NBER WORKING PAPER SERIES

THE JOINT IDENTIFICATION OF UTILITY AND DISCOUNT FUNCTIONS FROM STATED CHOICE DATA: AN APPLICATION TO DURABLE GOODS ADOPTION

Jean-Pierre H. Dube
Günter J. Hitsch
Pranav Jindal

Working Paper 18393
http://www.nber.org/papers/w18393

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2012

We thank Kilts Center for Marketing, Booth School of Business, University of Chicago, The Initiative on Global Markets, Booth School of Business, University of Chicago Research grant #4-1525 from the Marketing Science Institute The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Joint Identification of Utility and Discount Functions From Stated Choice Data: An Application to Durable Goods Adoption
Jean-Pierre H. Dube, Günter J. Hitsch, and Pranav Jindal
NBER Working Paper No. 18393
September 2012
JEL No. C14,D12,D9,M31

ABSTRACT

We present a survey design that generalizes static conjoint experiments to elicit inter-temporal adoption decisions for durable goods. We show that consumers' utility and discount functions in a dynamic discrete choice model are jointly identified using data generated by this specific design. In contrast, based on revealed preference data, the utility and discount functions are generally not jointly identified even if consumers' expectations are known. The separation of current-period preferences from discounting is necessary to forecast the diffusion of a durable good under alternative marketing strategies. We illustrate the approach using two surveys eliciting Blu-ray player adoption decisions. Both model-free evidence and the estimates based on a dynamic discrete choice model indicate that consumers make forward-looking adoption decisions. In both surveys the average discount rate is 43 percent, corresponding to a substantially higher degree of impatience than the rate implied by aggregate asset returns. The estimates also reveal a large degree of heterogeneity in the discount rates across consumers, but only little evidence for hyperbolic discounting.

Jean-Pierre H. Dube
University of Chicago
Booth School of Business
5807 South Woodlawn Avenue
Chicago, IL 60637
and NBER
jdube@chicagobooth.edu

Pranav Jindal
Penn State
Smeal College of Business
210 Business Building
East Park Avenue
University Park, PA 16802
Pranav.Jindal@PSU.edu

Günter J. Hitsch
University of Chicago
Booth School of Business
5807 South Woodlawn Avenue
Chicago, IL 60637
guenter.hitsch@ChicagoBooth.edu

1 Introduction

Predicting the adoption of a durable good is one of the most important tasks of marketing research. Modeling such adoption decisions is difficult because of the inherent inter-temporal tradeoff between buying now or buying at some future date. This tradeoff arises, for example, because durable goods often become cheaper over time or because the availability of complementary goods increases. Because of this inter-temporal tradeoff, consumers' adoption decisions depend not only on their static preferences among alternative products, but also on the extent to which they discount future utility flows and on their subjective expectations about future market conditions. In this sense, adopting a durable good is a *dynamic decision problem*.

The new product diffusion literature, a key contribution to marketing that originated with the seminal work of Bass (1969), directly addresses the question of when consumers will adopt a new product and attempts to predict the dynamic adoption path. New product diffusion models typically fit the historical sales evolution of a new product well. Diffusion models have also been successfully applied to help a firm predict new product sales over time (Bass, Gordon, Ferguson, and Githens 2001). However, traditional diffusion models are not based on a model of consumer choice that allows for the inherent tradeoff between buying now or in future that we described before. Hence, an analysis of how individual preferences, discount factors, and expectations affect the adoption of a new product is not directly possible. This is not only a conceptual limitation, but more importantly it also limits the applicability of diffusion models to some important marketing tasks. For example, firms need to decide on the initial price level when a product is launched and the subsequent schedule of price changes over time. Without a model that captures both consumers' product preferences and the degree to which they are willing to trade off buying today for buying tomorrow, an evaluation of how different pricing strategies affect the adoption path is generally not possible.

In contrast, a more recent diffusion literature, beginning with the work of Horsky (1990), has recognized the necessity of predicting the aggregate sales evolution from individual consumer decisions. During the last decade, this literature has adopted the dynamic discrete choice approach of Miller (1984), Pakes (1986), and particularly Rust (1987) to model and estimate durable goods demand (e.g. Melnikov 2001, Song and Chintagunta 2003, Nair 2007). Dynamic discrete choice models allow for inter-temporal tradeoffs, and clearly lay out how discounting and expectations affect consumers' decision making. Hence, these models are conceptually attractive and are now widely used in marketing and economics (see the discussions in Bronnenberg, Dubé, Mela, Albuquerque, Erdem, Gordon, Hanssens, Hitsch, Hong, and Sun 2008 and Arcidiacono and Ellickson 2011).

A weakness of dynamic discrete choice models is that they suffer from a fundamental identification problem if estimated from revealed preference data. Using data on discrete choices only, consumers' utility functions, discount factors, and subjective beliefs about future market conditions are not jointly identified (Magnac and Thesmar 2002). Hence, with few exceptions, researchers using field data on consumer choices *assume* a value for the discount factor, typically to reflect some aggregate asset return, and assume that consumers have rational (self-fulfilling) expectations[1]. We consider this approach unsatisfactory and expect that these assumptions are likely to be wrong in many empirical contexts. Numerous studies in psychology and behavioral economics have shown that the rate at which consumers discount the future can vary tremendously across individuals and can differ substantially from the economy-wide asset returns (Frederick, Loewenstein, and O'Donoghue 2002). Also, to our knowledge it is unknown if consumers are able to form mutually-consistent, self-fulfilling beliefs about future market outcomes such as prices in new durable goods markets.

In this paper we propose a new approach that retains the attractive structure of dynamic discrete choice models but allows us to infer consumers' discount functions and utility functions jointly. Our approach is based on stated choice data and, thus, is similar to conjoint analysis. Conjoint analysis, originating with Green and Rao (1971), is another key contribution to marketing and has often been applied to durable goods purchases (see Green and Srinivasan 1990, Green, Krieger, and Wind 2001, and Huber 1997 for recent surveys). However, even though the inter-temporal aspect of durable goods adoption decisions is well-understood in marketing, it has largely been ignored in the conjoint literature. Implicitly, the past conjoint literature has assumed that consumers are myopic. This assumption, if wrong, can lead to severe bias in the preference estimates since non-purchase is entirely ascribed to preferences rather than to beliefs about more favorable future market conditions.

Our approach extends the conjoint literature to account for inter-temporal decision making by forward-looking consumers. Our proposed experimental design elicits product adoption choices conditional on a prediction of future market conditions (states) that are relevant for the consumer's decision. We prove that data sampled from this design non-parametrically identify consumers' discount factors and, more generally, discount functions. Intuitively, we can observe how changes in the market conditions, such as a price cut, impact consumer choices at different points in time. The difference in the reaction to such changes at different dates identifies how future utilities are discounted. By providing the subjects with a forecast of future market conditions we implicitly *control* for their subjective expectations when estimating preferences and discount functions. However, we do not attempt to estimate how the subjects form expectations about the future, which is a topic beyond the scope of this paper. A similar survey-based approach to estimating discount factors has recently been proposed in Viscusi, Huber, and Bell (2008) to estimate the subjects' time preferences for environmental quality. Our papers differ both in the substantive application and in many details of the survey design. In particular, two key contributions of this paper over Viscusi, Huber, and Bell (2008) are the

[1]Two exceptions are Chung, Steenburgh, and Sudhir (2011) and Yao, Mela, Chiang, and Chen (2012) who have tried to devise approaches to estimate discount factors using field data.

formal discussion of the fundamental identification problem in dynamic discrete choice models and the proof of non-parametric identification of discount factors (discount functions) from stated choice data obtained from a specific survey design.

The ability to identify discount factors is a key conceptual advantage over traditional approaches to estimating dynamic discrete choice models using revealed preference data. But our approach also has an important practical advantage for the purpose of forecasting the diffusion of a new product.Similar to conjoint analysis, the approach can be implemented before sales data from a new product are available.

To illustrate the applicability of our survey design we conduct two studies using subjects from a panel maintained by a large U.S. marketing research company. The survey elicits dynamic adoption choices of Blu-ray players under different predictions of future product prices and availability of Blu-ray movies.

A model-free analysis of the raw data provides direct evidence that the subjects systematically change their adoption timing in response to price changes across choice tasks. These changes occur in a manner that is largely consistent with forward-looking behavior. In particular, the subjects tend to delay the adoption of a Blu-ray player if prices decrease in the more distant future, and accelerate the adoption if prices decrease closer to the present period. However, some of the stated choices cannot be easily rationalized, possibly reflecting mistakes made by the survey respondents.

We then separately estimate discount factors and utility functions based on a dynamic discrete choice product adoption model. We obtain three key results. First, the data provide strong evidence for forward-looking behavior. The general dynamic adoption model fits the data better than two alternative models where the subjects are either myopic or do not perceive a trade-off between current and future adoption choices. Second, the population average of the discount factor is 0.7, corresponding to an annual interest rate of 43 percent. In contrast, researchers estimating adoption decisions from field data typically assume much larger discount factors corresponding to aggregate interest rates. Third, we detect substantial heterogeneity in discounting across subjects, in contrast to the standard assumption of homogenous discount factors in the empirical dynamic discrete choice literature.

We also fail to detect strong support for hyperbolic discounting. This finding is consistent with Chevalier and Goolsbee (2005) who also estimate discount factors for a durable good (textbooks). Hence, to the best of our knowledge, empirical evidence for hyperbolic discounting has not been documented for durable goods adoption decisions.

The remainder of the document is organized as follows. In Section 2, we discuss the identification problems inherent to the estimation of dynamic discrete choice models using field data. In Section 3 we propose a survey-based sampling mechanism that allows us to jointly identify discount functions and utilities. Section 4 describes the two conjoint surveys. In Section 5, we provide model-free evidence for forward-looking behavior in the stated choice data. In Section

6 we discuss our estimation approach and the key results on discounting.

2 Dynamic discrete choice adoption models

Dynamic discrete choice model (e.g. Rust 1987) have been widely employed in the recent literature on durable goods demand estimation to predict adoption decisions. We first give a general overview of dynamic discrete choice models and then discuss the topic of identification of the model primitives from field data. We highlight some key identification challenges that have been previously recognized in the econometrics literature (e.g. Rust 1994, Bajari, Chernozhukov, Hong, and Nekipelov 2009 and Abbring 2010). These challenges are less well known in the applied empirical literature and amongst practitioners. For this reason, we provide a formal discussion of the model and its identification below. In Section 3 we propose a practical solution to some of the identification problems.

General structure of dynamic discrete choice models

Below, we present dynamic discrete choice in a very general form as in Rust (1987). But we interpret many model elements in the context of durable goods adoption decisions. The model predicts the decisions of a consumer who decides to adopt one of J products in each period $t = 0, 1, \ldots$ The consumer's choice in period t is denoted by $j \in \mathcal{A} = \{0, \ldots, J\}$, where $j = 0$ (the reference alternative) is the decision to postpone adoption until some future period. In each period, the consumer observes a state vector, $x_t \in \mathbb{X}$, which may include the prices of all products, the availability of complementary goods, or whether a consumer has already made an adoption decision in the past. The consumer also observes a latent utility component, ϵ_{jt}, for each possible choice. The flow utility from choice j in period t is given by $u_j(x_t) + \epsilon_{jt}$. The latent utility components, ϵ_{jt}, are *i.i.d.* with pdf $p(\epsilon)$, where $\epsilon = (\epsilon_0, \ldots, \epsilon_J)$. Following the convention in the literature, we normalize the utility from the reference alternative, $j = 0$, such that $u_0(x_t) \equiv 0$. The consumer believes that the state x_t evolves according to a Markov process with transition density $p(x_{t+1}|x_t, j)$.

The consumer's objective is to make adoption decisions over time $(j(0), j(1), \ldots)$ to maximize the expected present value of her utility $\mathbb{E}\left(\sum_{t=0}^{\infty} \delta^t (u_{j(t)}(x_t) + \epsilon_{j(t)})\right)$. $\delta \in [0, 1)$ is the consumer's discount factor. The adoption decisions, $j(t)$, are functions of the current state and the latent utility components. Under optimal behavior, the value of choosing action j, net of the random utility component, ϵ_{jt}, is given by the choice-specific value function of j,

$$v_j(x_t) = u_j(x_t) + \delta \int \max_{k \in \mathcal{A}} \{v_k(x') + \epsilon_k\} p(\epsilon) p(x'|x_t, j) d\epsilon dx'. \qquad (1)$$

For example, if $j = 0$ this equation states that the value of delaying adoption is given by the current flow utility, $u_0(x_t) = 0$, and the expected present value of making an optimal adoption choice tomorrow given knowledge of the realizations of x_{t+1} and ϵ_{t+1}. To solve the model we define the expected value function:

$$v(x) \equiv \int \max_{k \in \mathcal{A}} \{v_k(x) + \epsilon_k\} p(\epsilon) d\epsilon. \tag{2}$$

$v(x_t)$ is the value that the consumer expects conditional on the state x_t but before observing ϵ_t. Using the definition of the expected value function, the choice-specific value functions can be written in simpler form as

$$v_j(x_t) = u_j(x_t) + \delta \int v(x') p(x'|x_t, j) dx'. \tag{3}$$

Furthermore, the expected value function satisfies the recursive relationship

$$v(x_t) = \int \max_{k \in \mathcal{A}} \{u_k(x_t) + \epsilon_k + \delta \int v(x') p(x'|x_t, k) dx'\} p(\epsilon) d\epsilon. \tag{4}$$

Under mild regularity conditions the right-hand side of the Bellman equation (4) defines a contraction mapping. Thus, there exists a unique value function satisfying equation (4) and correspondingly unique choice-specific value functions satisfying (3).

In summary, the primitives of the dynamic discrete choice model are the utility functions, $u_j(x_t)$, the distribution of latent utility components, $p(\epsilon)$, the consumer's belief about the evolution of the state vector, $p(x_{t+1}|x_t, j)$, and the consumer's discount factor, δ. Given these primitives, the expected value function is uniquely defined by equation (4) and the choice-specific value functions are uniquely defined by equation (3). The model predicts that, under dynamically optimal behavior, the consumer chooses action j, given x_t and ϵ_t, if and only if $v_j(x_t) + \epsilon_{jt} \geq v_k(x_t) + \epsilon_{kt}$ for all $k \neq j$.

