implication of projection bias that the likelihood of returning an item should be larger the higher is the return-date temperature. The marginal effect associated with this temperature varies in sign across categories and is statistically significant only for winter boots. While we find slightly stronger support in other reduced-form specifications discussed below, we suspect the reason we find

[21] In our dataset, it is not uncommon to see a 30°F change in temperature within a zip code over a relatively short time span. To demonstrate this, we looked at daily temperatures over 14-day windows (recall that our daily temperature is the midpoint between each day's maximum and minimum temperatures). Specifically, for every weather station and for every 14-day window from 1990 through 1999, we calculated the difference between the highest and lowest daily temperatures over the 14 days. Across all weather stations, this difference was larger than 30°F in almost 15 percent of the cases, and the average difference was 20.64°F. Moreover, the 3.95 percent increase in return rate may understate the impact of the order-date temperature on returns, because colder temperatures are also associated with more orders.

only limited evidence for this second implication is not that people don't have projection bias, but rather that we are unable to identify exactly when people make the return decision. Our theoretical analysis assumes that there is a specific day on which the return decision is made, and it is the temperature on that day that matters. Our empirical specification assumes this return decision is made on the day the item is received. In our dataset, however, there is significant variation in the duration between when an item is received and when it is restocked. This duration ranges from 2 to 222 days, with a mean of 24.52 days, and about 75 percent of returns are restocked within 30 days.

To deal with this issue, in our robustness section we discuss using a wider window of days to proxy for the return-decision weather. Another option would be to use the restocking information to infer when the individual makes the return decision. However, this approach would require arbitrary assumptions about the lag between when a person makes the return decision and when she mails the item, and about the lag between when the company receives the return and when the item is restocked. In addition, for people who keep the item, we would need to make an assumption about when they made the return decision. Given the relatively low temperature correlations, even for days that are only three or four days apart, we would doubt the validity of any results using this approach, and hence we do not pursue it.

More generally, it may not even be the case that there is a specific day on which the return decision is made. Rather, we suspect that many people, once they have an item, assess whether to return the item over a number of days, perhaps trying the item out on multiple occasions. If so, then even if projection bias is influencing a person's feelings on every given day, it might have a relatively small effect if a person makes the decision by integrating their feelings over multiple days.

We next modify our analysis to control for household-specific fixed effects.[22] In order to obtain a sufficient number of observations, we use the dataset containing observations from all seven categories. We use a subset of this dataset that includes only those observations for which: (a) the household appears more than once in the dataset; (b) the household has not ordered multiple items from the specific category; and (c) the household has both returned some items and kept other items (because the coefficient estimates are now identified based on within-household variation). We drop multiple household orders from the same category because such orders are likely correlated—if a household orders a winter coat, whether they subsequently order and return another winter coat will likely depend on whether they returned the first winter coat.[23] We include all the same regressors as in Table 2, but since there are so many fixed effects, for computational reasons we are forced to estimate a linear model.[24]

Table 3 presents the coefficient estimates from this specification, both with and without household-specific fixed effects. In both cases, the coefficient estimate for average winter temperature is positive, the coefficient estimate for order-day temperature is negative, and both are statistically significant. Hence, we again see support for our conjecture that these are cold-weather items and support for the first implication of projection bias. In addition, the coefficient estimate for order-date temperature increases substantially when we control for household fixed effects. Hence, our negative marginal effects associated with order-date temperature in Table 2 cannot be explained by the current weather conditions influencing the types of households who order. Finally, again we do not find strong support for the return-date prediction—while the coefficient estimate is positive and larger when we control for household fixed effects, it is not statistically significant.

We next investigate whether similar results are obtained for non-cold-weather clothing. Our

[22] We cannot include individual-specific fixed effects because we have only a household identification number and not an individual identification number.

[23] In this reduced dataset, the average price of an item is \$74.23 and the average number of items ordered by a household is 35.32, compared to \$70.10 and 23.83, respectively, for the entire dataset. As expected, the fraction of items returned is much higher in this reduced dataset, 0.39 compared to 0.14. The averages of the other variables do not differ appreciably between the reduced and entire dataset.

[24] In terms of the comparability of Tables 2 and 3, the results in Table 2 change little if the specification is estimated as a linear model instead of as a probit.

model predicts that there should be no effect for non-weather-related items, and that there should be nonmonotonic effects for warmer-weather items.[25] To test for such differential effects, we use items in the parkas/coats and jackets categories that have temperature ratings greater than 0°F, and items in the boots category that are designated non-winter boots—we believe these items include a combination of warmer-weather items and non-weather-related items. Specifically, we include both the cold-weather and non-cold-weather items, and we interact the order-date and return-date temperatures with a variable indicating whether it is a non-cold-weather item. Table 4 contains the estimation results for parkas/coats/jackets and boots when these interactive terms are included in our base specification. The main effect of order-date temperature on likelihood of returning a cold-weather item is much as before (in Table 2). The marginal effect associated with the order-date interaction term is positive for both parkas/coats/jackets and boots. Indeed, looking at the sum of the two marginal effects, the order-date temperature appears to have a negligible effect for non-cold-weather items. Irrespective of temperature rating, the estimated marginal effect of the return-date temperature is negligible for parkas/coats/jackets. For winter boots, the estimated marginal effect of the return-date temperature is positive and statistically significant (as in Table 2), while the return-date temperature appears to have a negligible effect on non-winter boot returns. Hence, consistent with projection bias, we are seeing our predicted effects only for cold-weather items.

Tables 2 and 3 also provide information on the effects of the various other covariates.[26] Many of these effects are consistent with natural (although admittedly post hoc) interpretations of the variables in our theoretical model when the direct return-date effect dominates the indirect order-date effect. Consider for instance the price of the item: applying Proposition 2, if the return-date effect dominates, then the probability of return would be increasing in the price; and indeed the estimated marginal effect of price is positive and statistically significant for most specifications in Table 2 and for both specifications in Table 3.

Many of the control variables are likely correlated with an individual's preference for the good ($\mu(\mathbf{x})$). Applying Proposition 2, and

TABLE 3—LINEAR REGRESSIONS WITH AND WITHOUT HOUSEHOLD FIXED EFFECTS

	With household fixed effects	Without household fixed effects
Order-date temperature	−0.00082**	−0.00039**
	(0.00027)	(0.00013)
Receiving-date temperature	0.00017	0.00002
	(0.00029)	(0.00015)
Average winter temperature, 1990–1994	0.00276**	0.00067**
	(0.00090)	(0.00019)
Days between order and shipment	−0.00710**	−0.00373**
	(0.00222)	(0.00114)
Days between order and receipt	0.00387*	0.00175
	(0.00214)	(0.00109)
Ordered through Internet	−0.01597	−0.00858
	(0.01521)	(0.00648)
Ordered by female	0.03454**	0.01398**
	(0.00647)	(0.00205)
First-time buyer	−0.09726**	−0.00913**
	(0.00983)	(0.00314)
Number of prior purchases	0.00016	−0.00013**
	(0.00011)	(0.00002)
Percent of prior purchases returned	−0.45905**	−0.06196**
	(0.02438)	(0.00618)
Price of item	0.00106**	0.00070**
	(0.00026)	(0.00015)
Purchased with credit card	0.05714**	0.02638**
	(0.01583)	(0.00741)
Items in order	0.00551**	0.00250**
	(0.00121)	(0.00053)
Clothing-type fixed effects	YES	YES
Item fixed effects	YES	YES
Month-region fixed effects	YES	YES
Year-region fixed effects	YES	YES
Household fixed effects	YES	NO
Observations	162,580	162,580
R-squared	0.19	0.10

Notes: For each column, the dependent variable is whether an item is returned (= 1 if item returned, and 0 otherwise), and the table presents the coefficient estimates from a linear regression. Standard errors are in parentheses—the standard errors are robust to arbitrary heteroskedasticity and correlation within a household.

* Significant at the 10 percent level.

** Significant at the 5 percent level.

[25] For instance, if a light jacket is useful for temperatures between 40°F and 60°F, and most useful for 50°F, then the probability of return conditional on ordering should be largest when the order-date temperature is 50°F.

[26] The marginal effects associated with these other covariates for the specification in Table 4 are almost identical to the estimates presented in Table 2 for the corresponding category.

Table 4—Probit Regressions that Include Non-Winter Items

	Parkas, coats, and jackets	Boots
Order-date temperature	−0.00010	−0.00027**
	(0.00006)	(0.00007)
(Order-date temperature) × (temperature rating > 0°F)	0.00020*	
	(0.00011)	
(Order-date temperature) × (non-winter boot)		0.00025**
		(0.00009)
Receiving-date temperature	0.00007	0.00017**
	(0.00006)	(0.00008)
(Receiving-date temperature) × (temperature rating > 0°F)	−0.00002	
	(0.00011)	
(Receiving-date temperature) × (non-winter boot)		−0.00029**
		(0.00009)
Average winter temperature 1990–1994	0.00045**	0.00040**
	(0.00010)	(0.00009)
Observations	866,029	651,347
R-squared	0.03	0.06

Notes: For each column, the dependent variable is whether an item is returned (= 1 if item returned, and 0 otherwise), and the table presents the estimated marginal effects calculated at the sample means of the regressors. Standard errors are in parentheses—the standard errors are robust to arbitrary heteroskedasticity and correlation within a household. Although we do not report the results, all covariates included in Tables 2 and 3 are also included here.

