The Japanese Economic Review
The Journal of the Japanese Economic Association

The Japanese Economic Review
Vol. 69, No. 3, September 2018

doi: 10.1111/jere.12188

IDENTIFICATION OF DYNAMIC MODELS OF REWARDS PROGRAMME

By ANDREW T. CHING† and MASAKAZU ISHIHARA‡

†University of Toronto ‡New York University

"Frequent-buyer" rewards programmes are commonly used by companies as a marketing tool to compete for market share. They provide a unique environment for studying consumers' forward-looking behaviour. The consumer's problem on accumulating reward points can be formulated as a stationary infinite horizon discrete choice dynamic programming model. We show that the parameters of this model, including the discount factor, are well-identified. In particular, it is possible to identify state-dependent discount factors (i.e. discount factors can vary with the number of reward points). We discuss how this identification result is related to the goal-gradient hypothesis studied in the consumer psychology literature.

JEL Classification Numbers: C11, C35, C61, D91, M31.

1. Introduction

"Frequent-buyer" rewards programmes are commonly used as a marketing tool by companies to compete for market shares. Because the benefits of accumulating reward points today can only be realised in the future, this provides a unique environment for studying consumers' forward-looking behaviour. In particular, the consumer's problem on accumulating reward points can be formulated as a stationary infinite horizon discrete choice dynamic programming (DDP) model, which provides a tractable way to capture how consumers evaluate the expected future payoffs of reward points.

Although several studies have applied this framework to study field data (e.g. Lewis, 2004; Hartmann and Viard, 2008; Kopalle *et al.*, 2012), the identification issue of dynamic models of rewards programme has not been studied. It should be pointed out that the previous studies have not estimated the discount factor. Instead, they have assumed that it is not identified, and have used annual interest rates to calibrate it.[1] This may seem surprising to researchers who do not specialise in estimating dynamic models because the discount factor is a crucial parameter which determines the extent to which consumers consider future payoffs. The reason why previous studies have typically set the discount factor according to the interest rate is mainly due to the negative nonparametric identification result demonstrated in Rust (1994). However, it is important to note a limitation of Rust (1994), who assumes that there is no state variable that satisfies the exclusion restriction assumption stated in Fang and Wang (2015).[2] Here, we argue that as long as we are willing to assume that the number of reward points accumulated does not enter the current period utility function, this state variable provides exclusion restrictions that help identify the discount factor. This is because in this set up, the

[1] This approach is popular in many empirical applications of dynamic programming models mainly because most of them do not have state variables that provide the exclusion restriction, which we will formally state later.

[2] We note that Fang and Wang's (2015) exclusion restriction is different from that of Magnac and Thesmar's (2002), which is less intuitive. Our exclusion restriction is consistent with that of Fang and Wang's (2015).

accumulated reward points in general only influence consumer choice through their expected future payoffs,[3] and the extent of the influence is determined by the discount factor. Therefore, the correlation between the choice probabilities and the accumulated reward points allows us to identify the discount factor.

The rest of the paper is organised as follows. In Section 2, we review related published literature. In Section 3, we present a dynamic store choice model with rewards to illustrate our argument. In Section 4, we discuss the model's identification issue. In Section 5, we explain how the model is estimated. In Section 6, we discuss how our results are related to the goal-gradient hypothesis in the consumer psychology literature. Section 7 concludes.

2. Literature review

In this section, we review three streams of related literature. First, our paper contributes to the literature on consumer behaviour in a reward programme environment. Early studies in economics and marketing examined the role of reward programmes in generating switching costs and the implications for price competition (e.g. Banerjee and Summers, 1987; Caminal and Mautes, 1990; Kim *et al.*, 2001; Caminal and Claici, 2007; Hartmann and Viard, 2008; Chen and Pearcy, 2010; Fong and Liu, 2011; Caminal, 2012). Recently, detailed customer-level reward programme data have led researchers to investigate a wider range of empirical issues related to reward programmes. For example, researchers have examined short-term and long-term effects of reward programmes on consumer purchase behaviour (e.g. Lewis, 2004; Liu, 2007), effects on retailers sales and profits (Lal and Bell, 2003), joint effects of frequency reward and customer tier components of a loyalty programme on sales (Kopalle *et al.*, 2012), and effects of goal achievement and failure on repeat purchases (Wang *et al.*, 2016).[4]

An important issue in empirical studies is the measurement of the effect of reward programmes (Breugelmans *et al.*, 2015). Because reward programmes operate as a dynamic incentive mechanism (Lewis, 2004), consumers' forward-looking behaviour is a necessary component of any model that attempts to measure the impact of reward programmes on consumer purchase behaviour. However, these previous empirical studies have not paid close attention to whether the discount factor can be identified; instead of estimating the discount factor, they all fix it according to the interest rate. Here, we show that the discount factor can be separately identified from the other parameters of this model.

Our research also contributes to the literature on the identification of the discount factor. Recent studies propose an exclusion restriction approach for identifying the discount factor (Magnac and Thesmar, 2002; Fang and Wang, 2015). Several empirical papers have used this approach for estimating the discount factor (e.g. Chevalier and Goolsbee, 2009; Lee, 2013; Chung *et al.*, 2014; Ching and Osborne, 2015; Ishihara and Ching, 2016).[5] Our paper contributes to the literature by showing that in a reward programme

[3] The exception is when it just hits the cutoff for receiving a reward.