In the econometric analysis of dynamic discrete choice models we assume that the researcher observes some or all components of the state vector x_t. However, the latent utility components ϵ_{jt} are not observed to the researcher and serve as the econometric error term to reconcile the model predictions with the data. Hence, the model can only predict the probability but not the exact identity of a choice. We define the probability of decision j, given the state x_t, as

$$\sigma_j(x_t) = \Pr\{v_j(x_t) + \epsilon_{jt} \geq v_k(x_t) + \epsilon_{kt}, \quad \forall k \neq j\}.$$

$\sigma_j(x_t)$ is called the conditional choice probability (CCP) of j. If ϵ_{jt} has the Type I Extreme Value distribution, the conditional choice probabilities are given by the multinomial logit formula

$$\sigma_j(x_t) = \frac{\exp(v_j(x_t))}{\sum_{k=0}^{J} \exp(v_k(x_t))}.$$

We see that the predictions of dynamic discrete choice models are similar to static discrete choice models, with the choice-specific value functions taking the place of the utility functions in a static choice model. However, static and dynamic discrete choice models differ fundamentally in the extent to which the model primitives are identified.

Structural and reduced form of the dynamic discrete choice model

To discuss identification we assume that we can observe the CCPs $\sigma(x) = (\sigma_0(x), \ldots, \sigma_J(x))$ without error for all states $x \in \mathbb{X}$. This assumption embodies an idealized situation where we have access to an arbitrarily large data set that was generated by the dynamic discrete choice model. We do not attempt to infer the distribution of the latent utility terms, $p(\epsilon)$, from the data, but assume that this distribution is known. Our focus is on identification of all other model primitives, including the utility functions for each choice, the discount factor, and the state transition densities. These primitives constitute the structural form of the model, $\mathcal{S} = (u_j(x), \delta, p(x'|x, j))$.[2] We also make the following assumption, which is satisfied by all commonly employed distributions in empirical work, including the Type I Extreme Value and multivariate normal distributions.

Assumption. *The distribution of ϵ has a density $p(\epsilon)$ with respect to the Lebesgue measure on $\mathbb{R}^{J+1}$ and is everywhere strictly positive.*

We first present an important auxiliary result. Define $d_j(x) \equiv v_j(x) - v_0(x)$ for all $j \in \mathcal{A}$. $d_j(x)$ is the difference between the choice-specific values for choice j and the reference alternative 0. Note that $\sigma_j(x)$ is as a function of $d(x) = (d_1(x), \ldots, d_J(x))$ only. Now hold the state x fixed and define the function $\phi(d)$ with components $j \in \mathcal{A}$:

$$\phi_j(d) = \Pr\{d_j + \epsilon_j \geq d_k + \epsilon_k, \quad \forall k \neq j\}. \tag{5}$$

$\phi : \mathbb{R}^J \to \Delta^J$ maps choice-specific value function differences into choice probabilities in the unit simplex $\Delta^J = \{(\xi_0, \ldots, \xi_J) \in \mathbb{R}^{J+1} : \sum_{j=0}^{J} \xi_j = 1 \text{ and } \xi_j \geq 0 \text{ for all } j\}$. Indeed, if $p(\epsilon)$ is positive then all choice probabilities are positive and $\phi(d)$ will be in the interior of Δ^J, denoted by $\text{int}(\Delta^J)$. The following proposition summarizes results by Hotz and Miller (1993) and Norets and Takahashi (2012):

Proposition. *Under the assumptions on the distribution of ϵ stated above, ϕ has an inverse function $\phi^{-1} : int(\Delta^J) \to \mathbb{R}^J$.*

A key implication of this proposition is that we can directly infer the choice-specific value functions differences if we observe the CCPs for all $x \in \mathbb{X}$: $v_j(x) - v_0(x) = \phi_j^{-1}(\sigma(x))$. We call $\mathcal{R} = (v_j(x) - v_0(x))$ the reduced form of the model. If the CCPs are observed without error, as we assume for the discussion of identification, and if the distribution of the latent utility

[2]The notation $u_j(x)$ refers to the function $u : \mathcal{A} \times \mathbb{R} \to \mathbb{R}$, and similarly for $p(x'|x, j)$.

terms is known, then the reduced form of the model $\mathcal{R}$ is non-parametrically identified. In a static discrete choice model, knowledge of $\mathcal{R}$ allows a firm to predict the effect of a marketing action, such as a price change, on consumer demand. In contrast, in a dynamic discrete choice model, most marketing actions of interest cannot be evaluated based only on knowledge of the reduced form $\mathcal{R}$. For example, a firm may systematically change the dynamic pricing policy for a durable good. If consumers anticipate this shift in the pricing strategy their beliefs about future prices, captured by the transition density $p(x'|x,j)$, will change. Since the choice-specific value functions depend on the transition density, the new pricing strategy will also change the reduced form of the model. Thus, demand predictions based on past data will not correctly reflect the change in consumer demand caused by the change in the pricing policy. Instead, the firm needs to know the structural form of the model, $\mathcal{S}$, to predict accurately the effect of the new pricing policy on demand.

Identification of dynamic discrete choice models

The following proposition contains the key identification result.

Proposition 1. *Let the distribution of the random utility components, $p(\epsilon)$, the discount factor, δ, and the consumer's beliefs about the evolution of the state vector, $p(x'|x,j)$, be given. Suppose we observe the CCPs, $\sigma_j(x)$, for all states $x \in \mathbb{X}$ and all choices $j \in \mathcal{A}$. Then:*

(i) We can infer the unique choice-specific value functions, $v_j(x)$, consistent with the dynamic discrete choice model.

(ii) The utilities, $u_j(x)$, are identified for all states $x \in \mathbb{X}$ and all choices $j \in \mathcal{A}$.

Proof. The proof follows Bajari, Chernozhukov, Hong, and Nekipelov (2009). As discussed above, we can infer the choice-specific value function differences directly from the CCPs:

$$d_j(x) = v_j(x) - v_0(x) = \phi_j^{-1}(\sigma(x)). \tag{6}$$

To see why the proposition holds, first re-write the expected value function:

$$\begin{aligned} v(x) &= \int \max_{k \in \mathcal{A}} \{v_k(x) + \epsilon_k\} p(\epsilon) d\epsilon \\ &= \int \max_{k \in \mathcal{A}} \{d_k(x) + \epsilon_k\} p(\epsilon) d\epsilon + v_0(x). \end{aligned} \tag{7}$$

Using equation (7) and the inversion (6), we can express $v_0(x)$ in the following form (remember that $u_0(x) \equiv 0$):

$$\begin{aligned} v_0(x) &= u_0(x) + \delta \int v(x')p(x'|x,0)dx' \\ &= \delta \int \max_{k \in \mathcal{A}} \{d_k(x') + \epsilon_k\} p(\epsilon)p(x'|x,0)d\epsilon dx' + \delta \int v_0(x')p(x'|x,0)dx' \\ &= \delta \int \max_{k \in \mathcal{A}} \{\phi_k^{-1}(\sigma(x')) + \epsilon_k\} p(\epsilon)p(x'|x,0)d\epsilon dx' + \delta \int v_0(x')p(x'|x_t,0)dx' \qquad (8) \end{aligned}$$

Note that the first term on the right-hand side of the equation above is known given $p(\epsilon)$, $p(x'|x,j)$, and the CCPs. It is straightforward to show that (8) satisfies Blackwell's conditions and thus defines a contraction mapping with a unique solution $v_0(x)$. Hence, we can infer $v_0(x)$ along with all the other choice-specific value functions:

$$v_j(x) = \phi_j^{-1}(\sigma(x)) + v_0(x).$$

Given the choice-specific value functions, we can calculate the expected value function (2) and then infer the utility functions from the equation defining the choice-specific value functions:

$$u_j(x) = v_j(x) - \delta \int v(x')p(x'|x,j)dx.$$

□

The proposition says that if we are willing to treat $p(\epsilon)$, δ, and $p(x'|x,j)$ as known, the consumer's utility function is non-parametrically identified. However, the solution to the contraction mapping (8) depends on δ and $p(x'|x,j)$. If we assume a different discount factor, $\tilde{\delta}$, or a different transition density, $\tilde{p}(x'|x,j)$, we will infer a different set of choice-specific value functions, $\tilde{v}_j(x)$, and utilities, $\tilde{u}_j(x)$. The choice-specific value functions $\tilde{v}_j(x)$ also entirely rationalize the observed CCPs. We conclude that without further assumptions, the structural form of the dynamic discrete choice model, $\mathcal{S} = (u_j(x), \delta, p(x'|x,j))$, is not non-parametrically identified.

Assumptions to achieve identification

To overcome the identification problem the dynamic discrete choice literature has typically made two assumptions. The first assumption is *rational expectations*. Under rational expectations, the consumer's subjective beliefs, $p(x'|x,j)$, coincide with the actual evolution of the state vector. Thus, if x is observed, $p(x'|x,j)$ can be inferred directly from the data. However, as we showed above, even if $p(x'|x,j)$ is known, the utility functions and the discount factor are not jointly identified. The extant literature typically also assumes that the discount factor is known, or rather that it can be calibrated to reflect some economy-wide interest rate or asset return, r. For example, based on the consumption CAPM (see for instance the discussion in Cochrane 2001), the risk-free rate r satisfies the following relationship:

$$u'(c_t) = \delta \mathbb{E}_t \left[(1 + r) u'(c_{t+1}) \right].$$

Here, $u'(c)$ is the marginal utility of consumption. For example, at the time of writing, the yield on 3-month U.S. Treasury bills, an essentially risk-free security, is 0.08 percent at an annualized rate. Correspondingly, if consumers expect no change in consumption over the next three months such that $u'(c_t) = \mathbb{E}_t[u'(c_{t+1})]$, the corresponding annual discount factor would be $\delta = \frac{1}{1+0.0008} = 0.9992$. In practice, the identification problem is sufficiently severe that researchers are typically unable to estimate δ even if a specific parametric form for $u_j(x)$ is assumed.

Both assumptions, for the discount factor and for beliefs, are highly problematic. First, countless studies in psychology and behavioral economics cast doubt on the assumption of a uniform discount factor that corresponds to some economy-wide asset return (Frederick, Loewenstein, and O'Donoghue 2002). Second, there is no solid body of empirical research that justifies the rational expectations assumption. A priori, particularly in the case of a new product adoption where consumers have little or no prior experience with the category and little or no access to past data, it is doubtful that consumers would know the exact process by which prices and other components of the state variable x_t evolve.

The problem is particularly severe if the main purpose of the empirical analysis is to make predictions about future product adoptions. Incorrect assumptions about the discount factor and/or beliefs will generate incorrect inferences about the structural form of the model. However, as we highlighted above, an accurate prediction of future demand requires knowledge of the true structural form. In the next section we will present an approach based on stated choice data that allows us to identify jointly the consumer's utility functions and the discount factor. Our approach skirts the problem of belief estimation, which we leave for future research.

As a final comment, the assumption that $p(\epsilon)$ is known is standard in the applied discrete choice literature and, in our view, substantially less severe than the assumptions on the discount factor and beliefs. The inversion theorem above shows that, even for static discrete choice models ($\delta = 0$), the utility functions and the latent utility term distribution are not jointly identified without additional restrictions on either $u_j(x)$ or $p(\epsilon)$.

3 Identification of discount functions using stated choice data

As discussed in the preceding section, discount factors are generally not identified from revealed choice data. We now introduce an approach based on stated choice data that overcomes this identification problem. The approach allows us to identify discount factors non-parametrically and, more generally, discount functions $\rho(t)$ that allow future utilities to be discounted at some arbitrary rate $\rho(t) \in [0,1]$. The approach is based on a sampling mechanism that can be implemented using a survey design that is similar, although slightly more involved, than a typical conjoint design. We first describe this sampling mechanism and then show that discount functions and utilities are identified from the resulting choice data.

Sampling mechanism

We present each subject with a sequence of choice tasks. In each choice task we provide the subject with information about the current and future states over the time horizon $t = 0, 1, \ldots, T$, where $t = 0$ is the present period. We instruct the subjects to take the values of $x_t \in \mathbb{X}$ as given such that they do not face uncertainty over the future states up to and including period T. The subjects state that they will either adopt one of the J products in some period $t \leq T$, or that they will not adopt a product before period $T+1$. Delaying the adoption decision until period $T+1$ or beyond includes the possibility of adopting at time $t > T$ or never adopting at all. We label the decision not to adopt before period $T+1$ as 0, and the decision to adopt product j in period t as (j,t). Hence, the choice set is $\mathcal{A} = \{(1,0),(2,0),\ldots,(J-1,T),(J,T),0\}$ and includes $J \cdot (T+1)+1$ options.

We allow for a general discount function $\rho(t)$, where $\rho(0) = 1$ and $0 \leq \rho(t) \leq 1$ for all t. The case of geometric discounting that we used before is a special case with $\rho(t) = \delta^t$. Hyperbolic discounting with (β,δ)-preferences (e.g. Phelps and Pollak 1968) is another special case, where $\rho(0) = 1$ and $\rho(t) = \beta\delta^t$ for all $t \geq 1$.

The choice-specific value in period 0 from adopting product j in period t is

$$\omega_{jt}(x_t) = \rho(t) u_j(x_t),$$

and the value from choosing the reference alternative 0 is $\omega_0(x_T)$. Here we implicitly assume that the subjects believe that x_t follows a Markov process. Hence, x_T contains all the information to predict the value of delaying the product adoption until some period $t > T$. A key assumption is that $u_j(x)$ is stationary in the sense that the utility from choosing alternative j depends only on the state, but not generally on the period in which the adoption decision is made. Indeed, if the utility function changed over time we would not be able to distinguish discounting from changes in product preferences.

We assume that the stated choice maximizes the subject's utility. The utility from the choice (j,t) is given by $\omega_{jt}(x_t) + \epsilon_{jt}$, where ϵ_{jt} is a latent utility component. ϵ_0 denotes the latent utility component from the reference alternative. Define $\mathcal{X} = \times_{t=0}^{T}\mathbb{X}$. Then for $\boldsymbol{x} \in \mathbb{X}$ the conditional choice probability for choice $(j,t) \in \mathcal{A}$ is

$$\sigma_{jt}(\boldsymbol{x}) = \Pr\{\omega_{jt}(x_t) + \epsilon_{jt} \geq \omega_{ks}(x_s) + \epsilon_{ks} \text{ and } \omega_{jt}(x_t) + \epsilon_{jt} \geq \omega_0(x_T) + \epsilon_0 \quad \forall (k,s) \neq (j,t)\}.$$

The CCP of choosing 0 is defined analogously. The latent utility terms, $\boldsymbol{\epsilon} = (\epsilon_{10}, \ldots, \epsilon_{JT}, \epsilon_0)$, have a distribution with a density $p(\boldsymbol{\epsilon})$ that is measurable with respect to the Lebesgue measure and everywhere positive. As in the previous section, we assume that this distribution is known.