* Significant at the 10 percent level.

** Significant at the 5 percent level.

assuming that the return-date effect dominates, a positive (negative) marginal effect associated with a covariate might indicate a weaker (stronger) preference for the good. Under this interpretation, the estimates in Table 2 suggest that preferences are stronger for buyers who experience shipment delays, order on the Internet, are male, do not use a credit card, and are not first-time buyers.

The control variables might also be correlated with the cost incurred from returning an item (c). Applying Proposition 2, a positive (negative) marginal effect associated with a covariate might indicate a lower (higher) return cost. For instance, under this interpretation, using a credit card or having a larger return rate on prior purchases is associated with a lower return cost. We had expected that having more items in the order should decrease the return cost due to economies of scale in returning items. The estimates in Table 2 do not support this contention, while those in Table 3 do.[27]

In addition to testing for projection bias with regard to temperature, we test for projection bias with regard to snowfall. For snow-related items—that is, items for which the instrumental utility increases with the amount of snowfall—projection bias predicts that the likelihood of returning an item will be increasing in the order-date snowfall and decreasing in the return-date snowfall. We find minimal evidence for projection bias over snowfall. Using the full dataset, the first column of Table 5 contains the estimation results when we add to the base specification snowfall on the day of the order, average snowfall over the two-day window during which the item was likely received, and average snowfall from 1990 through 1994. The estimated marginal effect of order-date snowfall is positive and statistically significant at the 10 percent level. However, when separate estimations are done for each of the seven categories, the marginal effect estimates associated with

[27] One possible explanation why orders with multiple items may be less likely to be returned (especially for infrequent buyers) is that these orders are more likely instances where the orderer is different from the individual who actually uses the item.

Table 5—Probit Regressions that Include Snowfall and Prior Weather

	All seven categories	All seven categories	All seven categories
Order-date temperature	−0.00018**	−0.00018**	−0.00017**
	(0.00003)	(0.00003)	(0.00003)
Temperature 7 days prior to order		−0.00005*	−0.00005*
		(0.00003)	(0.00003)
Receiving-date temperature	0.00004	0.00004	0.00005*
	(0.00003)	(0.00003)	(0.00003)
Average winter temperature, 1990–1994	0.00050**	0.00052**	0.00052**
	(0.00006)	(0.00005)	(0.00006)
Order-date snowfall	0.00005*		0.00005*
	(0.00003)		(0.00003)
Snowfall 7 days prior to order			0.00003
			(0.00003)
Receiving-date snowfall	0.00001		0.00001
	(0.00004)		(0.00004)
Average winter snowfall, 1990–1994	0.00019		0.00012
	(0.00044)		(0.00044)
Observations	2,181,724	2,199,144	2,178,005
R-squared	0.07	0.07	0.07

Notes: For each column, the dependent variable is whether an item is returned (= 1 if item returned, and 0 otherwise), and the table presents the estimated marginal effects calculated at the sample means of the regressors. Standard errors are in parentheses—the standard errors are robust to arbitrary heteroskedasticity and correlation within a household. Although we do not report the results, all covariates included in Tables 2 and 3 are also included here.

* Significant at the 10 percent level.

** Significant at the 5 percent level.

snowfall on the order and return dates vary in sign and few are statistically significant. One explanation for why we don't find stronger evidence in regards to snowfall is that the instrumental utility an individual derives from some of these items does not depend on snowfall. The fact that the estimated marginal effect associated with average snowfall from 1990 through 1994 is not statistically significant suggests that this is indeed the case.[28]

Finally, we use orders for items in the rain-wear category to test for projection bias with regard to rainfall. We found no significant effect for order-date rainfall, receiving-date rainfall, and average annual rainfall—and hence no evidence of projection bias over the effects of rainfall.[29] Even so, as for snowfall, the fact that the probability of a return does not depend on average annual rainfall may suggest that the instrumental utility an individual derives from some of these rain-wear items does not depend significantly on the amount of rainfall. Another closely related explanation is that, unlike temperature, rainfall is often not an all-day event, but rather occurs over a small portion of the day. Hence, order-date rainfall would be relevant only for the subset of people who happened to experience that rainfall.

A. *Robustness*

We investigate the robustness of our empirical results along several dimensions.[30] First, we consider alternatives for our selection criterion wherein we dropped all items for which more than 50 percent of the orders occurred when the average temperature in the month the order was placed and in the month after the order was placed was greater than 40°F. If, instead, we

[28] Note that including the snowfall variables does not appreciably change the estimates associated with the temperature variables. Including the snowfall variables also does not appreciably change the marginal effects associated with the various other control variables (which are not presented in Table 5).

[29] All three marginal effects are estimated relatively precisely—these results are available from the authors.

[30] The empirical results discussed in this subsection are available from the authors.

use a cutoff temperature of 45°F, 35°F, or 30°F, the marginal effects in Tables 2 and 3 do not change significantly. Alternatively, if instead of dropping items, we drop orders from zip codes in which the temperature drops below 40°F, on average, fewer than 30 days a year, once again the marginal effects in Tables 2 and 3 do not change significantly.

We also consider the robustness of our results to including the observations for which the household ordered multiple units of the same item. When including these observations, we estimate a slightly modified model that incorporates two new indicator variables: one indicating whether the item was part of an order with multiple units of that item, and another indicating whether the item was also ordered again by the same household as part of a separate order. The marginal effects of the variables reported in Tables 2 and 3 again do not change significantly. The marginal effects for the multiple-unit indicator variables are positive and statistically significant, indicating that if an item is ordered multiple times by the same household (at the same time or at different times), the probability of a return is higher.[31]

We also test the robustness of our results to alternative specifications of the order-date weather and return-date weather. For instance, if instead of using order-date weather we use the average weather over the order date and the day prior to the order date, the marginal effects in Tables 2 and 3 do not change significantly. We consider an even broader window for the return date because of the problem of identifying when the return decision is made. In particular, instead of using the average temperature on the two-day window when the item is most likely received, we use the average temperature over the two weeks after the item is most likely received. The marginal effect of this new variable is positive in all categories except sports equipment, and it is statistically significant in four of those six categories. In addition, it is positive and statistically significant when we run the estimation combining observations from all seven categories. The fact that using a two-week window provides stronger evidence of projection bias on the return decision supports our earlier contention that the failure to find strong results on the return date reflects an inability to identify exactly when the return decision is made.

In our base specification, we control for the expected instrumental value of the item (i.e., the variable Ψ_R from our theoretical model) by including the average historical winter temperature in the zip code (based on data from 1990 through 1994). We also test whether our main results are robust to additional proxies for the expected instrumental value. Specifically, we use three additional proxies: (a) the average historical monthly temperatures for December, January, February, and March; (b) the average historical monthly temperatures for the first month after the order, the second month after the order, and the third month after the order; and (c) the actual average temperatures for the first month after the order, the second month after the order, and the third month after the order. For all three estimations, the marginal effects associated with the order-date temperature and the return-date temperature do not change significantly in any category, suggesting that our main results are robust to alternative proxies for the expected instrumental value.[32]

Finally, we note that our results are primarily attributable to orders placed in late fall and early winter. In particular, the marginal effects in Tables 2 and 3 change very little when these specifications are estimated after dropping all orders from February through September. This finding is not surprising given that 90 percent of our observations occur in October through January (14 percent occur in October, 32 percent occur in November, 36 percent occur in December, and 8 percent occur in January). When we estimate our base specification using

[31] These results are consistent with the explanation for such orders wherein households order multiple units of the same item planning to keep only one (or a subset).

[32] For (a), we replace our base proxy with this new proxy; for (b) and (c), we include these new proxies in addition to our base proxy. The marginal effects associated with the average historical monthly temperatures—either for the winter months or for the months subsequent to ordering—vary in sign and are often statistically significant. The largest effects are for the average January temperature in the former case and for the temperature in the first month after ordering in the latter case, both of which are positive and significant across all categories. The marginal effects for actual average temperatures vary in sign and are rarely statistically significant.

only orders placed from February through September, neither the order-date temperature nor the return-date temperature appears to systematically affect the probability of return. Note that this latter finding is still consistent with projection bias if the within-month temperature variation in these months is irrelevant to the value of a winter-clothing item. For instance, if a winter coat has no value on any day warmer than, say, 40°F, then, according to projection bias, the probability of return should be independent of whether the order-date temperature is 45°F, 55°F, or 65°F.[33] We'll return to this issue in Section IV.

B. *Alternative Explanations*

Our main empirical finding is that the colder the temperature on the order date, the more likely it is that the item is returned. We also find some weak evidence that the colder the temperature on the receiving date, the less likely it is that the item is returned. We next investigate whether these findings could be due to some other factor not incorporated into our theoretical analysis of projection bias.