Because the future states up to period T are given to the subjects in our survey design, and because we assume that consumers observe the latent utility terms, we have a model of choice under perfect foresight. This model can be analyzed as a multinomial discrete choice model, and no dynamic programming techniques are needed. We next show that the discount function and the utilities are identified if our survey allows for sufficient variation in the states presented to the subjects.

Identification of the discount function $\rho(t)$

Proposition 2. *Suppose that we observe the CCPs $\sigma(\boldsymbol{x}) = (\sigma_{10}(\boldsymbol{x}), \ldots, \sigma_{JT}(\boldsymbol{x}), \sigma_0(\boldsymbol{x}))$ without error for $\boldsymbol{x} \in \mathcal{B} \subseteq \mathcal{X}$ ($\mathcal{B}$ is a subset of all possible sequences of current and future states that can be presented to the subjects). Assume that $p(\boldsymbol{\epsilon})$ is known. Assume that for period t, $1 \leq t \leq T$, there are two states, $\bar{x}$ and $\bar{x}'$, such that $u_j(\bar{x}) \neq u_j(\bar{x}')$ for some $j \in \{1, \ldots, J\}$. Furthermore, assume that $\mathcal{B}$ contains four elements $\boldsymbol{x}^{(l)}$, $l = 1, \ldots, 4$, such that: (i) $x_T^{(1)} = x_T^{(2)}$ and $x_T^{(3)} = x_T^{(4)}$, (ii) $x_0^{(1)} = x_t^{(3)} = \bar{x}$, and (iii) $x_0^{(2)} = x_t^{(4)} = \bar{x}'$. Then $\rho(t)$ is identified.*

Proof. We can invert the conditional choice probabilities $\sigma(\boldsymbol{x}^{(l)})$ to infer $d_{ks}^{(l)} = \omega_{ks}(x_s^{(l)}) - \omega_0(x_T^{(l)}) = \phi_{ks}^{-1}(\sigma(\boldsymbol{x}^{(l)}))$ for all $k \in \{1, \ldots, J\}$ and $s \leq T$. Note that

$$
\begin{aligned}
d_{j0}^{(1)} - d_{j0}^{(2)} &= (\omega_{j0}(\bar{x}) - \omega_0(x_T^{(1)})) - (\omega_{j0}(\bar{x}') - \omega_0(x_T^{(2)})) = u_j(\bar{x}) - u_j(\bar{x}'), \\
d_{jt}^{(3)} - d_{jt}^{(4)} &= (\omega_{jt}(\bar{x}) - \omega_0(x_T^{(3)})) - (\omega_{jt}(\bar{x}') - \omega_0(x_T^{(4)})) = \rho(t)(u_j(\bar{x}) - u_j(\bar{x}')).
\end{aligned}
$$

Because $u_j(\bar{x}) - u_j(\bar{x}') \neq 0$ it follows that

$$
\rho(t) = \frac{d_{jt}^{(3)} - d_{jt}^{(4)}}{d_{j0}^{(1)} - d_{j0}^{(2)}}.
$$

$\square$

The intuition for why the discount function can be identified is as follows. Our sampling design presents a subject with a choice of product j at two different states $\bar{x}$ and $\bar{x}'$, either now or in some future period $t > 0$. Holding the value from delaying adoption beyond period T constant, we can infer how the value of adoption differs across the two states $\bar{x}$ and $\bar{x}'$. The extent to which the difference in the value of adoption at time $t > 0$ is different (typically smaller) than the difference in the value of adoption in the present period identifies the discount factor $\rho(t)$.

If we are willing to impose linearity on one component of the utility function the conditions for identification can be weakened further. Suppose $x = (\xi, z)$ where $z \in \mathbb{R}^J$, and that the utility function has the form $u_j(\xi, z) = \mu_j(\xi) + \alpha z_j$, $\alpha \neq 0$.

Proposition 3. *Suppose that we observe the CCPs* $\sigma(\boldsymbol{x}) = (\sigma_{10}(\boldsymbol{x}), \ldots, \sigma_{JT}(\boldsymbol{x}), \sigma_0(\boldsymbol{x}))$ *without error for* $\boldsymbol{x} \in \mathcal{B} \subseteq \mathcal{X}$*, and that* $p(\boldsymbol{\epsilon})$ *is known. Assume that* $\mathcal{B}$ *contains four elements* $\boldsymbol{x}^{(l)}$*,* $l = 1, \ldots, 4$*, such that for some* $j \in \{1, \ldots, J\}$*: (i)* $x_T^{(1)} = x_T^{(2)}$ *and* $x_T^{(3)} = x_T^{(4)}$*, (ii)* $z_{j0}^{(1)} \neq z_{j0}^{(2)}$ *and* $z_{jt}^{(3)} \neq z_{jt}^{(4)}$*, (iii)* $\xi_0^{(1)} = \xi_0^{(2)}$ *and* $\xi_t^{(3)} = \xi_t^{(4)}$*. Then* $\rho(t)$ *is identified.*

Proof. As before invert the conditional choice probabilities $\sigma(\boldsymbol{x}^{(l)})$ to infer $d_{ks}^{(l)} = \omega_{ks}(x_s^{(l)}) - \omega_0(x_T^{(l)}) = \phi_{ks}^{-1}(\sigma(\boldsymbol{x}^{(l)}))$ for all $k \in \{1, \ldots, J\}$ and $s \leq T$. Note that

$$\begin{aligned} d_{j0}^{(1)} - d_{j0}^{(2)} &= (\mu_j(\xi_0^{(1)}) + \alpha z_{j0}^{(1)} - \omega_0(x_T^{(1)})) - (\mu_j(\xi_0^{(2)}) + \alpha z_{j0}^{(2)} - \omega_0(x_T^{(2)})) = \alpha(z_{j0}^{(1)} - z_{j0}^{(2)}), \\ d_{jt}^{(3)} - d_{jt}^{(4)} &= (\rho(t)(\mu_j(\xi_t^{(3)}) + \alpha z_{jt}^{(3)}) - \omega_0(x_T^{(3)})) - (\rho(t)(\mu_j(\xi_t^{(4)}) + \alpha z_{jt}^{(4)}) - \omega_0(x_T^{(4)})) \\ &= \rho(t)\alpha(z_{jt}^{(3)} - z_{jt}^{(4)}). \end{aligned}$$

Hence

$$\rho(t) = \frac{z_{jt}^{(3)} - z_{jt}^{(4)}}{z_{j0}^{(1)} - z_{j0}^{(2)}}.$$

□

For example, the empirical literature often assumes that utility is linear in price. Proposition 3 states that $\rho(t)$ is identified as long as we have some price variation in the present period and in period $t > 0$, holding constant both ξ within each period and the value from delaying adoption beyond period T. However, unlike in Proposition 2, the within-period variation in prices need not be the same across periods (Proposition 2 assumes that subjects are presented with a choice at the *same* states $\bar{x}$ and $\bar{x}'$ both in the present period and at $t > 0$).

Identification of the utility functions $u_j(x)$

Proposition 4. *Suppose that we observe the CCPs* $\sigma(\boldsymbol{x}) = (\sigma_{10}(\boldsymbol{x}), \ldots, \sigma_{JT}(\boldsymbol{x}), \sigma_0(\boldsymbol{x}))$ *without error for* $\boldsymbol{x} \in \mathcal{B} \subseteq \mathcal{X}$*, and assume that* $p(\boldsymbol{\epsilon})$ *is known. Furthermore, suppose that there are two periods* $t, t' \leq T$ *such that* $\rho(t) \neq \rho(t')$*, and that the conditions in Proposition 2 (or Proposition 3) are met such that* $\rho(t)$ *and* $\rho(t')$ *are identified. Assume that* $\mathcal{B}$ *contains two elements,* $\boldsymbol{x}^{(1)}$ *and* $\boldsymbol{x}^{(2)}$*, such that (i)* $x_T^{(1)} = x_T^{(2)}$ *and (ii)* $x = x_t^{(1)} = x_{t'}^{(2)}$ *(the same state* x *is presented to the subjects in two different time periods). Then* $u_j(x)$ *is identified for all* $j = 1, \ldots, J$*.*

Proof. Invert $\sigma(\boldsymbol{x}^{(l)})$ to infer $d_{ks}^{(l)} = \omega_{ks}(x_s^{(l)}) - \omega_0(x_T^{(l)}) = \phi_{ks}^{-1}(\sigma(\boldsymbol{x}^{(l)}))$ for all $k \in \{1, \ldots, J\}$ and $s \leq T$. Note that

$$d_{jt}^{(1)} - d_{jt'}^{(2)} = (\omega_{jt}(x_t^{(1)}) - \omega_0(x_T^{(1)})) - (\omega_{jt'}(x_{t'}^{(2)}) - \omega_0(x_T^{(2)})) = (\rho(t) - \rho(t'))u_j(x).$$

Because $\rho(t) - \rho(t') \neq 0$ we have

$$u_j(x) = \frac{d_{jt}^{(1)} - d_{jt'}^{(2)}}{\rho(t) - \rho(t')}.$$

$\square$

This proposition shows that the utility functions are always identified unless the discount function $\rho(t)$ never varies for $t \leq T$, which is equivalent to the extreme situation of no discounting, i.e. $\rho(t) \equiv 1$ for all $t \leq T$.

Generalization

Consider the following generalization of the choice-specific value from adopting product j in period t:

$$\omega_{jt}(\boldsymbol{x}) = \rho(t)u_j(x_t, 0) + \sum_{s=t+1}^{T} \rho(s)u_j(x_s, 1) + \varphi_j(x_T).$$

To motivate this formulation, consider a case where x_t contains information on the product prices and the availability of complementary goods. At the time of adoption, t, the consumer's utility depends both on the price of product j and the available complementary goods, while the utility after adoption, in periods $s > t$, is only affected by the complementary goods. This difference in utility between the adoption period and later is captured by the state $\iota \in \{0, 1\}$ in the utility function, $u_j(x, \iota)$. $\varphi_j(x_T)$ captures the value that the consumer derives from the complementary goods beyond period T.

It is straightforward to modify Propositions 2-4 to show that the discount function $\rho(t)$ and the utility functions $u_j(x, \iota)$ are also identified for this more general model.

Discussion

Propositions 2-4 clearly indicate the variation in the data that is needed to identify the discount function $\rho(t)$ and the utilities $u_j(x)$. This variation is created by the states $\boldsymbol{x} = (x_0, \ldots, x_T) \in \mathcal{B}$ and is entirely under the control of the researcher. Our results also indicate the data requirement for the joint identification of the discount and utility functions. We need to present the subjects only with a small subset of all possible sequences of states $\boldsymbol{x} \in \mathcal{X}$. To see this, suppose that the state space $\mathbb{X}$ is finite and that it contains K elements. Then $|\mathcal{X}| = K^{T+1}$, and the number of possible sequences of states rises exponentially in T. However, Proposition 2 shows that we only need information on the CCPs at $2 \cdot (T+1)$ sequences of states to identify the discount function $\rho(t)$ for $t \leq T$. Proposition 4 shows that we can choose some state $x' \in \mathbb{X}$ and identify the utility functions $u_j(x')$ from the CCPs at two state sequences. We then only need information on the CCPs at $K - 1$ additional state sequences (holding x_T constant) to infer $u_j(x)$ for all states

x. Thus, the discount and utility functions are identified from $2 \cdot (T+1) + K + 1$ CCPs. In particular, if T is small relative to K, the data requirement for identification of the discount and utility functions is of the same order of magnitude as the data requirement for identification of the utilities in a static discrete choice model. This question regarding the data requirement is different from the question of how difficult it is for the subjects to understand and process our survey design compared to a static conjoint design.

The assumption on the latent utility terms is less innocuous. Our model assumes that the subjects observe all current and future ϵ_{jt} terms (and ϵ_0) at the time when the survey is administered, or at least act as if they observed these latent utility terms. Due to this assumption the stated choices can be analyzed as a multinomial choice model with a specific structure on the adoption values due to discounting. An alternative assumption would be that the subjects might face uncertainty over the future latent utility terms and incorporate this uncertainty into their stated choice. Generally, the treatment and interpretation of the econometric error term in a discrete choice model is of great importance, but the assumptions in the extant literature are largely based on internal consistency (Rust 1987, Rust 1994) and convenience, not on prior empirical work establishing how to best account for these error terms. In this sense, the assumption made in this paper is not more arbitrary than the assumptions in prior work. We view an exploration of alternative assumptions on the error term to be of great importance, but beyond the scope of this paper.

Our approach controls for the role of expectations in dynamic decision-making by providing the subjects with a forecast of the sequence of future states. A much more ambitious task would consist of estimating the process by which the subjects form expectations about the future. This is also a topic of great importance that is beyond the scope of this paper.

Alternative approaches to achieve identification of discount factors

Another approach to the identification problem is based on an exclusion restriction. Suppose that the state space $\mathbb{X}$ is discrete, and that there are two states $x_a \neq x_b$ such that $u_j(x_a) = u_j(x_b)$ for all $j \in \mathcal{A}$ but $p(x'|x_a, k) \neq p(x'|x_b, k)$ for some choice k. x_a and x_b provide variation that does not affect current utilities but affects future payoffs through its impact on the state transitions. Fang and Wang (2012) present a theorem stating that, under these assumptions, the discount parameters in a hyperbolic model with (β, δ)-preferences are identified.[3][4] A similar claim is often attributed to Magnac and Thesmar (2002). Their paper, however, provides an exclusion restriction that is based on the "current value function" which is not a model primitive but requires knowledge of the solution of the dynamic discrete choice model and thus limits

[3]The statement in Fang and Wang (2012) is actually somewhat more general and also allows for identification of a third parameter that indicates the extent to which a decision maker anticipates her future present-bias.

[4]Chung, Steenburgh, and Sudhir (2011) apply the same strategy to estimate discount factors for sales agents anticipating future incentive bonuses, although no formal identification proof is provided.

the applicability of the result.

This identification strategy is promising, but in the case of durable goods adoption decisions it will generally be difficult to find a state variable that satisfies the exclusion restriction. States that we typically observe include prices, the availability of complementary goods, and product qualities. None of these variables are generally excluded from the utility functions.