Current Temperature Helps Predict Future Temperatures.—Our theoretical analysis assumes that the order-date temperature is completely orthogonal to future temperatures and thus orthogonal to the usefulness of the item. An obvious alternative is that the order-date temperature is, in fact, informative about future temperatures and thus serves as a proxy for the usefulness of the item. For instance, if a colder order-date temperature is correlated with colder temperatures in the first few weeks (or months) after receiving the item, then a colder order-date temperature would be correlated with the item being of higher value. However, there are three problems with this alternative explanation. First, theoretically, if the direct effect dominates the indirect effect—as suggested by our results in Tables 2 and 3 and by our later structural estimation—it implies that lower order-date temperatures would be associated with a decreased likelihood of return, which is exactly the opposite from what we find. Second, as we discuss at the end of Section II, the actual serial correlation in (de-meaned) temperatures is small—e.g., the correlation is 0.07 for days seven days apart, and it is 0.03 for days two weeks apart.[34] Third, because a longer delay between order and receipt should reduce the information content of the order-day temperature, this alternative implies that the order-date temperature should have a smaller effect for larger order-receipt delays. When we test this prediction—by reestimating our base specification with an interaction term between the order-date temperature and the order-receipt delay, and by reestimating our base specification using only those observations where the item is received at least six days after ordering—the marginal effect of the order-date temperature does not appear to depend on the order-receipt delay.[35]

Learning about the Local Weather.—A closely related alternative explanation is that, even though the order-date temperature may contain no actual information relative to historical averages, people may not know the historical averages and thus use the current weather to make inferences about the expected future weather. If so, a lower temperature on the order date would shift a person's order-date beliefs about the local weather toward lower temperatures, which would make the person perceive the item to be more valuable, which in turn would make the person more likely to order. In principle, then, this learning story is consistent with our main empirical finding. However, the comparative static is not unambiguous—in particular, a lower temperature on the order date would also shift the person's *return-date* beliefs toward

[33] Formally, if $u(T) = 0$ for all $T \geq 40$°F, then $(1 - \alpha)u(T_d) + \alpha u(T_t) = (1 - \alpha)u(T_d)$ for any current temperature $T_t \geq 40$°F.

[34] The fact that our order-date result is robust to including actual temperatures in the months after ordering (see footnote 32) further suggests that this result is not driven by the order-date temperature providing information about future usefulness.

[35] For the receiving date, this alternative explanation makes the same prediction as projection bias, and moreover, if the current temperature, in fact, contains information about near-future temperatures, then the receiving-date temperature ought to be even more informative about future value. Hence, the fact that we find minimal evidence that the receiving-date temperature affects the probability of a return perhaps casts further doubt on this alternative.

lower temperatures, which has the opposite effect. Moreover, this learning explanation also implies that the temperature, say, a week prior to the order date should contain similar information about the local weather. It follows that, if we include prior-date temperatures in our regressions, we ought to observe marginal effects similar to those we observe for order-date temperature. Column 2 of Table 5 contains estimation results when the temperature seven days prior to the order is added to our base specification (using all seven categories). While the marginal effect associated with the temperature seven days prior to the order is negative and statistically significant at the 10 percent level, it is significantly smaller than the marginal effect associated with order-date temperature and does not change the marginal effect associated with order-date temperature. When this specification is estimated separately for each of the seven categories, the marginal effect associated with temperature seven days prior to the order is most often negative but never statistically significant. As a further check, we also include the snowfall variables, including snowfall seven days prior to the order, and the results are much the same (column 3 of Table 5). (Interestingly, in this last specification, the marginal effect associated with the return-date temperature is positive and statistically significant.) These results suggest that our main empirical findings are not driven by learning about the local weather.[36]

Learning about the Status of Current Clothing.—A third possibility is that people are learning about the adequacy of their current clothing. If the recent weather has been cold, it is more likely that a person has used her existing cold-weather clothing, and as a result the person will have a better idea of whether she needs new cold-weather clothing—e.g., a person puts on her hat on the first cold day of the year and learns that it has a hole. If so, then conditional on ordering, recent cold weather is likely to be correlated with a higher expected valuation (need) for the item. However, if the direct effect dominates the indirect effect—as suggested by our results in Tables 2 and 3 and by our later structural estimation—this story would imply that lower order-date temperatures would be associated with a decreased likelihood of return, which is exactly the opposite from what we find. Moreover, much as for learning about the local weather, under learning about the adequacy of one's current clothing, the temperature on the days prior to the order ought to have a similar effect. Hence, the results from Table 5 suggest that our main empirical findings are not driven by learning about one's current clothing.

Current Weather Affects Current Search.—A fourth possibility is that the current weather may affect people's information about available items by influencing their propensity to browse catalogs or search the Internet. For instance, our main comparative static could be explained by the combination of (a) colder temperatures being associated with an increased propensity to browse catalogs or search the Internet (because people are stuck inside); and (b) having increased information about available items being correlated with an increased propensity to return items (perhaps because people order more marginally enticing items). While this explanation is plausible—although again we note that it is not clear that the comparative static should go in this direction—this explanation is not unique to cold-weather items, and so we ought to observe the same result for non-cold-weather items. Because Table 4 indicates that the result does not hold for non-cold-weather items, our main empirical findings do not seem to be driven by current weather affecting information gathering.

Current Weather Affects One's Mood.—A fifth possibility is that the current weather might influence people's mood, and their mood may influence their propensity to order. In particular, a low temperature may make it more likely that a person is depressed, and a depressed individual may be prone to order an item to raise her spirits. If this "extra utility" from the item is no longer present at the return date, then colder order-date temperatures would be associated

[36] It's worth reemphasizing our fundamental inability to distinguish projection bias versus mispredicting the weather. We are interpreting the fact that the order-date temperature has a much larger effect on returns than the temperature seven days prior to reflect that, on the order date, people are biased by today's utility from the item when predicting the future utility from the item. But the same finding could reflect that people are biased by today's temperature when predicting future temperatures.

with an increased likelihood of return, as we have found. But as for the previous story, this effect should be the same for cold-weather items and non-cold-weather items, and hence again the results from Table 4 contradict this story.[37]

Cold Weather Reminds People to Buy.—A sixth possibility is that cold weather reminds people to consider buying the item. Of course, if people are reminded in an idiosyncratic way—so that a colder temperature leads more people to consider buying but does not change the distribution $G(\gamma)$—then nothing changes. If, however, people are reminded in a selective way, then our theoretical conclusions might change. Suppose, for instance, that people with large γ's consider ordering no matter what—perhaps because they really need the item—whereas people with small γ's consider buying only when they are reminded by cold temperatures. Then on cold days, the set of people who consider ordering contains more marginal individuals, and as a result the likelihood of return conditional on ordering might be larger (even if everyone is fully rational). Once again, it is not obvious that the comparative static should go in this direction. Indeed, it seems equally plausible that it is the people with high γ's who are most likely to be reminded by cold weather, because their need for the item will correspond to a higher cost of not having it. In addition, the fact that the order-date results do not change appreciably when household-specific fixed effects are included in the specification makes this explanation less plausible.

Alternative Purchases during Shipping Delay.—A seventh possibility is that the current weather affects behavior during the shipping delay. In particular, after ordering an item, a person might go to a local store and buy a substitute before the item arrives. If cold weather during the shipping delay makes this behavior more likely, then a colder order-date temperature would indeed be associated with an increased likelihood of return. However, this explanation implies that the return-day temperature ought to have a similar effect, and so again our limited evidence that the return-date temperature has the opposite effect casts some doubt on this explanation.

Risk Aversion and Time Preferences.—Finally, it is worth emphasizing that our order-date results cannot be explained by risk aversion or by a preference for immediate gratification (hyperbolic discounting). On the former, the degree of risk aversion is likely to be orthogonal to weather conditions. On the latter, a preference for immediate gratification would not have any effect on order-date decisions, because a person is making decisions that affect only future utility.

IV. Estimating a Structural Model

While the reduced-form results associated with temperature on the day of order are consistent with projection bias, they do not reveal the magnitude of the bias. By adding some distributional and functional-form assumptions to our economic model from Section I, we are able to estimate this magnitude.

Our goal is to estimate a vector of structural parameters $\boldsymbol{\theta}$ (one of which is the degree of projection bias α).[38] To do so, we construct a likelihood function based on the decision to return conditional on ordering. Specifically, our economic model will permit us to derive a parametric expression for the probability of return conditional on ordering as a function of observed data and structural parameters. Formally, let $\mathbf{Z}_i$ denote the vector of observed data for order-item i; let $P(\mathbf{Z}_i, \boldsymbol{\theta})$ denote the derived probability of return conditional on ordering for order-item i; let $y_i = 1$ if order-item i is returned and $y_i = 0$ if order-item i is not returned; and let N denote the number of order-items in our sample. Assuming the decision to return an item is independent

[37] The fact that we do not find strong effects for snowfall or rainfall would seem to provide further support against the information and mood hypotheses, because snow or rain would seem to be a more important determinant of mood or being "stuck inside" than temperature. Indeed, in the finance literature, researchers have focused on how cloud cover might influence stock-market returns (Edward M. Saunders 1993; David Hirshleifer and Tyler Shumway 2003; Mark J. Kamstra, Lisa A. Kramer, and Maurice D. Levi 2003).