Another approach, proposed in Yao, Mela, Chiang, and Chen (2012), is to utilize data where consumers are observed in two different choice situations. In both situations the consumers have the same preferences. However, in one situation choices are made dynamically, while in the other situation current choices do not affect future payoffs such that static decision making is optimal. The paper presents an example where mobile phone customers were initially on a linear usage plan and then later switched to a three-part tariff. Optimal decision making (placing and accepting calls) is static under the linear plan but dynamic under the three-part tariff. Finite-horizon choice problems provide a similar context because decision-making in the terminal period is static. For the case of continuous controls Yao, Mela, Chiang, and Chen (2012) prove that geometric discount factors are identified from data with this specific structure. No proof is provided for discrete choices, although it seems plausible that the identification argument would also hold in this more general context.

This is also a promising approach, but unlikely to be applicable to the case of durable goods adoption decisions. In particular, the adoption of a durable good always has dynamic consequences, both because current product adoption affects future utility flows and because there is an option value from delaying adoption unless the market is completely "static," in the sense that prices, the available products and qualities, etc., never change over time.

4 Survey and data

Survey design

We designed two surveys to elicit adoption decisions for Blu-ray players. At the time the surveys were conducted, Blu-ray was a nascent technology providing high-definition video. Blu-ray had just won a standards war against the competing HD DVD format and received much attention in the press. Thus, the technology was likely of interest to many survey participants, making this a good context in which to to implement and illustrate the sampling mechanism described in the previous section. The two surveys differ in the complexity of the choice tasks faced by the subjects. Below we provide a summary of the surveys. Additional details are contained in Appendix A.

In several introductory screens, we first give the subjects an overview of the Blu-ray technology and compare the benefits of Blu-ray over regular DVD movies. We then present the subjects with the choice tasks. Figure 1 shows a screen in the first survey. The screen provides the subjects with information about the evolution of prices for a medium quality Blu-ray player from March 2009 (referred to as “Now” on the screen) to December 2011. We ask the subjects if and when they would adopt the Blu-ray player. Subjects can choose the “Will not buy” option which indicates that either they will never buy the Blu-ray player or that they might buy the Blu-ray player after December 2011. On each screen, we remind the subjects of the number of Blu-ray movie titles available in each period. In the first survey, the number of titles remains the same across choice tasks.

In the second survey, the subjects choose among two Blu-ray player brands (Sony and Samsung). The left-hand side of the screen (Figure 3) provides the subjects with information about the evolution of prices for the two brands from “Now” (December 2008) to December 2012. The right-hand side of the screen provides corresponding information on the number of available Blu-ray movie titles. We ask the subjects if and when they would adopt one of the available players.

The subjects in these conjoint experiments face a complex task and need to process a large amount of information. For this reason, we do not vary all product attributes (current and future prices and movie titles) simultaneously across choice tasks. Instead, we vary one factor at a time. The survey consists of two or more *blocks*. Blocks are randomly assigned to the subjects. The first screen in each block presents a particular base scenario (Figure 1), defined by a sequence of prices and the number of movie titles. In the subsequent screens, we randomly vary current or future product prices in one time period only. For example, Figure 2 shows a screen where the price of a medium quality Blu-ray player in December 2009 is lower than in the base scenario in Figure 1, while all other prices remain constant. In the second survey, we also vary the number of movie titles (in at most two time periods) across screens. Across different blocks, we vary the sequence of prices in *all* time periods, and, in survey two, we also vary the sequence of movie titles.

The particular design features of our survey can also capture how the subjects change their adoption timing decisions in response to future price or title changes *within* a given choice scenario. This allows us to conduct simple, model-free tests of forward-looking behavior, which we discuss in Section 5.

Data description

The data were collected using the online panel of Market Tools, Inc, a national market research company. The panel is meant to represent the US population and is used by many large companies such as Canon. Tables 1 and 2 summarize the survey. We conducted survey 1 in February 2009, capturing the responses of 1000 subjects.[5] A fraction was randomly sampled from the Market Tools panel, while the remaining fraction was obtained using oversampling based on expressed interest in high definition (HD) products. We oversampled on HD interest

[5]The sampling was conducted to ensure that the age distribution of the respondents matched the corresponding distribution in the U.S. population.

to ensure that our sample would contain sufficiently many potential costumers for the Blu-Ray technology. Each respondent was exposed to two blocks and answered four questions in each block, resulting in 8,000 choices. We conducted survey 2 in November 2008. 505 respondents completed the survey. Again, we selected a fraction of the respondents using random sampling and the remaining fraction based on expressed interest in HD products. Each respondent was exposed to either two or three blocks and answered six questions in each block, resulting in a data set of 6,576 choices. We ensured that there was no overlap in the subjects sampled across the two experiments. Table 3 shows the distribution of the demographics and other attributes across the subjects in both studies. Due to the oversampling scheme employed, 77% of respondents in the first survey and 82% of respondents in the second survey expressed interest in high definition products.

Tables 1 and 2 also report the distribution of the modal choices across subjects. The data reveal a generally high purchase intent for a Blu-ray player—only 18% of subjects in survey 1 and 17% of subjects in survey 2 most frequently choose "Will not buy." We observe subjects with modes in all survey periods, but overall the modal choices are concentrated in the last two periods in survey 1 and the middle two periods in survey 2.

Within-subject variation in choices is important to document dynamic adoption timing patterns in the data (see Section 5) and to estimate consumer heterogeneity. Tables 1 and 2 show the distributions of the number of distinct choices made by the subjects. For example, if a subject chooses Sony in December 2009 once, Sony in December 2010 three times, and "Will not buy" in all other choice tasks, the number of distinct choices made by that subject is three. The highest possible number of distinct choices in the first experiment is 5 (4 periods and the "Will not buy" option) and in the second experiment is 13 (a combination of 2 brands, 6 periods and the "Will not buy" option). We find that about one third of the subjects never vary their choice across tasks within a block. In survey 1, 48% of the subjects make two distinct choices and 14% make three distinct choices. In survey 2, we observe 51% of the subjects making two or three distinct choices and 12% making four or five distinct choices. While 73% of the subjects always choose the same brand (or the "Will not buy" option), 63% make choices in at least two different time periods.

5 Internal consistency of stated choices

In this section, we provide direct evidence for forward-looking consumer behavior and dynamic adoption timing without resorting to a model that places more structure on consumer behavior. Our survey design allows us to capture within-subject changes in the adoption time that directly reveals whether subjects respond to price changes over time in a rational manner. Suppose a subject is exposed to the price sequence $P_0, \dots, P_T$ in the base screen and chooses to adopt the product in period t. In each subsequent task, exactly one of the $T + 1$ prices increases or decreases. If the subject is exposed to a price decrease or increase in period t, then we classify the observation as exhibiting a *current* price increase or decrease. If the subject is exposed to a price increase or decrease in any period $s < t$, then we classify the observation as exhibiting a *past* price increase or decrease. We define *future* price increases or decreases in the same manner. Suppose the subject's indirect utility function is additively separable across time periods, a standard assumption that we maintained throughout Sections 2 and 3, and that the per-period utility function is decreasing in the product price. Rationality then implies that the subject should not change her adoption choice if the current price decreases. On the other hand, a current price increase is consistent with any choice, including a change of the adoption time to another period or a switch to the "will not buy" option. If a past price decreases, rationality implies that the subject either should not change her choice or should adopt the product earlier, in period $s < t$. Conversely, if a future price decreases, the subject either should not change her choice or should delay the adoption of the product to period $s > t$. However, for both past and future price increases, the subject should not change her choice.

These predictions only hold if the subject's preferences do not vary across choice tasks, i.e. if the random utility terms ϵ_{jt} remain constant. Suppose these terms vary across tasks, for example because the subject makes mistakes. If the mistakes are random, then the changes in adoption times across periods should still be systematically related to the predictions of rational behavior as discussed above.

We first focus on survey 1, where the subjects can only substitute across time, but not across products (although they can choose the outside option of not buying). Table 4 shows how the subjects respond to current, past, and future price increases. Overall, 90% of all observations are correctly classified, i.e. can be explained by rationality without having to resort to changes in the error terms, ϵ_{jt}, across tasks. For a past price decrease, 30% of the subjects buy the product earlier than indicated in the base screen, while 7% buy the product later and 2% switch to "no buy." Conversely, for a past price increase, 5% buy earlier, 5% buy later, and 3% switch to "no buy." If there is a future price decrease, 42% of subjects delay their original purchase, while 5% buy earlier and 1% switch to "no buy." If there is a future price increase, on the other hand, 16% buy later, while 4% buy earlier and 2% switch to "no buy." Overall, with the exception of future price increases, we see that the subjects mostly change their adoption times in a manner that is consistent with the predictions of rational choice.

In survey 2, where the subjects can substitute across brands and across time, we also find the basic pattern of earlier adoption for past price decreases and later adoption for future price decreases. However, compared to survey 1, a smaller fraction (84% of all observations) is correctly classified. A possible reason for this difference is the higher complexity that the subjects face when processing survey 2, which might lead to a higher incidence of mistaken choices. In spite of this difference, subjects still mostly make choices that are consistent with rationality.

6 Estimation and results

Adoption model

The empirical model that we fit to the survey data is a special case of the general dynamic discrete choice model outlined in Section 2. The state vector, $x_t = (P_t, N_t)$, includes the prices of all products considered, P_t, and the number of available movie titles, N_t. We present the subjects with predictions of future prices and the number of available movie titles for the periods $t = 0, 1, \ldots, T$. The subjects either state that they will adopt product j in period $t \leq T$, a choice denoted by $y = (j, t)$, or that they will not buy any product by period T, a choice denoted by $y = 0$. The choice set is $\mathcal{A} = \{(1, 0), (2, 0), \ldots, (J - 1, T), (J, T), 0\}$.

We specify the utility function as follows:

$$u_j(x_t) = \gamma_j + \alpha P_{jt} + \eta_t.$$

γ_j is the intercept for product j, which includes the consumer's valuation of all tangible and intangible product attributes. α is the marginal utility of income, and η_t is the utility derived from watching some (or all) of the available movies from period t onwards:

$$\eta_t = \frac{1}{\rho(t)} \sum_{k=t}^{\infty} \rho(k)(\lambda N_k).$$

The choice-specific value in the survey period, $t = 0$, from adopting product j in period t is then given by

$$\omega_{jt}(x_t) = \rho(t)(\gamma_j + \alpha P_{jt} + \eta_t). \tag{9}$$

The total utility from choice (j, t) is $\omega_{jt}(x_t) + \epsilon_{jt}$, and the total utility from the choice of the reference alternative 0 is defined analogously. By assumption, the consumers anticipate all random utility components, ϵ_{jt}, for the survey periods $t \leq T$, and also the random utility component for the outside option, ϵ_0.

Define $\theta \equiv (\gamma_1, \ldots, \gamma_J, \alpha, \lambda, \rho(1), \ldots, \rho(T))$, a vector indicating the subject's preferences. The choice-specific values are a function of θ. The consumer adopts product j in period t if and only if $\omega_{jt}(x_t; \theta) + \epsilon_{jt} \geq \omega_{ks}(x_s; \theta) + \epsilon_{ks}$ for all $(k, s) \neq (j, t)$ and $\omega_{jt}(x_t; \theta) + \epsilon_{jt} \geq \omega_0(x_T; \theta) + \epsilon_0$. Assuming that the random utility terms are i.i.d. Type I Extreme Value distributed, consumer choices are given by a multinomial logit model with $T \cdot J + 1$ options, where the adoption probability of product j in period t is given by

$$\Pr\{y = (j, t) | \boldsymbol{x}, \theta\} = \frac{\exp(\omega_{jt}(x_t))}{\exp(\omega_0(x_T)) + \sum_{s=0}^{T} \sum_{k=1}^{J} \exp(\omega_{ks}(x_s))}. \tag{10}$$

We allow for preference heterogeneity across subjects. The preferences of subject h are described by a subject-specific parameter vector, θ_h. The corresponding choice probabilities describing the behavior of subject h are given by $\Pr\{y|\boldsymbol{x}, \theta_h\}$. We assume that the subject-level parameters are drawn from a normal population distribution: $\theta_h \sim N(\bar{\theta}, V_\theta)$. The priors on the hyper-parameters, $\bar{\theta}$ and V_θ, are specified as follows:

$$\bar{\theta}|V_\theta \sim N(0, a^{-1}V_\theta),$$
$$V_\theta \sim \text{Inverse-Wishart}(\nu, \nu I).$$

The parameter choices $a = 1/16$ and $\nu = \dim(\theta)+3$ ensure proper but very diffuse prior settings. We estimate the model using a hybrid MCMC approach with a customized Metropolis step as discussed in Rossi, Allenby, and McCulloch (2005) (Chapter 5) and applied in Dubé, Hitsch, and Rossi (2010).

Comparison models

We estimate the model specification above assuming geometric discounting, $\rho(t) = \delta^t$. This model is the dynamic analog of the multinomial logit (MNL) model, and we refer to it as the *Dynamic MNL* model. To assess the role of discounting, we also estimate several alternative specifications that modify some of the assumptions of the baseline Dynamic MNL model.

The Dynamic MNL model nests the case of consumer myopia as a special case: $\delta_h \equiv 0$. For the purposes of comparison, we separately estimate a model in which the subjects are fully myopic and refer to it as the *Current Adoption MNL* model. In this model, the choice-specific value from product adoption in the current period, $t = 0$, is given by

$$\omega_{j0}(x_0) = \gamma_j + \alpha P_{j0} + \lambda N_0,$$

and $\omega_{jt}(x_t) = 0$ for all $t > 0$.

The second comparison model, the *MNL* model, is a standard multinomial logit model in which the inter-temporal choice aspect of durable goods product adoption is absent. Thus,

$$\omega_{jt}(x_t) = \gamma_j + \alpha P_{jt} + \lambda N_t$$

for all periods $t = 0, \dots, T$. This model is equivalent to the case where the consumer has a discount factor of $\delta = 1$ for all future adoption values and a separate discount factor of $\delta = 0$ on the expected future utility from watching movies. This specification captures the typical approach to conjoint modeling of durable goods adoption decisions used by marketing research practitioners.

We also estimate a third specification for comparison, the *Dynamic MNL with Hyperbolic Discounting* model, to test whether consumers exhibit a preference for immediate rewards (see Frederick, Loewenstein, and O'Donoghue 2002 for examples). Following a specification that is widely used in the literature, we parametrize the discount function using the specification of Phelps and Pollak (1968): $\rho(0) = 1$ and $\rho(t) = \beta\delta^t$ for $t > 0$. The parameter β reflects a "present-bias" or inherent preference for immediacy if $\beta < 1$, which is a priori plausible for products such as consumer electronics. δ is the standard long-run discount factor. This specification nests the baseline Dynamic MNL model with geometric discounting when $\beta = 1$.