[38] The text provides an overview of the estimation procedure and the results. A more complete description of the estimation procedure appears in Appendix B.

across order-items, the log-likelihood is given by

$$\log L(\boldsymbol{\theta}) = \sum_{i=1}^{N}[y_i \log(P(\mathbf{Z}_i,\boldsymbol{\theta})) + (1-y_i)\log(1-P(\mathbf{Z}_i,\boldsymbol{\theta}))].$$

To compute $P(\mathbf{Z}_i,\boldsymbol{\theta})$, we must add some distributional and functional-form assumptions to our economic model from Section I. Recall that the daily utility function is $v(\mu(\mathbf{x}),\gamma,\varepsilon,\omega_d) = [\mu(\mathbf{x})+\gamma+\varepsilon]u(\omega_d)$. For $\mu(\mathbf{x})$, we use a linear specification $\mu(\mathbf{x}) = \mathbf{x}'\boldsymbol{\beta}_\mu$, where $\mathbf{x}$ includes the same explanatory variables used in the reduced-form model except: (a) the weather variables, which enter through $u(\omega_d)$; (b) the price, which enters the model in a specific way; and (c) the number of items in the order, which we think is more relevant for the return cost (see below). Following the theoretical model, we assume that γ and ε are independent random variables with mean zero; we also impose the additional assumption that they are normally distributed with variances of σ_γ^2 and σ_ε^2. Because a person's behavior is invariant to multiplicative transformations of $[\mu(\mathbf{x})+\gamma+\varepsilon]$, σ_γ^2, σ_ε^2, and $\boldsymbol{\beta}_\mu$ cannot be individually identified. Therefore, we impose $\sigma_\gamma^2 = 1$, which is just a scale normalization.

Recall that the instrumental utility $u(\omega_d)$ represents the marginal utility of using an item relative to the next-best clothing item (and is therefore always nonnegative). Our maintained assumption is that the items in our dataset are cold-weather items, and in particular are the coldest-weather items that buyers might use. Hence, buyers will use these items when the temperature is sufficiently cold, and the marginal utility of the item relative to the next-best clothing item gets larger as the temperature gets colder. If the item is a heavy winter coat, the buyer might be indifferent between that winter coat and her fall jacket when the temperature is, say, 40°F, in which case the instrumental utility would be zero. As the temperature falls, however, the winter coat becomes more and more valuable relative to the fall jacket. For our estimation, we use the following functional form:

$$u(T_d) = \beta_T(T_d - \bar{T}) \quad \text{if } T_d \leq \bar{T};$$

$$u(T_d) = 0 \quad \text{if } T_d > \bar{T}.$$

In other words, we assume that people derive utility from the item only when the temperature is below some $\bar{T}$, and moreover that the utility derived from the item is linear in the deviation of the temperature from $\bar{T}$.

We also must address two issues with regard to projection bias. First, a literal application of simple projection bias would imply that a person might believe that she will get utility from an item on a day when she can confidently predict that she won't use the item. For instance, suppose a person uses an item only when the temperature is colder than 40°F (i.e., $\bar{T} = 40$°F), and consider her prediction for the value of the item on a 80°F day. A literal application of simple projection bias makes the implausible prediction that, if she makes this prediction on, say, a 20°F day, she would predict positive utility from the item on the 80°F day.[39] To eliminate such effects, we assume that the person correctly predicts *when* she would use the item, and mispredicts her utility only for those days. Formally, we assume that if there were no uncertainty about the future temperature T_d, and if the current temperature were T_t, then her perception would be

$$\tilde{u}(T_d|T_t) = (1-\alpha)\beta_T(T_d - \bar{T}) + \alpha\beta_T(T_t - \bar{T}) \quad \text{if } T_d \leq \bar{T} \text{ and } T_t \leq \bar{T};$$

$$\tilde{u}(T_d|T_t) = (1-\alpha)\beta_T(T_d - \bar{T}) \quad \text{if } T_d \leq \bar{T} \text{ and } T_t > \bar{T};$$

$$\tilde{u}(T_d|T_t) = 0 \quad \text{if } T_d > \bar{T}.$$

The second issue with regard to projection bias is our inability to precisely identify the day on which a person decides whether to return an item. The combination of assuming projection bias in the return decision and imposing an incorrect return date could potentially cause a significant bias in our estimation. To avoid this source of bias, we estimate the model under the assumption that the return decision is not affected by projection bias—formally, we assume the return-date cutoff is the fully rational cutoff $\bar{\Lambda}$ rather than projection-bias cutoff

[39] Formally, she would predict utility equal to $[\mu(\mathbf{x}) + \gamma + \varepsilon]\alpha u(20°\text{F})$.

$\tilde{\Lambda}(T_R)$. As we discuss in Section III, this assumption might even be accurate to the extent that a person tries out an item on multiple occasions and makes the return decision based on some cumulative impression of the item.

For the return cost c, we exogenously impose that the cost of returning a single item is \$4 (approximately the cost of postage). But we also allow that there might be economies of scale in returning items, and so the return cost might be decreasing in the number of items in the order. Specifically, we use a nonlinear specification $c = 4\exp[\beta_c((\text{items in order}) - 1)]$, which implies for $\beta_c < 0$ that the return cost decreases toward zero as there are more items in the order. Finally, for computational purposes (explained in Appendix B), we must exogenously specify some of the parameters. We assume that the number of days that an item lasts is $M = 1{,}825$ (i.e., five years); that the daily discount factor is $\delta = 0.9997$ (i.e., an annual discount factor of 0.9); and that utility is derived from the items only when the temperature is below $\bar{T} = 40$°F.

Under these specifications, it is straightforward to compute $P(\mathbf{Z}_i, \boldsymbol{\theta})$ and then numerically obtain the MLE estimator of $\boldsymbol{\theta}$—the details are given in Appendix B. Because our model has a large number of parameters to estimate (primarily because of the fixed effects) it takes some work to get the numerical optimization routine to converge. And, in many cases when convergence was obtained, the Hessian matrix was singular and standard errors could not be calculated. This is likely due to the large dimension of the parameter space and/or lack of sufficient curvature in the likelihood function. We found that by restricting the dataset to reasonable subsamples, we could obtain MLE estimates for which standard errors could be computed (at least for the key parameters of interest). In one subset, we dropped all items ordered from a zip code in which the temperature drops below 40°F for fewer than 30 days per year (from 1990 through 1994). In an effort to reduce the parameter space without significantly reducing our sample size, for vests and jackets (winter boots and parkas/coats), we dropped all items for which the total number of orders is fewer than 500 (1,000) and we dropped all orders in a specific month-region when the total number of orders for that month-region is fewer than 500 (1,000). For the jackets, parkas/coats, and vests categories, these restrictions enabled us to obtain MLE estimates with standard errors (at least for the key parameters of interest). For the hats and winter boot categories, we were able to obtain MLE estimates with standard errors (at least for the key parameters of interest) only when we dropped all orders made in February through August.[40] For the gloves/mittens and sports equipment categories, even with these data restrictions (and several others we tried), we were unable to obtain MLE estimates with standard errors.[41]

The estimates of the structural parameters are given in Table 6.[42] First, note that the coefficient estimates for β_T are negative and significant. These results provide further support for our maintained hypothesis that we are studying cold-weather items. With this confirmed, our primary interest is the estimates for the projection bias parameter α. Except for the vests category, the coefficient estimates for α range between 0.31 and 0.50, and these estimates are quite precise. The null hypothesis that there is no projection bias ($\alpha = 0$) is rejected for these categories. For the vests category, in contrast, the coefficient estimate for α is very close to zero. One possible explanation is that vests are less likely than items in the other categories to be the coldest-weather items that buyers might use. Indeed, the estimate for β_T is much smaller for vests than for the other categories. With the exception of the vests category, the estimates of α suggest that people's predictions of future tastes are roughly one-third

[40] For almost all categories, the point estimates for α are similar even if we do not drop orders where the total number of orders for that item and in that month-region are minimal and do not drop February-to-August orders. Dropping these select orders allows us to obtain standard errors for our key parameters—most notably α.

[41] The number of individual items in each of these two categories is quite large and hence there is a significant number of item-fixed-effect parameters to estimate. We suspect the reason we were unable to obtain standard errors for the gloves/mittens and sports equipment categories is because of the higher dimensionality of the parameter space.

[42] For computational reasons discussed in Appendix B, σ_ε and α are estimated via reparameterizations that restrict σ_ε to be positive and α to be between -1 and 2. Table 6 presents the final estimates for σ_ε and α, where the standard errors are computed using the delta method.

Table 6—Estimates of Structural Parameters

	Winter boots	Hats	Parkas / coats	Vests	Jackets
Determinants of individual-specific tastes ($\boldsymbol{\beta}_\mu$):					
Days between order and shipment	0.0049	0.0096**	0.0042	−0.0070	−0.0010
	(0.0045)	(0.0014)	(0.0027)	(0.0074)	(0.0053)
Days between order and receipt	−0.0002	−0.0016	−0.0001	0.0119*	0.0041
	(0.0043)	(.)	(0.0026)	(0.0071)	(0.0051)
Ordered through Internet	0.0792**	0.0775**	0.0807**	0.0387	0.0418
	(0.0303)	(0.0246)	(0.0151)	(0.0495)	(0.0265)
Ordered by female	−0.0686**	−0.0635	−0.0536**	−0.0117	0.0127
	(0.0099)	(.)	(0.0058)	(0.0170)	(0.0114)
First-time buyer	−0.0994**	−0.0794	0.0221**	−0.1562**	−0.0456**
	(0.0123)	(.)	(0.0068)	(0.0209)	(0.0148)
Percent of items buyer returns	−1.5532**	−1.5095	−0.3030**	−2.1922**	−1.4485**
	(0.0086)	(.)	(0.0011)	(0.0136)	(0.0032)
Number of buyer's prior purchases	−0.0012**	−0.0014	−0.0009**	−0.0003	−0.0003**
	(0.0001)	(.)	(0.0001)	(0.0002)	(0.0001)
Item purchased with credit card	−0.2389**	−0.2311	−0.0270	−0.1301**	0.0248
	(0.0330)	(.)	(0.0214)	(0.0487)	(0.0358)
Item, clothing type, month-region, year-region FEs	YES	YES	YES	YES	YES
β_c	−0.2199**	0.0595	−1.0625	0.0010	−1.1877
	(0.0904)	(.)	(.)	(0.0117)	(.)
β_T	−1.5147**	−0.1196**	−0.7776**	−0.0447**	−2.2398**
	(0.4562)	(0.0003)	(0.0020)	(0.0044)	(0.0004)
σ_ε	0.3483**	0.8310**	0.6010	0.3877**	0.5859**
	(0.0003)	(0.000001)	(.)	(0.0065)	(0.00003)
α	0.3084**	0.4698**	0.3814**	0.0002	0.4992**
	(0.0570)	(0.00001)	(0.0352)	(0.0056)	(0.0002)
Log likelihood	−81,235	−132,585	−231,559	−43,842	−53,774
Observations	202,323	415,976	446,541	119,407	119,043

Notes: Table presents MLE estimates for the parameters of the structural model (see Appendix B for details). Standard errors are in parentheses; (.) indicates missing standard errors.

* Significant at the 10 percent level.

** Significant at the 5 percent level.

to one-half the way between actual future tastes and current tastes.[43]

Finally, consider the estimates for $\boldsymbol{\beta}_\mu$ and β_c. The $\boldsymbol{\beta}_\mu$ estimates indicate that for most categories tastes are stronger for individuals who order through the Internet, are male, are repeat buyers with low return rates on prior purchases, and did not pay with a credit card. Moreover, the fact that all of these coefficients have the opposite sign as the corresponding coefficient in the reduced-form regression suggests that increasing $\mu(\mathbf{x})$ decreases $\Pr[\text{return}|\text{order}]$ (which in turn suggests that the direct return-date effect indeed dominates the indirect order-date effect). The β_c estimates provide minimal support for economies of scale associated with returning items.

V. Discussion

Our analysis provides strong empirical evidence that people experience projection bias with regard to the weather when they purchase cold-weather apparel and gear. In this section, we conclude by discussing some limitations of our analysis and some broader implications.

[43] Because the life expectancy of items is likely to vary across individuals and items, we test the robustness of these results to changes in the number of days an item lasts—specifically, we assume that items last for two years or 1.5 years. For vests, the estimates of α are not significantly different from zero in all specifications. For winter boots, hats, and parkas/coats, the estimates of α do change some, but they remain in roughly the same range (0.3 to 0.55). For jackets, the estimates of α change more substantially, and are on the order of 0.1.

A major limitation of our analysis is that we abstract away from the determinants of when people make the order decision. In other words, whereas we treat the order date as exogenous, it clearly is not—as is implicit in some of our discussions of alternative explanations. In terms of our main results, we are not overly concerned about this limitation because of our focus on the likelihood of return conditional on ordering (and in Section III we are able to rule out several order-date selection stories that might account for our results). Even so, to fully understand the implications of projection bias for purchase decisions, we would want to account for how projection bias influences the order date. Indeed, Loewenstein, O'Donoghue, and Rabin (2003) emphasize how, when a person has many opportunities to buy, all it takes is one high-valuation day to ensure overbuying. Applied here, all it takes is one cold day to induce the person to buy the jacket. We leave such issues for future research.

A second limitation of our analysis is that our empirical results suggest significant differences across categories in terms of the magnitude of the projection bias. These differences raise the question of how to think about people having different degrees of projection bias for different types of decisions. On one hand, it sounds quite likely that people are better at understanding certain types of taste changes than they are at understanding other types of taste changes. At the same time, it would be unappealing to require an estimate of the degree of projection bias for every new domain to which it is applied. It is hoped that, as more evidence of projection bias is assembled, we will discover systematic patterns for when people are more or less prone to experience projection bias.

A closely related question is what is the external validity of our analysis. Have we identified something unique to catalog orders of winter clothing items, or have we identified something that applies more generally? As we have discussed throughout, our empirical analysis cannot distinguish whether we are seeing evidence of projection bias or whether we are seeing evidence that people mispredict their local weather. Even so, given that our evidence confirms the experimental evidence reviewed by Loewenstein, O'Donoghue, and Rabin (2003), we believe it is reasonable to suggest that projection bias is playing a major role, and thus is likely to operate more generally in economics. It is worth investigating further—theoretically but especially empirically—whether projection bias is playing a significant role in consumption of addictive goods, saving-consumption decisions, labor-market decisions, and so forth.

Projection bias is potentially quite important for economics. It is, of course, quite challenging to find a convincing identification strategy to adequately test for projection bias in complicated economic environments, and even more challenging to quantify the impact of projection bias in such environments. By taking the first step of analyzing the relatively simple environment of catalog orders, we hope this paper has laid the groundwork for further empirical analyses of projection bias in these more complicated economic environments.

Appendix A: Derivations and Proofs

Derivation of Order-Date Cutoff for Rational People.—The claim is that rational people will order the item when

$$\gamma \geq \bar{\Lambda} - \bar{\varepsilon}(c/\Psi_R) \equiv \bar{\gamma},$$

where $\bar{\varepsilon}(a)$ is the $\bar{\varepsilon}$ such that $\Pr(\varepsilon \geq \bar{\varepsilon})[\mathrm{E}(\varepsilon|\varepsilon \geq \bar{\varepsilon}) - \bar{\varepsilon}] = a$.

PROOF:

Defining $\hat{\varepsilon}(\gamma) \equiv \bar{\Lambda} - \gamma$, it follows from the text that the person will order when

$$\Pr[\varepsilon \geq \hat{\varepsilon}(\gamma)][(\mu(\mathbf{x}) + \gamma + \mathrm{E}(\varepsilon|\varepsilon \geq \hat{\varepsilon}(\gamma)))\Psi_R - p] + \Pr[\varepsilon < \hat{\varepsilon}(\gamma)][-c] \geq 0.$$

Because $\bar{\Lambda} \equiv [(p - c)/\Psi_R] - \mu(\mathbf{x})$ implies $[\mu(\mathbf{x}) + \gamma]\Psi_R - p = -c - [\bar{\Lambda} - \gamma]\Psi_R = -c - \hat{\varepsilon}(\gamma)\Psi_R$, we can rewrite this condition as

$$-c + \Pr[\varepsilon \geq \hat{\varepsilon}(\gamma)][\mathrm{E}(\varepsilon|\varepsilon \geq \hat{\varepsilon}(\gamma)) - \hat{\varepsilon}(\gamma)]\Psi_R \geq 0.$$

For any ε', $\Pr[\varepsilon \geq \varepsilon'][\mathrm{E}(\varepsilon|\varepsilon \geq \varepsilon') - \varepsilon'] = \int_{\varepsilon'}^{\infty}(\varepsilon - \varepsilon')f(\varepsilon)\ d\varepsilon$, which is decreasing in ε' (f is the pdf of ε). Hence, the left-hand side is decreasing in $\hat{\varepsilon}(\gamma)$, and because $\hat{\varepsilon}(\gamma)$ is decreasing in γ, the left-hand side is increasing in γ. It follows that there exists a cutoff $\bar{\gamma}$ such that the person orders if and only if $\gamma \geq \bar{\gamma}$. Moreover, this condition will hold with equality when $\gamma = \bar{\gamma}$. Hence, the definition of $\bar{\varepsilon}(a)$ in the claim implies $\hat{\varepsilon}(\bar{\gamma}) = \bar{\varepsilon}(c/\Psi_R)$, and given that $\hat{\varepsilon}(\bar{\gamma}) \equiv \bar{\Lambda} - \bar{\gamma}$, the formula for $\bar{\gamma}$ follows.

Derivation of Order-Date Cutoff for People with Projection Bias.—The claim is that people with projection bias will order the item when

$$\gamma \geq \tilde{\Lambda}(\omega_O) - \bar{\varepsilon}(c/\tilde{\Psi}_R(\omega_O)) \equiv \tilde{\gamma}(\omega_O),$$

where $\bar{\varepsilon}(a)$ is again the $\bar{\varepsilon}$ such that $\Pr(\varepsilon \geq \bar{\varepsilon})[\mathrm{E}(\varepsilon|\varepsilon \geq \bar{\varepsilon}) - \bar{\varepsilon}] = a$.