We initially estimate all four model specifications using the assumption that the value from not adopting by period T is 0, $\omega_0(x_T) \equiv 0$. Later, we conduct a sensitivity analysis to check the robustness of our results to this assumption. We also restrict the parameters δ and β to lie between 0 and 1. We impose this restriction by expressing the parameters using a logistic transformation based on an unrestricted parameter. For example, we estimate $\tilde{\delta}$ and express the corresponding geometric discount factor as $\delta = \exp(\tilde{\delta})/(1 + \exp(\tilde{\delta}))$.

Estimation strategy

The results in Section 3 show that the discount and utility functions are non-parametrically identified from the data generated by our sampling mechanism for the case of homogeneous utility parameters.

The extension of these identification results to the case of heterogeneous utility parameters requires no additional assumptions beyond those typically used for the case of a standard, static discrete choice model. Therefore, the conditions required to identify the heterogeneous analog of our adoption model are the same as in the extant choice modeling literature. For a long panel with a large number of observations per individual, we could estimate separate utility functions for each individual. In practice, researchers typically do not have access to a long panel and, instead, have to resort to pooling observations from different cross-sectional units. Under pooling, each subject's parameter vector, θ_h, is assumed to be drawn from a population distribution, $F(\theta)$. Fox, Kim, Ryan, and Bajari (2012) provide conditions for the non-parametric identification of $F(\theta)$ using a random coefficients logit model estimated with purely cross-sectional data. Identification requires the assumptions that the random utility terms are additive and Type I Extreme Value distributed, that utility has the form $u_j(x; \theta_h) = g(x^T\theta_h)$, and that the state variables are continuous. Other approaches exist to obtain non-parametric identification of a discrete choice model with unobserved heterogeneity (see the survey by Matzkin 2007 for some of the early work in this area, as well as Briesch, Chintagunta, and Matzkin 2010). However, these approaches do not typically nest the random coefficients logit specification that we estimate herein.

Results

For each of the two surveys, we compare results for the four model specifications: Dynamic MNL, Current Adoption MNL, MNL, and Dynamic MNL with Hyperbolic Discounting. For each model we report results for homogeneous preferences and heterogeneous preferences with normally distributed taste parameters. We report quantiles of the posterior distribution of the population parameters to assess the parameter magnitudes and precisions. We use the Newton and Raftery (1994) approach to compute the log marginal likelihood of each model. To address potential overflow concerns, we also report the trimmed marginal likelihood, dropping the upper and lower 2.5 percentile draws. Comparing log marginal likelihoods across models is equivalent to computing a Bayes factor to assess relative posterior model fit (see Rossi, Allenby, and McCulloch 2005).

Survey 1 In survey 1, we focus on a simple choice context where subjects choose if and when to adopt a single Blu-ray player of medium or average quality. Our goal is to test whether subjects are forward-looking and the extent to which they discount future consumption. To keep the survey simple, we do not vary the number of Blu-ray movies across tasks (although the number of movies as reported to the subjects varies across periods). Hence, we do not attempt to estimate the value of watching movies. Instead, we estimate a period-specific intercept to control for the effect of movie titles on choices. Below, in survey 2, we will explicitly consider the effect of movie availability on adoption choices. Results for the four model specifications are reported in Table 5. Note that we estimate the sign of the price coefficient, α, freely.

Comparing the log marginal densities of the four models we see that controlling for between-subject heterogeneity strongly increases the model fit. Hence, below we will only focus on the estimates for heterogeneous preferences.

The Dynamic MNL specification fits better than the two comparison models, Current Adoption MNL and MNL, that do not account for dynamic adoption timing. This is of course consistent with the model-free evidence for forward-looking adoption behavior presented in Section 5. Since the Dynamic MNL model nests the Current Adoption model, a comparison of the log marginal likelihoods provides strong evidence against the restriction that consumers are myopic, $\delta_h \equiv 0$. Besides improving the model fit, we also see substantive changes in the estimated preferences if we allow for dynamic adoption timing. The Dynamic MNL model exhibits a larger degree of heterogeneity, as can be seen by comparing the population standard deviations of taste parameters across models. The Dynamic MNL model also generates a posterior distribution on the price coefficient that is centered around a larger (more negative) value than the benchmark MNL specification.

The Dynamic MNL model has a slightly better fit (based on a comparison of the log marginal densities) than the Dynamic MNL model with Hyperbolic Discounting, in spite of the fact that the latter includes one additional free parameter. The distribution of $\tilde{\beta}$, the parameter determining the "present-bias" $\beta = \exp(\tilde{\beta})/(1 + \exp(\tilde{\beta}))$, is centered close to 1 as can be seen in Figure 6. Therefore, the two models generate substantively similar behavioral predictions.

Our results provide strong evidence that the consideration of inter-temporal trade-offs in adoption timing plays an important role in consumer decision-making. In Figure 4 we report the distribution of the posterior means of the subjects' discount factors. Researchers estimating dynamic discrete choice problems routinely assume that consumers discount the future with a common discount factor corresponding to some aggregate interest rate or asset return. For annual decision-making, this assumption would imply a discount factor of approximately 0.95 if the interest rate is 5 percent. However, the estimates in Figure 4 show that the behavior of the subjects in our survey is characterized by discount factors that are much smaller than typically assumed. The average subject has a discount factor of roughly 0.7, corresponding to an annual interest rate of 43%. Moreover, there is a large degree of heterogeneity in discounting across subjects.

Survey 2 The data from survey 2 allow us to check if the findings regarding discounting and forward-looking adoption behavior are robust to a more complex and possibly more realistic environment. In survey 2 we add a choice between Blu-ray players of two leading brands, Sony and Samsung, and we also vary the number of available movie titles across choice tasks. Results for the four model specifications are reported in Table 6. Our basic findings are consistent with those from survey 1. In particular, allowing for preference heterogeneity strongly improves the model fit, and the Dynamic MNL model fits much better than the two comparison models that do not allow for dynamic adoption timing.

We see the importance of allowing for an unrestricted discount factor even more strongly in survey 2. For the Current Adoption MNL and the MNL models the posterior distribution of the coefficient on titles, λ, is centered well below zero, implying that the majority of the subjects prefer fewer available movie titles. However, we obtain a positive sign on the titles coefficient once we free up the discount factor parameter in the Dynamic MNL model. Also, as in survey 1, we find that the Dynamic MNL model predicts more preference heterogeneity and that the posterior distribution of the price coefficient is centered around a larger (more negative) value than for the two models that do not allow for dynamic decision making.

Unlike in survey 1, the Dynamic MNL model with Hyperbolic Discounting has a slightly better fit than the Dynamic MNL model. Thus, the comparison of log marginal likelihoods implies that we reject the parameter restriction $\beta_h \equiv 1$. However, the differences in the estimates for the two models are very small. In Figure 7, the distribution of the posterior means of the present-bias parameters β has most of its mass close to 1.

In Figure 5, we report the distribution of the posterior means of the subjects' discount factors. Exactly as in survey 1, the average subject has a discount factor of approximately 0.7, corresponding to an interest rate of 43%, and we observe considerable heterogeneity in discounting across the subjects.

Model fit for assumed discount factors

We already highlighted that researchers estimating dynamic discrete choice problems using field data typically assume a homogenous value for the discount factor based on some aggregate interest rate or asset return. We now show how the preference estimates change if a specific, homogenous discount factor is assumed. We re-estimate the Dynamic MNL model for the discount factors $\delta = 0, 0.2, 0.4, 0.6, 0.8, 0.9$ and report the corresponding posterior model fit of the baseline Dynamic MNL specification in Table 7. The discount factors of $\delta = 0.6$ and $\delta = 0.8$ provide a much better model fit than $\delta = 0.9$, even though the latter more closely corresponds to a realistic long-run annual interest rate of about 11 percent. A discount factor of $\delta = 0.6$ corresponds to an annual interest rate of about 66 percent, indicating a considerable degree of impatience. Generally the Dynamic MNL model fits considerably better than any of the assumed, homogenous discount factor models. Although not reported, we also find that restricting the discount factor to a fixed value results in an upward bias in the degree of heterogeneity in the model parameters.

Some tentative managerial implications

In Table 8, we report the estimated correlation matrix for the population distribution of the taste parameters. We observe a strong negative correlation between $\tilde{\delta}$ (recall that $\delta = \exp(\tilde{\delta})/(1 + \exp(\tilde{\delta}))$) and the price coefficient. Thus, higher patience (i.e. a larger discount parameter) is associated with higher price sensitivity (i.e. a larger price coefficient in absolute value). To a lesser extent, we also observe a negative correlation between the discount parameter and the titles coefficient, implying that subjects with a relatively low utility from titles exhibit more patience. Both these correlations have interesting managerial implications. Suppose a monopolist was selling a Blu-Ray player to this population of consumers. Coase argued that skimming (inter-temporal price discrimination) would unravel if high willingness-to-pay consumers could anticipate future price discounts and wait until the price discounts became available. In the limit case where consumers were perfectly patient ($\delta = 1$), even a monopolist selling a durable good would lose all market power. As demonstrated by Nair (2007), if consumers are impatient ($\delta < 1$) price skimming is possible. But prices and profits decrease for larger values of δ. Our findings indicate that high willingness-to-pay consumers (i.e. those with a low sensitivity to prices and a high valuation for the number of available titles) are also the most impatient, which works even more strongly against the Coasean view. We could not have obtained this insight from an analysis based on field data, as field data do not allow us to distinguish between early adoption due to a high intrinsic utility from the product versus impatience.

Sensitivity analysis

Movie consumption after period T We use the study 2 data to explore the sensitivity of the key results to the assumptions about consumer beliefs after the final period T in the survey. The Dynamic MNL model assumes that $N_t = N_T$ for all $t > T$. Suppose we assume instead that subjects only consider movie consumption utility up until T: $\eta_t = \sum_{k=t}^{T} \delta^{k-t}(\lambda N_k)$. The posterior fit of this specification is better than that of our Dynamic MNL. However, the two specifications produce very similar parameter estimates. For instance, in Figure 8 we show that the distribution of the discount factor is statistically indistinguishable from the Dynamic MNL case. Therefore our findings do not appear to be sensitive to the assumption of consumers obtaining utility from the availability of movies beyond the last survey period, T.

Delaying adoption after period T In the model specifications above we assume that the value from choosing the reference alternative 0 is $\omega_0(x_T) \equiv 0$ regardless of the value of the state in the final survey period. This assumption would be justified if the subjects thought of the reference alternative as the choice of never adopting the Blu-ray technology. We relax this assumption to allow for the possibility that subjects treat the no-purchase alternative as an option to delay adoption to some period later than T. For example, a consumer might want to delay the adoption of a Blu-ray player if she expects that prices decline after period T. To capture this option value we assume that the subjects anticipate that they will make sequential adoption decisions beyond period T according to the dynamic discrete choice model outlined in Section 2. We restrict the analysis to the case of geometric discounting. The choice-specific values from adopting product j or delaying the adoption, $v_k(x_t)$, are then given by equation (1). Thus, at time $t = 0$ the option value from delaying the adoption decision beyond period T is given by

$$\omega_0(x_T) = \delta^T v_0(x_T) = \delta^{T+1} \int \max_{k \in \{0,\ldots,J\}} \{v_k(x) + \epsilon_k\} p(\epsilon) p(x|x_T) d\epsilon dx.$$

$p(x_{t+1}|x_t)$ denotes the subjects' expectations about the evolution of the states beyond period T.

We estimate three specifications of the model with delayed adoption. In the first specification, we assume that the subjects believe that prices remain fixed after period T, such that $P_t = P_T$ for all $t > T$. In this specification, the benefit from delaying adoption is due to the option to re-draw the latent utility components, ϵ_{kt}, in each future period. In two additional specifications, we assume that the subjects anticipate that prices decline until they reach a long-run, constant price level ten periods after T. These two specifications differ in the rate at which prices decline. We hold the number of available movie titles fixed at $N_t = N_T$ for all $t > T$.

In Table 9, we compare the log marginal likelihood of the baseline *Dynamic MNL* specification and the three models allowing for delayed adoption. The baseline specification has a higher posterior fit, and the estimates for the comparison models are qualitatively similar to the results for the baseline model. For example, in Figure 9 we display the distribution of the discount factor, δ, for the baseline and comparison specifications. The marginal densities for the comparison model discount factors are largely within the 90 percent pointwise posterior credibility region that envelopes the marginal density of the discount factor corresponding to the baseline model. Only the marginal density corresponding to the model with fast price decline puts more mass on smaller discount factors. In summary, the main results do not appear to be sensitive to the assumption on the value from choosing the reference alternative.

7 Conclusions

The main focus of this paper is methodological. We describe a sampling mechanism that can be implemented as a conjoint design to capture stated choices across options within a period and across time. Based on data obtained from this sampling mechanism, the subjects' discount and utility functions are separately identified. Thus, this sampling mechanism overcomes a fundamental identification problem for dynamic discrete choice models when estimated from field data.

Our method is intended to be a practical approach for durable goods demand estimation when consumers are forward-looking and consider the tradeoff between buying now or buying at some future date. It should be of interest both to marketing researchers studying durable goods adoption and marketing practitioners who use conjoint studies to plan the launch and other marketing activities for a new durable good.

We illustrate our approach using two conjoint studies eliciting dynamic adoption choices of Blu-ray players. Both the raw data patterns and the model estimates reveal that the subjects in our survey make forward-looking adoption choices. However, the implied discount rates corresponding to the estimated discount factors are around 43 percent and, thus, are much larger than typically observed asset returns. Therefore, calibrating discount factors based on aggregate interest rates or asset returns, a common and, due to the identification problem, necessary practice in empirical work estimating dynamic discrete choice models from field data, may result in biased and misleading preference estimates. The stated choice data also reveal a large degree of heterogeneity in the discount rates, again casting doubt on the assumption of a common, homogenous discount factor that is typically used in empirical work based on field data. We find little evidence for hyperbolic discounting.

The results from the two studies are insufficient to make general statements about discounting in other dynamic discrete choice contexts, including but not limited to durable goods adoption decisions. At a minimum, however, our findings suggest that discount factors should ideally be estimated, not assumed, and that empirical work using dynamic discrete choice models would benefit from combining field and survey data.

Although this paper makes some progress on a fundamental question in marketing and economics, many issues are left unresolved. First, using our survey design, we can estimate current-period preferences and discount functions by “endowing” the subjects with deterministic beliefs. That the subjects indeed take these beliefs as deterministically given is an assumption that we have not tested. In practice, consumers have to form non-deterministic beliefs based on available information. How consumers form such beliefs is a question of great importance that is largely unresolved. Second, the analysis of our stated choice data is based on very specific assumptions on the latent utility shocks that serve as the econometric error term. The same is of course true for all discrete choice models. How the utility shocks in our survey relate to the utility shocks that might influence subjects when making “real world” decisions is also unresolved. Third, our survey is fairly complex to process, and this poses the question of how the survey design influences the quality of the stated choice data. For example, we saw that the incidence of stated choices that are difficult to rationalize was lower in the simpler survey 1 than in the more complex survey 2. All three issues are important topics for future research.

References

Abbring, J. H. (2010): "Identification of Dynamic Discrete Choice Models," *Annual Review of Economics*, 2, 367–394.

Arcidiacono, P., and P. Ellickson (2011): "Measuring long-run marketing effects and their implications for long-run marketing decisions," *Annual Review of Economics*, 3, 363–394.

Bajari, P., V. Chernozhukov, H. Hong, and D. Nekipelov (2009): "Nonparametric and Semiparametric Analysis of a Dynamic Game Model," *manuscript*.

Bass, F. M. (1969): "A New Product Growth Model for Consumer Durables," *Management Science*, 15(5), 215–227.

Bass, F. M., K. Gordon, T. L. Ferguson, and M. L. Githens (2001): "DIRECTV: Forecasting Diffusion of a New Technology Prior to Product Launch," *Interfaces*, 31, 82–93.

Briesch, R. A., P. Chintagunta, and R. Matzkin (2010): "Nonparametric Discrete Choice Models with Unobserved Heterogeneity," *Journal of Business and Economic Statistics*, 28(1), 291–307.

Bronnenberg, B. J., J. P. Dubé, C. F. Mela, P. Albuquerque, T. Erdem, B. Gordon, D. Hanssens, G. Hitsch, H. Hong, and B. Sun (2008): "Measuring long-run marketing effects and their implications for long-run marketing decisions," *Marketing Letters*, 19, 367–382.

Chevalier, J., and A. Goolsbee (2005): "Are Durable Goods Consumers Forward Looking? Evidence from College Textbooks," *NBER Working Paper 11421*.

Chung, D. J., T. Steenburgh, and K. Sudhir (2011): "Do Bonuses Enhance Sales Productivity? A Dynamic Structural Analysis of Bonus-Based Compensation Plans," *manuscript*.

Cochrane, J. H. (2001): *Asset Pricing*. Princeton University Press, Princeton, NJ.

Dubé, J.-P., G. J. Hitsch, and P. E. Rossi (2010): "State Dependence and Alternative Explanations for Consumer Inertia," *The RAND Journal of Economics*, 47 (6), 998–1010.

Fang, H., and Y. Wang (2012): "Estimating Dynamic Discrete Choice Models with Hyperbolic Discounting, with an Application to Mammography Decisions," *Manuscript*.

Fox, J., K. Kim, S. Ryan, and P. Bajari (2012): "The Random Coefficients Logit Model Is Identified," *Journal of Econometrics*, 166(2), 204–212.

Frederick, S., G. Loewenstein, and T. O'Donoghue (2002): "Time Discounting and Time Preference: A Critical Review," *Journal of Economic Literature*, 40(2), 351–401.

Green, P. E., A. M. Krieger, and Y. Wind (2001): "Thirty Years of Conjoint Analysis: Reflections and Prospects," *Interfaces*, 31(3), S56–S73.

Green, P. E., and V. R. Rao (1971): "Conjoint Measurement for Quantifying Judgmental Data," *Journal of Marketing Research*, 8(3), 355–363.

Green, P. E., and V. Srinivasan (1990): "Conjoint Analysis in Marketing: New Developments with Implications for Research and Practice," *Journal of Marketing*, 54(4), 3–19.

Horsky, D. (1990): "A Diffusion Model Incorporating Product Benefits, Price, Income and Information," *Marketing Science*, 9(4), 342–365.

Hotz, V. J., and R. A. Miller (1993): "Conditional Choice Probabilities and the Estimation of Dynamic Models," *Review of Economic Studies*, 60(3), 497–529.

Huber, J. (1997): "What We Have Learned from 20 Years of Conjoint Research: When to Use Self-Explicated, Graded Pairs, Full Profiles or Choice Experiments," *Sawtooth Software Research Paper Series*.

Magnac, T., and D. Thesmar (2002): "Identifying Dynamic Discrete Decision Processes," *Econometrica*, 70(2), 801–816.

Matzkin, R. (2007): "Heterogeneous Choice," *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress*, III, Ch. 4.

Melnikov, O. (2001): "Demand for Differentiated Durable Products: The Case of the U.S. Computer Printer Market," *manuscript*.

Miller, R. (1984): "Job Matching and Occupational choice," *Journal of Political Economy*, 92, 1086–1120.

Nair, H. S. (2007): "Intertemporal Price Discrimination with Forward-Looking Consumers: Application to the US Market for Console Video-Games," *Quantitative Marketing and Economics*, 5, 239–292.

Newton, M. A., and A. E. Raftery (1994): "Approximate Bayesian Inference with the Weighted Likelihood Bootstrap," *Journal of the Royal Statistical Society, Series B*, 56(1), 3–48.

Norets, A., and S. Takahashi (2012): "On the Surjectivity of the Mapping Between Utilities and Choice Probabilities," *Manuscript*.

Pakes, A. (1986): "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks," *Econometrica*, 54(4), 755–784.

Phelps, E. S., and R. A. Pollak (1968): "On Second-Best National Saving and Game-Equilibrium Growth," *Review of Economic Studies*, 35(2), 185–199.

Rossi, P., G. Allenby, and R. McCulloch (2005): *Bayesian Statistics and Marketing.* John Wiley & Sons.

Rust, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55(5), 999–1033.

——— (1994): "Structural Estimation of Markov Decision Processes," in *Handbook of Econometrics*, ed. by R. F. Engle, and D. L. McFadden, vol. IV, pp. 3081–3143. Amsterdam, London and New York: Elsevier, North-Holland.

Song, I., and P. K. Chintagunta (2003): "A Micromodel of New Product Adoption with Heterogeneous and Forward-Looking Consumers: Application to the Digital Camera Category," *Quantitative Marketing and Economics*, 1, 371–407.

Viscusi, W. K., J. Huber, and J. Bell (2008): "Estimating discount rates for environmental quality from utility-based choice experiments," *Journal of Risk and Uncertainty*, 37, 199–220.

Yao, S., C. F. Mela, J. Chiang, and Y. Chen (2012): "Determining Consumers' Discount Rates With Field Studies," *manuscript.*

A Survey description

In several introductory screens, we provide subjects information about the Blu-ray player and the design of the study. The objective of these screens is to educate the subjects about Blu-ray players and familiarize them with the survey design and the choice tasks they will encounter. First, we show subjects differences in picture quality when using a Blu-ray player versus a traditional DVD. We then provide an in-depth description of the Blu-ray technology. Specifically, we provide the subjects information on how the Blu-ray discs and players work, the leading manufacturers of Blu-ray technology, and similarities between a Blu-ray disc and a traditional DVD. Subjects are then shown the Blu-ray player they will have the option of choosing across different choice tasks. In the first survey, the subjects are told that they have the option of choosing a mid-high quality Blu-ray player, while in the second survey, they have a choice between a medium quality player (Samsung) and a state of the art top quality Blu-ray player (Sony).

In the subsequent screens, we provide subjects details on how to respond to the choice tasks and navigate the survey. We ask the subjects about their current Blu-ray player ownership status and test their understanding of the price (and number of titles) path over time using graphs from a typical choice task. Based on their responses, subjects are provided feedback about the graphs and the choice tasks to make sure they understand the trade-offs involved across different choices. Figures 1 and 3 show typical choice tasks in the first and second survey, respectively. In each choice task, subjects are provided with the prices (and number of titles) of the Blu-ray player(s) over four (six) periods. Subjects are told that the prices and number of titles shown are expert predictions and they should expect to see these prices in the future. Additionally, they are asked to ignore changes in prices due to inflation. Further details about the variation in prices and number of titles available across different choice tasks is present in Section 4.

Table 1: Survey 1 Description

Survey Overview	
Price variation over time	
Across subjects	Yes
Within subjects	Yes
Variation in no. of titles over time	
Across subjects	No
Within subjects	Yes
Adoption decision	Only inter-temporal choice
No. products (brands)	1
No. of survey time periods	4
Total no. of choices(*)	5
Survey Summary	
No. subjects	1,000
No. blocks per subject	2
No. questions per block (including baseline)	4
No. of price manipulations	3
No. of title manipulations	0
Total number of choices in data	8,000
Distribution of Choices	
Mar-09	3%
Dec-09	11%
Dec-10	27%
Dec-11	41%
Will not buy	19%
Modal Choices	
Mar-09	2%
Dec-09	11%
Dec-10	28%
Dec-11	41%
Will not buy	18%
Number of distinct choices	
1	36.5%
2	48.3%
3	14.2%
4	0.7%
5	0.3%

Note. The table summarizes the prices and titles variation over time, both within and across subjects, in survey 1. Additionally, the table summarizes the survey data and reports the distribution of choices, modal choices and the number of distinct choices made by subjects in survey 1.

Table 2: Survey 2 Description

Survey Overview		
Price variation over time		
Across subjects	Yes	
Within subjects	Yes	
Variation in no. of titles over time		
Across subjects	Yes	
Within subjects	Yes	
Adoption decision	Inter-temporal and brand choice	
No. products (brands)	2	
No. of survey time periods	6	
Total no. of choices(*)	13	
Survey Summary		
No. subjects	505	
No. blocks per subject	2 or 3	
No. questions per block (including baseline)	6	
No. of price manipulations	3	
No. of title manipulations	2	
Total number of choices in data	6,576	
Distribution of Choices	**Sony**	**Samsung**
Dec-08	3%	2%
Jun-08	4%	5%
Dec-09	6%	11%
Dec-10	8%	15%
Dec-11	5%	9%
Dec-12	5%	10%
Will not buy	18%	
Modal Choices	**Sony**	**Samsung**
Dec-08	3%	1%
Jun-08	4%	5%
Dec-09	9%	11%
Dec-10	8%	16%
Dec-11	4%	7%
Dec-12	5%	9%
Will not buy	17%	
Number of distinct choices		
1	34.3%	
2	32.5%	
3	18.8%	
4	6.9%	
5	4.8%	
6	1.6%	
7	0.8%	
8	0.2%	
9	0.2%	
Number of distinct brands chosen		
1	73%	
2	25%	
3	2%	

Number of distinct time periods chosen	
1	36.8%
2	33.9%
3	19.6%
4	7.7%
5	1.2%
6	0.8%

Note. The table summarizes the prices and titles variation over time, both within and across subjects, in survey 2. Additionally, the table summarizes the survey data and reports the distribution of choices, modal choices, number of distinct choices, and the number of distinct brands and time periods chosen by subjects in survey 2.

Table 3: Survey Demographics

	Survey 1	**Survey 2**
Number of respondents	1000	505
% Males	49%	62%
Interest in HD products?	77%	82%
Age Distribution		
20-25	8%	7%
26-30	12%	13%
31-35	16%	16%
36-40	11%	11%
41-45	11%	11%
45+	43%	41%
Ethnicity		
White/Caucasian	87%	84%
African American	5%	5%
Hispanic	2%	4%
Asian	4%	5%
Other	2%	2%
Education Level		
Less than high school	1%	0%
High school	28%	22%
College	48%	50%
Graduate degree	21%	24%
Other	2%	3%
Household Income		
Less than $25,000	9%	10%
$25,000 - $50,000	28%	28%
$50,000 - $75,000	23%	22%
$75,000 - $100,000	17%	16%
$100,000 - $150,000	9%	13%
More than $150,000	5%	4%
Future Income Expectation		
Decrease in near future	14%	13%
No change	70%	68%
Increase in near future	16%	19%
Ownership		
TV	92%	91%
Satellite/Cable TV	78%	78%
DVD player	95%	92%
Flat panel	47%	41%
Other HD	22%	23%
DVR	40%	37%

Note. The table reports the distributions of different demographics and ownership status of subjects for both surveys.

Table 4: Direct Evidence of Forward-Looking Behavior

	Current Price Decrease	Current Price Increase	Past Price Decrease	Past Price Increase	Future Price Decrease	Future Price Increase
Survey 1						
No. obs.	398	483	960	954	332	311
No change	89%	18%	62%	87%	53%	78%
Buy earlier	7%	19%	30%	5%	5%	4%
Buy later	3%	36%	7%	5%	42%	16%
Switch to no buy	1%	28%	2%	3%	1%	2%
Correctly classified	89%	100%	92%	87%	95%	78%
Total correctly classified			90%			
Survey 2						
No. obs.	84	44	583	432	40	43
No change	63%	34%	63%	69%	35%	58%
Buy same product earlier	7%	5%	20%	8%	13%	2%
Buy same product later	19%	48%	7%	14%	33%	33%
Switch to other product	11%	11%	9%	7%	20%	7%
Switch to no buy	0%	2%	2%	3%	0%	0%
Correctly classified	74%	100%	92%	75%	88%	65%
Total correctly classified			84%			

Note. The table provides model free evidence that consumers are forward looking by analyzing internal consistency of choices. The top (bottom) panel reports results from survey 1 (2). For each block of questions, the columns detail changes in past, current, or future prices (in subsequent choice tasks) relative to the time of adoption in the base screen/choice task. The rows summarize deviation in choices in the subsequent tasks relative to the choice made in the base choice task. For each survey, we report the number of observations in each price scenario (column), distribution of choice deviation patterns by type of price variation, and the percentage of correctly classified responses. The correctly classified responses are choices which can be explained by rationality, without having to resort to changes in error terms, ϵ_{jt}, across tasks.