PROOF:

The proof is analogous to that for rational people. Defining $\tilde{\varepsilon}(\gamma, \omega) \equiv \tilde{\Lambda}(\omega) - \gamma$, the person will order when

$$\Pr[\varepsilon \geq \tilde{\varepsilon}(\gamma, \omega_O)][(\mu(\mathbf{x}) + \gamma + \mathrm{E}(\varepsilon|\varepsilon \geq \tilde{\varepsilon}(\gamma, \omega_O)))\tilde{\Psi}_R(\omega_O) - p] + \Pr[\varepsilon < \tilde{\varepsilon}(\gamma, \omega_O)][-c] \geq 0,$$

which we can rewrite as

$$-c + \Pr[\varepsilon \geq \tilde{\varepsilon}(\gamma, \omega_O)][\mathrm{E}(\varepsilon|\varepsilon \geq \tilde{\varepsilon}(\gamma, \omega_O)) - \tilde{\varepsilon}(\gamma, \omega_O)]\tilde{\Psi}_R(\omega_O) \geq 0.$$

Because the left-hand side is decreasing in $\tilde{\varepsilon}(\gamma, \omega_O)$, and because $\tilde{\varepsilon}(\gamma, \omega_O)$ is decreasing in γ, the left-hand side is increasing in γ. It follows that there exists a cutoff $\tilde{\gamma}(\omega_O)$ such that the person orders if and only if $\gamma \geq \tilde{\gamma}(\omega_O)$. Moreover, this condition will hold with equality when $\gamma = \tilde{\gamma}(\omega_O)$. Hence, the definition of $\bar{\varepsilon}(a)$ in the proposition implies $\tilde{\varepsilon}(\tilde{\gamma}(\omega_O), \omega_O) = \bar{\varepsilon}(c/\tilde{\Psi}_R(\omega_O))$, and given that $\tilde{\varepsilon}(\tilde{\gamma}(\omega_O), \omega_O) \equiv \tilde{\Lambda}(\omega_O) - \tilde{\gamma}(\omega_O)$, the formula for $\tilde{\gamma}(\omega_O)$ follows.

PROOF OF LEMMA 1:

To simplify notation, we use $\tilde{\gamma} \equiv \tilde{\gamma}(\omega_O)$ and $\tilde{\Lambda} \equiv \tilde{\Lambda}(\omega_R)$. Recall that the cdf for ε is $F(\varepsilon)$ and the cdf for γ is $G(\gamma)$, and that we assume F and G are both continuous, differentiable, and strictly increasing on the real line. Hence, $f(\varepsilon) \equiv dF(\varepsilon)/d\varepsilon > 0$ for all ε, and $g(\gamma) \equiv dG(\gamma)/d\gamma > 0$ for all γ.

Note that

$$\Pr[\text{return}|\text{order}] = \frac{\Pr[\gamma \geq \tilde{\gamma} \text{ and } \gamma + \varepsilon < \tilde{\Lambda}]}{\Pr[\gamma \geq \tilde{\gamma}]} = \frac{\int_{\tilde{\gamma}}^{\infty} F(\tilde{\Lambda} - \gamma)g(\gamma)\ d\gamma}{1 - G(\tilde{\gamma})}.$$

Differentiating with respect to $\tilde{\Lambda}$ yields

$$\frac{d[\Pr[\text{return}|\text{order}]]}{d\tilde{\Lambda}} = \frac{\int_{\tilde{\gamma}}^{\infty} f(\tilde{\Lambda} - \gamma)g(\gamma)\ d\gamma}{1 - G(\tilde{\gamma})} > 0$$

since $f(\tilde{\Lambda} - \gamma) > 0$ for all γ.

Differentiating with respect to $\tilde{\gamma}$ yields

$$\frac{d[\Pr[\text{return}|\text{order}]]}{d\tilde{\gamma}} = \frac{[1 - G(\tilde{\gamma})][-F(\tilde{\Lambda} - \tilde{\gamma})g(\tilde{\gamma})] + [\int_{\tilde{\gamma}}^{\infty} F(\tilde{\Lambda} - \gamma)g(\gamma)\ d\gamma][g(\tilde{\gamma})]}{[1 - G(\tilde{\gamma})]^2}.$$

Because F is increasing, $F(\tilde{\Lambda} - \gamma) \leq F(\tilde{\Lambda} - \tilde{\gamma})$ for all $\gamma \geq \tilde{\gamma}$, and therefore $[\int_{\tilde{\gamma}}^{\infty} F(\tilde{\Lambda} - \gamma)g(\gamma)\ d\gamma][g(\tilde{\gamma})] \leq [\int_{\tilde{\gamma}}^{\infty} F(\tilde{\Lambda} - \tilde{\gamma})g(\gamma)\ d\gamma][g(\tilde{\gamma})] = F(\tilde{\Lambda} - \tilde{\gamma})[1 - G(\tilde{\gamma})][g(\tilde{\gamma})]$. It follows that $d[\Pr[\text{return}|\text{order}]]/d\tilde{\gamma} < 0$.

PROOF OF PROPOSITION 1:

Note that

$$\tilde{\Lambda}(\omega_R) = \frac{p - c}{(1 - \alpha)\Psi_R + \alpha mu(\omega_R)} - \mu(\mathbf{x})$$

and

$$\tilde{\gamma}(\omega_O) = \frac{p - c}{(1 - \alpha)\Psi_R + \alpha mu(\omega_O)} - \mu(\mathbf{x}) - \bar{\varepsilon}\left(\frac{c}{(1 - \alpha)\Psi_R + \alpha mu(\omega_O)}\right),$$

where $\Psi_R \equiv \sum_{d=R+1}^{R+M} \delta^{d-R} E_{H_d}[u(\omega_d)]$ (which is independent of ω_O and ω_R), $m \equiv \sum_{d=R+1}^{R+M} \delta^{d-R}$, and $\bar{\varepsilon}(a)$ is the $\bar{\varepsilon}$ such that $\Pr(\varepsilon \geq \bar{\varepsilon})[\mathrm{E}(\varepsilon | \varepsilon \geq \bar{\varepsilon}) - \bar{\varepsilon}] = a$.

(i) For $\alpha = 0$, it is clear that $\tilde{\Lambda}(\omega_R)$ and $\tilde{\gamma}(\omega_O)$ are both independent of $u(\omega_R)$ and $u(\omega_O)$, and so Lemma 1 implies that $\Pr[\text{return}|\text{order}]$ is independent of both $u(\omega_R)$ and $u(\omega_O)$.

(ii) For $\alpha > 0$, it is clear that $\tilde{\Lambda}(\omega_R)$ is decreasing in $u(\omega_R)$ (recall that we assume $p > c$) while $\tilde{\gamma}(\omega_O)$ is independent of $u(\omega_R)$, and so Lemma 1 implies that $\Pr[\text{return}|\text{order}]$ is decreasing in $u(\omega_R)$.

(iii) For $\alpha > 0$, it is clear that $\tilde{\Lambda}(\omega_R)$ is independent of $u(\omega_O)$, but $u(\omega_O)$ has two effects on $\tilde{\gamma}(\omega_O)$. Note that $\Pr(\varepsilon \geq \bar{\varepsilon})[\mathrm{E}(\varepsilon | \varepsilon \geq \bar{\varepsilon}) - \bar{\varepsilon}] = a$ can be rewritten as $\int_{\bar{\varepsilon}}^{\infty}(\varepsilon \geq \bar{\varepsilon})f(\varepsilon)\ d\varepsilon = a$, from which it follows that $d\bar{\varepsilon}/da = -1/[1 - F(\bar{\varepsilon})]$. Hence,

$$\frac{d\tilde{\gamma}(\omega_O)}{du(\omega_O)} = \frac{-(\alpha m)(p - c)}{[(1 - \alpha)\Psi_R + \alpha mu(\omega_O)]^2} - \frac{-1}{1 - F(\bar{\varepsilon})}\left(\frac{-(\alpha m)c}{[(1 - \alpha)\Psi_R + \alpha mu(\omega_O)]^2}\right) < 0.$$

It then follows from Lemma 1 that $\Pr[\text{return}|\text{order}]$ is increasing in $u(\omega_O)$.

PROOF OF PROPOSITION 2:

The results follow from the following derivatives and Lemma 1 (again recall that $p > c$).

$$\frac{d\tilde{\Lambda}(\omega_R)}{dc} = \frac{-1}{(1 - \alpha)\Psi_R + \alpha mu(\omega_R)} < 0$$

$$\frac{d\tilde{\Lambda}(\omega_R)}{d\Psi_R} = \frac{-(1 - \alpha)(p - c)}{[(1 - \alpha)\Psi_R + \alpha mu(\omega_R)]^2} < 0 \begin{cases} < 0 \text{ if } \alpha < 1 \\ > 0 \text{ if } \alpha > 1 \end{cases}$$

$$\frac{d\tilde{\Lambda}(\omega_R)}{d\mu(\mathbf{x})} = -1 < 0$$

$$\frac{d\tilde{\Lambda}(\omega_R)}{dp} = \frac{1}{(1 - \alpha)\Psi_R + \alpha mu(\omega_R)} > 0$$

$$\frac{d\tilde{\gamma}(\omega_O)}{dc} = \frac{F(\bar{\varepsilon})}{1 - F(\bar{\varepsilon})} \frac{1}{[(1 - \alpha)\Psi_R + \alpha mu(\omega_O)]} > 0$$

$$\frac{d\tilde{\gamma}(\omega_O)}{d\Psi_R} = \frac{-(1 - \alpha)(p - c)}{[(1 - \alpha)\Psi_R + \alpha mu(\omega_O)]^2} + \frac{1}{1 - F(\bar{\varepsilon})}\left(\frac{-(1 - \alpha)c}{[(1 - \alpha)\Psi_R + \alpha mu(\omega_O)]^2}\right) \begin{cases} < 0 \text{ if } \alpha < 1 \\ > 0 \text{ if } \alpha > 1 \end{cases}$$

$$\frac{d\tilde{\gamma}(\omega_O)}{d\mu(\mathbf{x})} = -1 < 0$$

$$\frac{d\tilde{\gamma}(\omega_O)}{dp} = \frac{1}{(1-\alpha)\Psi_R + \alpha m u(\omega_O)} > 0$$

Appendix B: Details of the Structural Estimation

In this appendix, we provide the details for how to compute $P(\mathbf{Z}_i, \boldsymbol{\theta})$, and we describe some technical details from the estimation.