Table 5: Survey 1: Model Estimates

	Homogeneous Tastes			Heterogeneous Tastes					
				Pop. mean			Pop. SD		
	2.5%	50%	97.5%	2.5%	50%	97.5%	2.5%	50%	97.5%
Dynamic MNL Model									
γ_1	6.58	7.05	7.53	32.67	38.82	43.23	14.83	20.00	22.72
γ_2	6.27	6.65	7.00	31.81	39.47	44.55	16.62	21.42	23.94
γ_3	5.80	6.08	6.34	32.78	40.88	45.66	15.86	21.16	23.70
γ_4	4.48	4.68	4.88	26.94	34.29	38.97	10.95	14.54	17.07
Price (α)	-3.14	-2.97	-2.81	-22.32	-19.96	-15.87	7.51	9.87	11.28
Discount ($\tilde{\delta}$)	4.51	9.58	22.86	0.97	1.15	1.64	0.94	1.06	1.27
Log marginal likelihood		-10334.69				-4009.70			
Trimmed log m. l.		-10331.46				-3933.54			
Current Adoption Model									
γ_1	-2.20	-1.93	-1.65	-7.35	-5.62	-4.38	2.99	3.71	4.62
Price (α)	-36.59	-22.08	-10.99	-113.22	-81.01	-39.37	3.93	10.58	29.16
Log marginal likelihood		-12293.19				-11785.32			
Trimmed log m. l.		-12292.01				-11774.41			
MNL model									
γ_1	6.56	7.04	7.49	14.77	16.99	19.55	12.27	13.89	15.71
γ_2	6.29	6.64	6.97	17.70	19.10	20.67	11.50	12.87	14.36
γ_3	5.78	6.06	6.33	18.11	19.37	20.89	9.51	10.54	11.77
γ_4	4.48	4.67	4.86	14.09	15.03	16.14	6.17	6.94	7.80
Price (α)	-3.12	-2.97	-2.81	-9.75	-9.04	-8.38	4.46	5.06	5.76
Log marginal likelihood		-10333.08				-4291.01			
Trimmed log m. l.		-10331.39				-4216.51			
Dynamic MNL with Hyperbolic Discounting									
γ_1	6.57	7.33	7.50	25.53	36.07	41.54	12.92	17.35	20.12
γ_2	6.25	6.83	7.04	26.43	36.13	41.00	14.52	19.52	22.23
γ_3	5.74	6.20	6.41	26.91	37.37	42.39	13.37	19.39	22.47
γ_4	4.47	4.75	4.91	21.42	31.40	35.81	8.37	12.72	14.53
Price (α)	-3.14	-3.06	-2.80	-20.65	-18.25	-13.02	6.12	8.92	10.19
Present bias ($\tilde{\beta}$)	3.24	7.93	20.56	5.54	7.49	8.78	1.56	4.02	5.02
Discount ($\tilde{\delta}$)	4.39	7.89	14.18	1.02	1.31	2.19	0.90	1.04	1.52
Log marginal likelihood		-10334.02				-4065.00			
Trimmed log m. l.		-10331.36				-3998.43			

Note. The table summarizes the parameter estimates from different models estimated using the survey 1 data. We report estimates from both, the homogeneous and the heterogeneous models. For the homogeneous model, we report estimates corresponding to the 95% confidence region along with the median estimate. For the heterogeneous model, we report the same statistics for both, the mean and the standard deviation of the distribution of population heterogeneity.

Table 6: Survey 2: Model Estimates

	Homogeneous Tastes			Heterogeneous Tastes					
				Pop. mean			Pop. SD		
	2.5%	50%	97.5%	2.5%	50%	97.5%	2.5%	50%	97.5%
Dynamic MNL Model									
γ_{Sony}	3.50	4.51	5.59	30.66	37.26	45.27	40.32	45.96	52.06
$\gamma_{Samsung}$	2.77	3.60	4.47	26.24	31.85	38.40	34.78	39.55	44.61
Price (α)	-2.43	-2.18	-1.93	-22.98	-20.26	-18.22	13.71	15.46	17.79
Titles (λ)	0.12	0.18	0.25	1.10	1.30	1.52	1.05	1.24	1.46
Discount ($\tilde{\delta}$)	0.03	0.16	0.29	0.72	0.88	1.04	0.93	1.05	1.19
Log marginal likelihood		-10443.79				-4109.48			
Trimmed log m. l.		-10441.87				-4049.63			
Current Adoption Model									
γ_{Sony}	0.80	2.08	3.37	-5.13	3.07	12.97	4.14	8.11	12.01
$\gamma_{Samsung}$	0.40	1.38	2.40	-4.05	1.79	8.62	4.21	7.02	9.77
Price (α)	-1.17	-0.84	-0.48	-6.39	-3.10	-0.87	1.61	2.68	4.17
Titles (λ)	-1.11	-0.60	-0.11	-2.50	-1.23	0.27	0.86	1.60	2.46
Log marginal likelihood		-10827.43				-10299.89			
Trimmed log m. l.		-10826.12				-10279.34			
MNL Model									
γ_{Sony}	4.57	4.92	5.24	29.09	32.88	37.66	19.58	22.25	25.70
$\gamma_{Samsung}$	4.09	4.37	4.64	26.40	29.72	33.89	16.82	19.22	21.98
Price (α)	-1.78	-1.67	-1.56	-11.46	-10.27	-9.23	6.47	7.43	8.48
Titles (λ)	-0.17	-0.16	-0.15	-1.27	-1.13	-1.00	0.89	1.00	1.13
Log marginal likelihood		-10372.45				-4385.77			
Trimmed log m. l.		-10370.30				-4350.95			
Dynamic MNL with Hyperbolic Discounting									
γ_{Sony}	3.67	4.56	5.42	26.26	34.97	43.37	43.56	53.60	61.60
$\gamma_{Samsung}$	2.92	3.61	4.26	22.81	29.81	36.61	37.84	46.24	52.71
Price (α)	-2.42	-2.18	-1.93	-23.36	-20.43	-17.97	14.54	17.19	19.59
Titles (λ)	0.11	0.19	0.25	1.10	1.32	1.61	1.01	1.24	1.52
Present bias ($\tilde{\beta}$)	4.05	11.03	24.90	5.39	7.02	9.84	3.61	4.80	7.23
Discount ($\tilde{\delta}$)	0.06	0.16	0.30	0.89	1.05	1.22	0.94	1.06	1.20
Log marginal likelihood		-10444.61				-4022.80			
Trimmed log m. l.		-10441.77				-3995.47			

Note. The table summarizes the parameter estimates from different models estimated using the survey 2 data. We report estimates from both, the homogeneous and the heterogeneous models. For the homogeneous model, we report estimates corresponding to the 95% confidence region along with the median estimate. For the heterogeneous model, we report the same statistics for both, the mean and the standard deviation of the distribution of population heterogeneity.

Table 7: Impact of an assumed discount factor on model fit (Dynamic MNL)

	Log Marginal Likelihood
Dynamic MNL	-4109.476
$\delta = 0$	-10292.95
$\delta = 0.2$	-5170.171
$\delta = 0.4$	-4348.224
$\delta = 0.6$	-4234.535
$\delta = 0.8$	-4270.569
$\delta = 0.9$	-4309.22

Note. The table compares the log marginal likelihoods from the Dynamic MNL model with alternate specifications estimated assuming a fixed discount factor.

Table 8: Survey 2: Correlation Matrix for Population Distribution of Tastes (Dynamic MNL Model)

	γ_{Sony}	$\gamma_{Samsung}$	Price (α)	Titles (λ)	Discount ($\tilde{\delta}$)
γ_{Sony}	1				
	(1,1)				
$\gamma_{Samsung}$	0.93	1			
	(0.9,0.95)	(1,1)			
Price (α)	-0.90	-0.92	1		
	(-0.93,-0.86)	(-0.94,-0.88)	(1,1)		
Titles (λ)	-0.14	-0.23	0.12	1	
	(-0.3,0.01)	(-0.38,-0.08)	(-0.02,0.29)	(1,1)	
Discount ($\tilde{\delta}$)	0.15	0.23	-0.40	-0.34	1
	(-0.01,0.33)	(0.06,0.39)	(-0.54,-0.25)	(-0.49,-0.18)	(1,1)

Note. The table reports the correlation matrix (along with the 95% confidence region) for the population distribution of preferences from survey 2.

Table 9: Posterior Fit of the Delayed Adoption Specifications

Models	**Log Marginal Likelihood**
Dynamic MNL	-4049.63
Delayed adoption and slow price decline	-4066.61
Delayed adoption and fast price decline	-4080.11
Delayed adoption and no price decline	-4074.31

Note. The table compares the log marginal likelihoods from the Dynamic MNL model with alternate specifications which allow for delayed adoption and different rates of price decline.

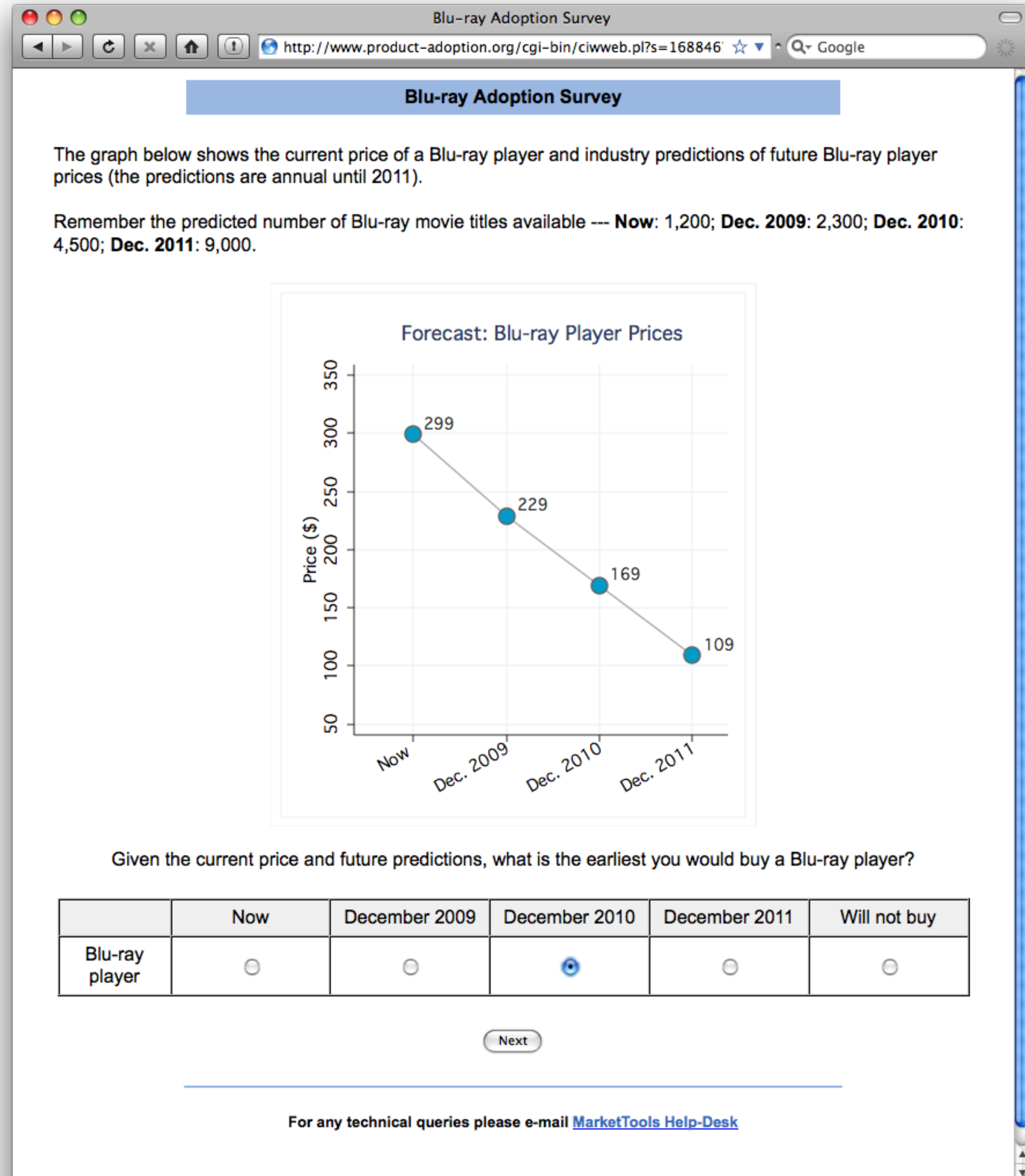
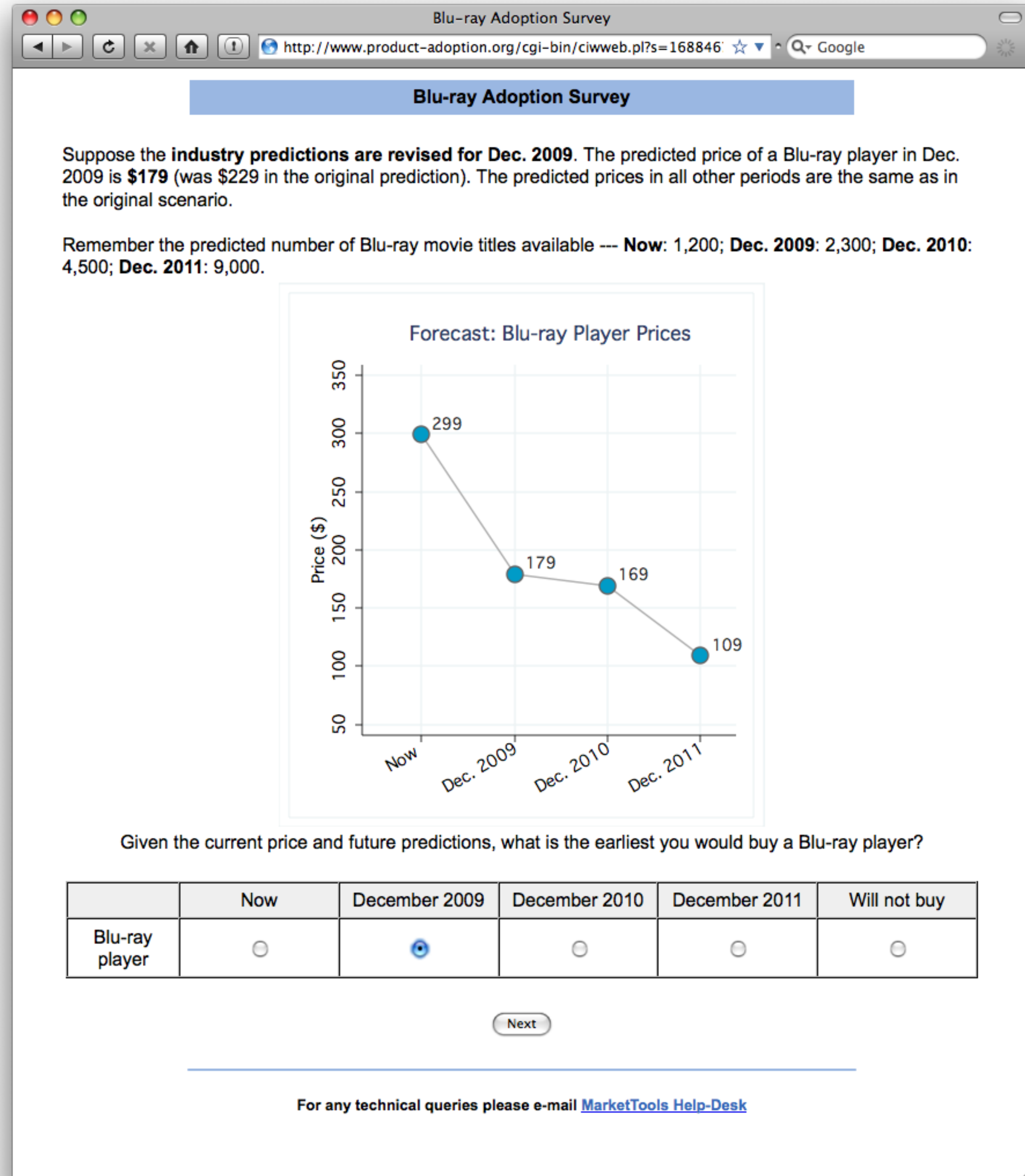


Figure 1: Survey screen: Base scenario

Note. The figure shows a sample base choice task from survey 1. The screen lists the number of titles available over time, provides a forecast of the Blu-ray player prices over time, and allows subjects to choose the time of Blu-ray player adoption.