Given our assumption that the return-date decision is not affected by projection bias while the order-date decision is, the return-date cutoff will be $\bar{\Lambda}$ and the order-date cutoff will be $\tilde{\gamma}(T_O)$ (each of these is defined in Section I), and therefore

$$P(\mathbf{Z}_i, \boldsymbol{\theta}) = \frac{\Pr(\gamma \geq \tilde{\gamma}(T_O) \cap \gamma + \varepsilon < \bar{\Lambda})}{\Pr(\gamma \geq \tilde{\gamma}(T_O))}.$$

Hence, to compute $P(\mathbf{Z}_i, \boldsymbol{\theta})$, we must first compute $\bar{\Lambda}$ and $\tilde{\gamma}(T_O)$.

From Section I, $\bar{\Lambda} \equiv [(p-c)/\Psi_R] - \mu(\mathbf{x})$. The price p is observed data. As mentioned in the text, the specific functional form that we use for the return cost is $c = 4 * exp[\beta_c((\text{items in order}) - 1)]$ and for $\mu(\mathbf{x})$ is $\mu(\mathbf{x}) = exp(\mathbf{x}'\boldsymbol{\beta}_\mu)$. The variables included in $\mathbf{x}$ are listed in Table 6. It remains to compute Ψ_R. Using the functional form for $u(T_d)$ given in Section IV,

$$\Psi_R \equiv \sum_{d=R+1}^{R+M} \delta^{d-R} E_{H_d}[u(T_d)] = \beta_T \psi_R,$$

where β_T is a parameter to be estimated and

$$\psi_R \equiv \sum_{d=R+1}^{R+M} \delta^{d-R} \Pr_{H_d}(T_d \leq \bar{T}) E_{H_d}(T_d - \bar{T} \mid T_d \leq \bar{T}).$$

To compute ψ_R, we first compute $\Pr_{H_d}(T_d \leq \bar{T}) E_{H_d}(T_d - \bar{T} \mid T_d \leq \bar{T})$ using historical weather data—specifically, we use the actual temperatures for a seven-day window around the date for the years 1990 to 1994. Consider, for instance, how to compute $\Pr_{H_d}(T_d \leq \bar{T}) E_{H_d}(T_d - \bar{T} \mid T_d \leq \bar{T})$ for January 4 in zip code 55403. We first construct the empirical temperature distribution by using the temperature observations in zip code 55403 for January 1–7 of the years 1990 to 1994 (for a total of 35 observations). Once we have this temperature distribution, it is simple to compute $\Pr_{H_d}(T_d \leq \bar{T}) E_{H_d}(T_d - \bar{T} \mid T_d \leq \bar{T})$. In principle, M, δ, and $\bar{T}$ are parameters to be estimated; however, because the computation of ψ_R is burdensome (particularly with several hundred thousand observations), we exogenously impose these parameters so that ψ_R need not be computed at every iteration of the numerical optimization algorithm (it is computed just once). As we discuss in the text, we use $M = 1825$, $\delta = 0.9997$, and $\bar{T} = 40$°F.

Also from Section I, $\tilde{\gamma}(T_O) \equiv \tilde{\Lambda}(T_O) - \bar{\varepsilon}(c/\tilde{\Psi}_R(T_O))$ where $\tilde{\Lambda}(T_O) \equiv [(p - c/\tilde{\Psi}_R(T_O)] - \mu(\mathbf{x})$. We have already described the computation of p, c, and $\mu(\mathbf{x})$. Using the functional form for $\tilde{u}(T_d|T_O)$ given in Section V,

$$\tilde{\Psi}_R(T_O) \equiv \sum_{d=R+1}^{R+M} \delta^{d-R} E_{H_d}[\tilde{u}(T_d|T_O)] = \beta_T[(1-\alpha)\psi_R + \alpha(T_O - \bar{T})\Delta_R] \quad \text{if } T_O \leq \bar{T};$$

$$\tilde{\Psi}_R(T_O) \equiv \sum_{d=R+1}^{R+M} \delta^{d-R} E_{H_d}[\tilde{u}(T_d|T_O)] = \beta_T[(1-\alpha)\psi_R] \quad \text{if } T_O > \bar{T},$$

where β_T and ψ_R are as above, α is a parameter to be estimated, T_O is observed data, and

$$\Delta_R \equiv \sum_{d=R+1}^{R+M} \delta^{d-R} \Pr_{H_d}(T_d \leq \bar{T}).$$

Given the exogenously specified M, δ, and $\bar{T}$, we compute Δ_R in the same way that we computed ψ_R (and again we compute Δ_R only once). Finally, we need to compute $\bar{\varepsilon}\left(c/\tilde{\Psi}_R(T_O)\right)$. Recall that $\bar{\varepsilon}(a)$ is the $\bar{\varepsilon}$ such that $\Pr(\varepsilon \geq \bar{\varepsilon})[\mathrm{E}(\varepsilon|\varepsilon \geq \bar{\varepsilon}) - \bar{\varepsilon}] = a$. While there is no closed-form solution for the function $\bar{\varepsilon}(a)$, in Appendix C we describe how numerical methods were used to obtain the following accurate approximation:

$$\bar{\varepsilon}(a) = -a + \sigma_\varepsilon \exp\Big[0.93367604 + 0.00001128757\,(a/\sigma_\varepsilon)^{-1} - 2.9422294\sqrt{a/\sigma_\varepsilon} + 0.18647818\,(a/\sigma_\varepsilon) - 0.47321191(a/\sigma_\varepsilon)^2\Big].$$

Recall that σ_ε^2 is the variance of ε, which is a parameter to be estimated.

After computing the cutoffs $\bar{\Lambda}$ and $\tilde{\gamma}(T_O)$, we must numerically compute

$$P(\mathbf{Z}_i, \theta) = \frac{\Pr(\gamma \geq \tilde{\gamma}(T_O) \cap \gamma + \varepsilon < \bar{\Lambda})}{\Pr(\gamma \geq \tilde{\gamma}(T_O))}.$$

To do so, we define new random variables

$$\eta = \frac{\varepsilon + \gamma}{\sqrt{\sigma_\varepsilon^2 + 1}} \text{ and } \nu \equiv -\gamma.$$

Given that $\gamma \sim N(0,1)$, $\varepsilon \sim N(0,\sigma_\varepsilon^2)$, and ε and γ uncorrelated, it is easy to show that η and ν are standard normal random variables with correlation $-1/\sqrt{\sigma_\varepsilon^2 + 1}$. Because $\gamma \geq \tilde{\gamma}(T_O)$ is equivalent to $\nu < -\tilde{\gamma}(T_O)$, it follows that

$$\Pr(\gamma \geq \tilde{\gamma}(T_O)) = \Pr(\nu < -\tilde{\gamma}(T_O)) = \Phi(-\tilde{\gamma}(T_O)),$$

where $\Phi(x)$ denotes the cdf of a standard normal random variable. Because in addition $\varepsilon + \gamma < \bar{\Lambda}$ is equivalent to $\eta < \bar{\Lambda}/\sqrt{\sigma_\varepsilon^2 + 1}$, it follows that

$$\begin{aligned}\Pr(\gamma \geq \tilde{\gamma}(T_O) \cap \gamma + \varepsilon < \bar{\Lambda}) &= \Pr\left(\nu < -\tilde{\gamma}(T_O) \cap \eta < \bar{\Lambda}/\sqrt{\sigma_\varepsilon^2 + 1}\right) \\ &= \Phi_2\left(-\tilde{\gamma}(T_O), \bar{\Lambda}/\sqrt{\sigma_\varepsilon^2 + 1}; -1/\sqrt{\sigma_\varepsilon^2 + 1}\right),\end{aligned}$$

where $\Phi_2(x, y; \rho)$ denotes the cdf of a standard bivariate normal random variable with correlation ρ. Since most statistical packages (including STATA) have the functions Φ_1 and Φ_2 built in, we can numerically compute

$$P(\mathbf{Z}_i, \theta) = \frac{\Phi_2\left(-\tilde{\gamma}(T_O), \bar{\Lambda}/\sqrt{\sigma_\varepsilon^2 + 1}; -1/\sqrt{\sigma_\varepsilon^2 + 1}\right)}{\Phi(-\tilde{\gamma}(T_O))}.$$

Finally, we note two technical details behind the estimation. First, we constrain σ_ε^2 to be positive through the reparameterization $\sigma_\varepsilon^2 = exp(\beta_\varepsilon)$. Given the MLE estimate of β_ε, the MLE estimate of σ_ε^2 is easily obtained, and its standard error can be computed using the delta method. Second, to facilitate numerical convergence, we constrain α to lie in the compact interval $[-1, 2]$, where the endpoints are intentionally chosen to lie well outside the predicted range for α, which is between zero and one. We use the reparameterization $\alpha = [2\,exp(\beta_\alpha) - 1]/[exp(\beta_\alpha) + 1]$. Given the MLE estimate of β_α, the MLE estimate of α is easily obtained, and its standard error can be computed using the delta method. Table 6 reports the estimates for σ_ε^2 and α (and not the estimates for β_ε and β_α).