Blu-ray Adoption Survey

http://www.product-adoption.org/cgi-bin/ciwweb.pl?s=168846

Blu-ray Adoption Survey

Suppose the **industry predictions are revised for Dec. 2009**. The predicted price of a Blu-ray player in Dec. 2009 is **$179** (was $229 in the original prediction). The predicted prices in all other periods are the same as in the original scenario.

Remember the predicted number of Blu-ray movie titles available --- **Now**: 1,200; **Dec. 2009**: 2,300; **Dec. 2010**: 4,500; **Dec. 2011**: 9,000.

Given the current price and future predictions, what is the earliest you would buy a Blu-ray player?

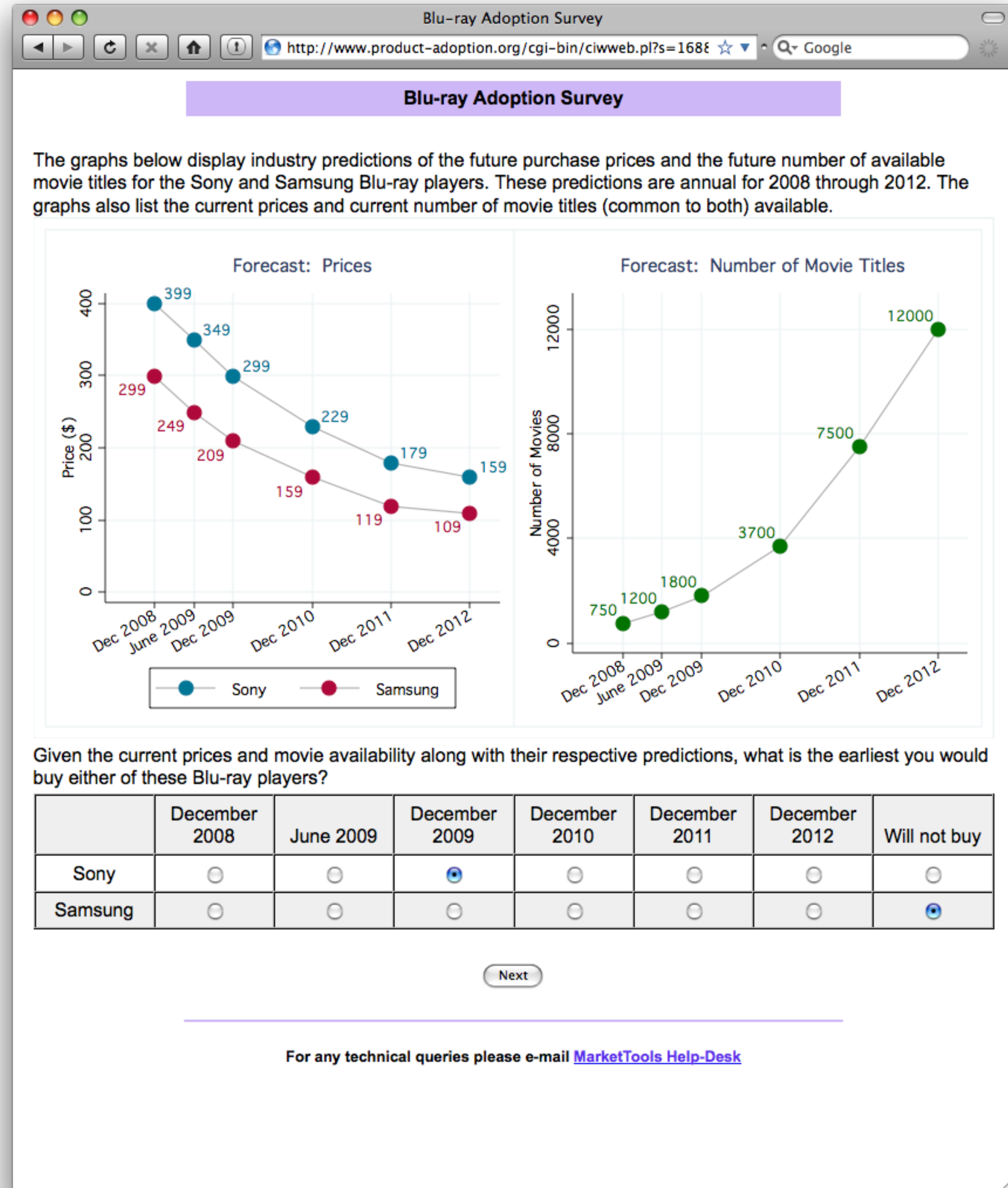
	Now	December 2009	December 2010	December 2011	Will not buy
Blu-ray player	○	◉	○	○	○

Next

For any technical queries please e-mail MarketTools Help-Desk

Figure 2: Survey screen: Price decrease in December 2009 relative to base scenario

Note. The figure shows a subsequent sample choice task from survey 1. The subsequent choice tasks differ relative to the base screen (shown in 1) in the variation in prices over time.



Blu-ray Adoption Survey

http://www.product-adoption.org/cgi-bin/ciwweb.pl?s=168ξ

Google

Blu-ray Adoption Survey

The graphs below display industry predictions of the future purchase prices and the future number of available movie titles for the Sony and Samsung Blu-ray players. These predictions are annual for 2008 through 2012. The graphs also list the current prices and current number of movie titles (common to both) available.

Given the current prices and movie availability along with their respective predictions, what is the earliest you would buy either of these Blu-ray players?

	December 2008	June 2009	December 2009	December 2010	December 2011	December 2012	Will not buy
Sony	○	○	●	○	○	○	○
Samsung	○	○	○	○	○	○	●

Next

For any technical queries please e-mail MarketTools Help-Desk

Figure 3: Base screen in second survey

Note. The figure shows a sample base choice task from survey 2. The screen provides a forecast of the Sony and Samsung Blu-ray player prices over time, a forecast of the number of titles available over time, and allows subjects to choose the time of adoption and brand of Blu-ray player.

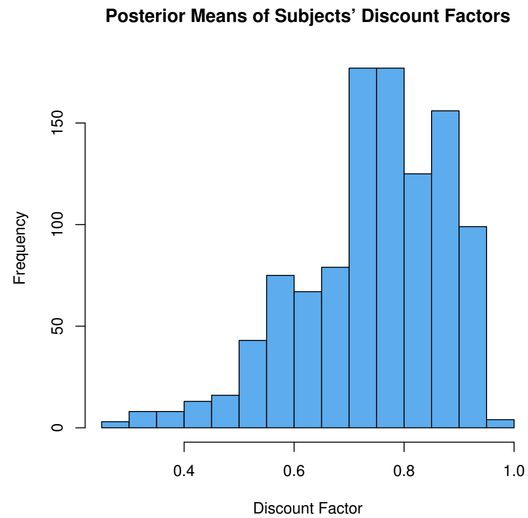


Figure 4: Survey 1: Distribution of Subjects' Posterior Means of the Discount Factor (Dynamic MNL Model)

Note. The figure plots a histogram of the posterior means of the subjects' discount factors estimated using survey 1 data. The horizontal axis indicates the estimated discount factor and the vertical axis indicates the frequency of subjects.

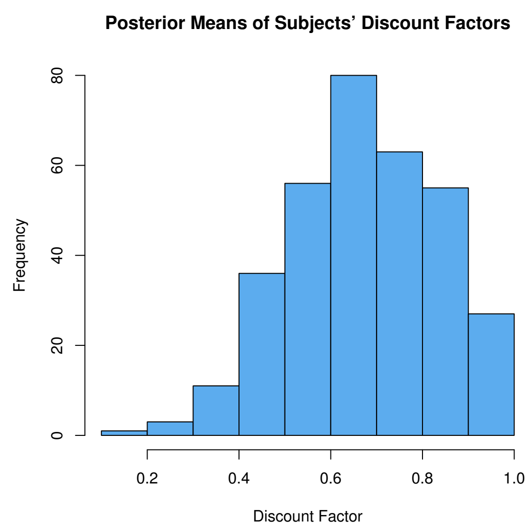


Figure 5: Survey 2: Distribution of Subjects' Posterior Means of the Discount Factor (Dynamic MNL Model)

Note. The figure plots a histogram of the posterior means of the subjects' discount factors estimated using survey 2 data. The horizontal axis indicates the estimated discount factor and the vertical axis indicates the frequency of subjects.

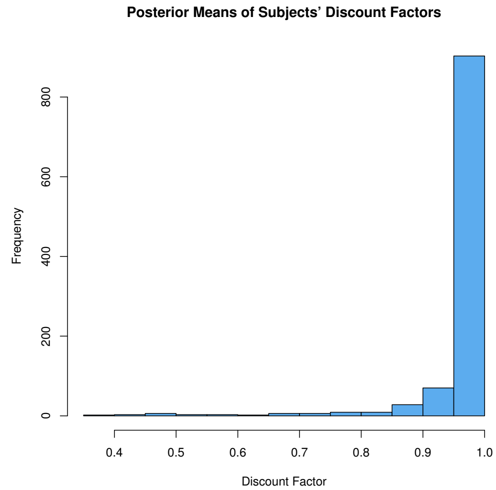


Figure 6: Survey 1: Distribution of Subjects' Posterior Means of the Hyperbolic Discount Factor β

Note. The figure plots a histogram of the posterior means of the subjects' hyperbolic discount factor estimated using survey 1 data. The hyperbolic discount factor is a measure of present bias. The horizontal axis indicates the estimated discount factor and the vertical axis indicates the frequency of subjects.

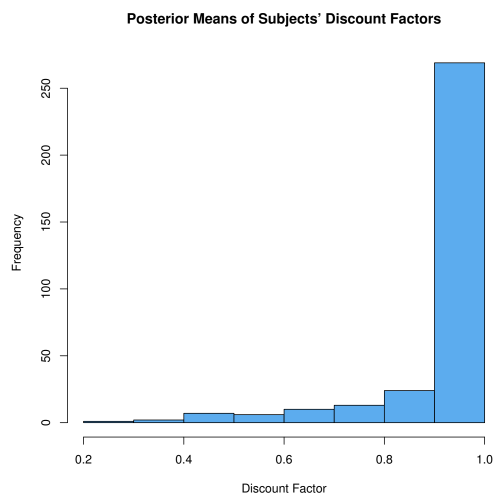


Figure 7: Survey 2: Distribution of Subjects' Posterior Means of the Hyperbolic Discount Factor β

Note. The figure plots a histogram of the posterior means of the subjects' hyperbolic discount factor estimated using survey 2 data. The hyperbolic discount factor is a measure of present bias. The horizontal axis indicates the estimated discount factor and the vertical axis indicates the frequency of subjects.

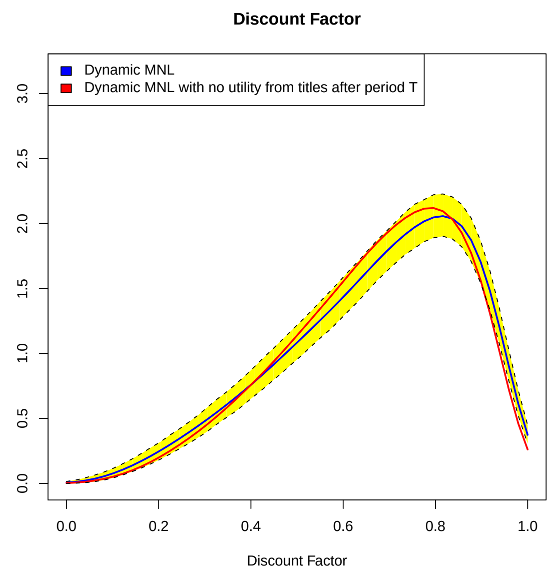


Figure 8: Robustness of the Discount Factor to Value of Titles

Note. The figure displays the pointwise posterior mean and 90% credibility region of the marginal density of the discount factor. The results are based on the Dynamic MNL model. For comparison purposes, we also show the results from the Dynamic MNL model with no utility from titles after period T.

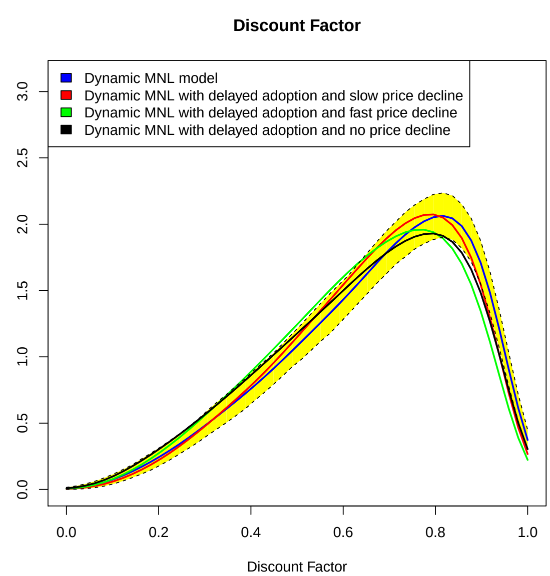


Figure 9: Robustness of the Discount Factor to Delayed Adoption

Note. The figure displays the pointwise posterior mean and 90% credibility region of the marginal density of the discount factor. The results are based on the Dynamic MNL model. For comparison purposes, we also show results from Dynamic MNL models with delayed adoption and slow, fast and no price decline.