Appendix C: Computation of $\bar{\varepsilon}(a)$

Let ε be a normal random variable with mean 0 and variance σ_ε^2. The goal is to calculate $\bar{\varepsilon}(a)$ which is the $\bar{\varepsilon}$ that satisfies

$$P(\varepsilon > \bar{\varepsilon})(\mathrm{E}[\varepsilon | \varepsilon > \bar{\varepsilon}] - \bar{\varepsilon}) = a.$$

The first step it to rewrite this formula in terms of the random variable

$$z \equiv \varepsilon/\sigma_\varepsilon \Leftrightarrow \varepsilon = z\sigma_\varepsilon,$$

and define

$$\bar{z} = \bar{\varepsilon}/\sigma_\varepsilon \Leftrightarrow \bar{\varepsilon} = \bar{z}\sigma_\varepsilon.$$

Plugging in for ε and $\bar{\varepsilon}$ it directly follows that

$$\begin{aligned} \mathrm{E}[\varepsilon | \varepsilon > \bar{\varepsilon}] - \bar{\varepsilon} &= \mathrm{E}[z\sigma_\varepsilon | z\sigma_\varepsilon > \bar{z}\sigma_\varepsilon] - \bar{z}\sigma_\varepsilon \\ &= \mathrm{E}[z\sigma_\varepsilon | z > \bar{z}] - \bar{z}\sigma_\varepsilon \\ &= \sigma_\varepsilon \mathrm{E}[z | z > \bar{z}] - \bar{z}\sigma_\varepsilon \\ &= \sigma_\varepsilon(\mathrm{E}[z | z > \bar{z}] - \bar{z}). \end{aligned}$$

Let $\Phi(z)$ denote the cdf for z. Using the usual standardization trick for normals we have

$$P(\varepsilon > \bar{\varepsilon}) = P(\varepsilon/\sigma_\varepsilon > \bar{\varepsilon}/\sigma_\varepsilon) = P(z > \bar{z}) = 1 - \Phi(\bar{z}).$$

Combining these results gives

$$P(\varepsilon > \bar{\varepsilon})(\mathrm{E}[\varepsilon | \varepsilon > \bar{\varepsilon}] - \bar{\varepsilon}) = (1 - \Phi(\bar{z}))\sigma_\varepsilon(\mathrm{E}[z | z > \bar{z}] - \bar{z}).$$

Therefore, the original formula becomes

$$(1 - \Phi(\bar{z}))\sigma_\varepsilon(\mathrm{E}[z | z > \bar{z}] - \bar{z}) = a$$

or

$$(1 - \Phi(\bar{z}))(\mathrm{E}[z | z > \bar{z}] - \bar{z}) = a/\sigma_\varepsilon.$$

Let $d = a/\sigma_\varepsilon$. Let the inverse of the functional relationship $(1 - \Phi(\bar{z}))(\mathrm{E}[z | z > \bar{z}] - \bar{z}) = d$ be denoted by $h(d) = \bar{z}$. If we know $h(d)$, then given a and σ_ε we can compute $\bar{\varepsilon}$ as

$$\bar{\varepsilon} = \bar{z}\sigma_\varepsilon = h(a/\sigma_\varepsilon)\sigma_\varepsilon.$$

The $h(d)$ function was calculated numerically as follows. First, using the formula for the density of the standard normal, it is simple to show that

$$(1 - \Phi(\bar{z}))(\mathrm{E}[z|z > \bar{z}] - \bar{z}) = (1/\sqrt{2\pi})\, exp(-\bar{z}^2/2) - (1 - \Phi(\bar{z}))\bar{z}.$$

Therefore, we need to find $\bar{z}$ in terms of d according to the formula

$$(1/\sqrt{2\pi})\, exp(-\bar{z}^2/2) - (1 - \Phi(\bar{z}))\bar{z} = d.$$

For a grid of values for $d \in [0.00001, 3.0]$ with the increments of the grid 0.00001, $\bar{z}$ was computed numerically. This generates 300,000 pairs of z, d values. Then, $\log(\bar{z} + d)$ was regressed on 1, $1/d$, $\sqrt{d}, d$ and d^2 and the fitted model (t-stats in parentheses)

$$\underset{(8102)}{0.93367604} + \underset{(206)}{0.00001128757(1/d)} - \underset{(-8094)}{2.9422294\sqrt{d}} + \underset{(632)}{0.18647818d} - \underset{(-10{,}657)}{0.47321191d^2},$$

was obtained. The R^2 for this regression is 0.99999039. Then, a fitted model for $h(d)$ can be calculated as

$$\begin{aligned}\bar{z} &= h(d)\\ &= exp\big(0.93367604 + 0.00001128757(1/d) - 2.9422294\sqrt{d} + 0.18647818d - 0.47321191d^2\big) - d.\end{aligned}$$

The R^2 of this fitted model is 0.99981976. This approximate function is very accurate even for $d > 3$ because the true $h(d)$ function is very close to $-d$ in this range and so is the fitted $h(d)$.

The final, and quite accurate, formula for $\bar{\varepsilon}$ as a function of a and σ_ε is

$$\begin{aligned}\bar{\varepsilon} = -a + \sigma_\varepsilon exp\big[&0.93367604 + 0.00001128757(a/\sigma_\varepsilon)^{-1} - 2.9422294\sqrt{a/\sigma_\varepsilon}\\ &+ 0.18647818(a/\sigma_\varepsilon) - 0.47321191(a/\sigma_\varepsilon)^2\big].\end{aligned}$$

REFERENCES

Angeletos, George-Marios, David Laibson, Andrea Repetto, Jeremy Tobacman, and Stephen Weinberg. 2001. "The Hyberbolic Consumption Model: Calibration, Simulation, and Empirical Evaluation." *Journal of Economic Perspectives*, 15(3): 47–68.

Ariely, Dan, and Klaus Wertenbroch. 2002. "Procrastination, Deadlines, and Performance: Self-Control by Precommitment." *Psychological Science,* 13(3): 219–24.

Camerer, Colin, Linda Babcock, George Loewenstein, and Richard Thaler. 1997. "Labor Supply of New York City Cabdrivers: One Day at a Time." *Quarterly Journal of Economics*, 112(2): 407–41.

DellaVigna, Stefano, and Ulrike Malmendier. 2006. "Paying Not to Go to the Gym." *American Economic Review*, 96(3): 694–719.

DellaVigna, Stefano, and M. Daniele Paserman. 2005. "Job Search and Impatience." *Journal of Labor Economics*, 23(3): 527–88.

Fehr, Ernst, and Lorenz Götte. 2007. "Do Workers Work More If Wages Are High? Evidence from a Randomized Field Experiment." *American Economic Review*, 97(1): 298–317.

Genesove, David, and Christopher Mayer. 2001. "Loss Aversion and Seller Behavior: Evidence from the Housing Market." *Quarterly Journal of Economics*, 116(4): 1233–60.

Hirshleifer, David, and Tyler Shumway. 2003. "Good Day Sunshine: Stock Returns and the Weather." *Journal of Finance*, 58(3): 1009–32.

Kamstra, Mark J., Lisa A. Kramer, and Maurice D. Levi. 2003. "Winter Blues: A Sad Stock Market Cycle." *American Economic Review*, 93(1): 324–43.

Krueger, Joachim, and Russell W. Clement. 1994. "Memory-Based Judgments about Multiple

Categories: A Revision and Extension of Tajfel's Accentuation Theory." *Journal of Personality and Social Psychology*, 67(1): 35–47.

Laibson, David, Andrea Repetto, and Jeremy Tobacman. 2005. "Estimating Discount Functions with Consumption Choices over the Lifecycle." http://www.economics.harvard.edu/faculty/laibson/papers/EstimatingDiscountFunctions16aug2005.pdf.

Loewenstein, George. 1996. "Out of Control: Visceral Influences on Behavior." *Organizational Behavior and Human Decision Processes,* 55(3): 272–92.

Loewenstein, George, Ted O'Donoghue, and Matthew Rabin. 2003. "Projection Bias in Predicting Future Utility." *Quarterly Journal of Economics*, 118(4): 1209–48.

Loewenstein, George, and David Schkade. 1999. "Wouldn't It Be Nice? Predicting Future Feelings." In *Well-Being: The Foundations of Hedonic Psychology*, ed. Daniel Kahneman, Edward Diener, and Norbert Schwarz, 85–105. New York: Russell Sage Foundation.

Saunders, Edward M., Jr. 1993. "Stock Prices and Wall Street Weather." *American Economic Review*, 83(5): 1337–45.

Shui, Haiyan, and Lawrence M. Ausubel. 2004. "Time Inconsistency in the Credit Market." http://papers.ssrn.com/sol3/papers.cfm?abstract_id=586622.

The Directory of Mail Order Catalogs. 2006. New York: Grey House Publishing.