[4] For a recent survey of the literature, see Breugelmans *et al.* (2015).

[5] Ching *et al.* (2014) combine exclusion restrictions with Geweke and Keane's (2000) approach to estimate consumers' expected future payoffs in a model of inventories, learning and category consideration, without imposing dynamic programming. In a sense, their approach is more general because they do not impose a dynamic programming solution to consumers. However, with less structure in their model, their approach cannot identify the discount factor.

context, we could identify state-dependent discount factors (i.e. the discount factor can be different for different levels of accumulated reward points). This result offers an opportunity to empirically study a connection between the rational forward-looking consumer theory in economics and the goal-gradient theory in consumer psychology.

The goal-gradient hypothesis argues that the motivation to achieve a goal changes as people get closer to the goal. The original hypothesis states that the motivation increases monotonically as the goal is neared (e.g. Hull, 1932; Lewin, 1938; Brown, 1948; Losco and Epstein, 1977; Brendl and Higgins, 1996). This hypothesis has been experimentally tested on both humans and animals (e.g. Hull, 1934), and in both single-goal and multiple-goal pursuits (e.g. Kivetz *et al.*, 2006; Louro *et al.*, 2007). In particular, Kivetz *et al.* (2006) use a field experiment at a coffee shop and test the hypothesis in a reward programme context. We argue that this increasing motivation to reach a goal can also be derived from a rational forward-looking consumer model with exponential discounting: as consumers get closer to the goal, the opportunity cost of not achieving the goal (and obtaining a reward) becomes higher, resulting in more "impatient" behaviour. This important connection has not been pointed out in the economics and psychology literature.

Moreover, our identification argument that a reward programme environment offers exclusion restrictions to identify state-dependent discount factors allows us to test whether the motivation, indeed, increases monotonically as the goal is neared. Bonezzi *et al.* (2011) show in a series of experiments that the motivation to reach a goal may change non-monotonically. They find that it can be higher when people are either far from or close to the goal, and lower when they are in the middle. In a rational forward-looking consumer model, standard exponential and hyperbolic discount factors do not allow for such a non-monotonic change in motivation, but state-dependent discount factors do.

3. The model

Suppose that there are J coffee shop chains in a city ($j = 1, \ldots, J$). Each chain offers a stamp card, which can be exchanged for a gift upon completion. The stamp card for a chain is valid for all stores in the same chain. Consumers receive one stamp for each visit at any store of a chain with a purchase.

Coffee shop chains' reward programmes differ in terms of: (i) the number of stamps required for a gift ($\bar{S}_j$); and (ii) the mean value of the gift (G_j). Consumers receive a gift in the same period that they complete the stamp card. Once consumers receive a gift, they will start with a blank stamp card again in the next period.

In each period, a consumer chooses which coffee shop chain to visit. Let $s_{ijt} \in \mathcal{S}_j \equiv \{0, 1, \ldots, \bar{S}_j - 1\}$ denote the number of stamps collected for chain j in period t before consumer i makes a decision. Note that s_{ijt} does not take the value $\bar{S}_j$ because of our assumption that consumers receive a gift in the same period that they complete the stamp card. The state space of this dynamic model is $\mathcal{S} \equiv \mathcal{S}_1 \times \cdots \times \mathcal{S}_J$.

Consumer i's single period utility of visiting coffee shop chain j in period t at $s_{it} = (s_{i1t}, \ldots, s_{iJt})$ is given by:

$$U_{ijt}(s_{it}, \epsilon_{ijt}) = \begin{cases} \alpha_j + \epsilon_{ijt} & \text{if } s_{ijt} < \bar{S}_j - 1 \\ \alpha_j + G_j + \epsilon_{ijt} & \text{if } s_{ijt} = \bar{S}_j - 1, \end{cases}$$

where α_j captures the brand equity for chain j, G_j is the valuation of the gift for chain j, and ϵ_{ijt} is the *i.i.d.* idiosyncratic random utility term. Note that we do not explicitly model prices here. However, one can interpret that ϵ_{ijt} captures prices which distribute *i.i.d.* across time and alternatives. We assume ϵ_{ijt} is unobserved to researchers and is type I extreme value distributed (location = Euler constant and scale = 1). In each period, consumers may choose not to visit a coffee shop. The single period mean utility of no shopping is normalised to zero (i.e. $U_{i0t}(s_{it}) = \epsilon_{i0t}$). Let the set of all options be denoted by $\mathcal{J} = \{0, 1, \ldots, J\}$.

Consumer i's objective at any time t is to maximise the sum of the present discounted future utility:

$$E\left[\max_{\{b_{i\tau}\}_{\tau=t}^{\infty}} \sum_{\tau=t}^{\infty} \beta^{\tau-t} \sum_{j=0}^{J} b_{ij\tau} \cdot U_{ij\tau}(s_{i\tau}, \epsilon_{ij\tau}) \middle| s_{it}, \epsilon_{ijt}\right],$$

where $b_{i\tau} = (b_{i1\tau}, \ldots, b_{iJ\tau})$; $b_{ij\tau} = 1$ if consumer i chooses chain j in period τ and $b_{ij\tau} = 0$ otherwise; β is the discount factor. The evolution of state, s_{it}, is deterministic and depends on consumers' choice. Given the state s_{ijt}, the next period state, s_{ijt+1}, is determined as follows:

$$s_{ijt+1} = \begin{cases} s_{ijt} + 1 & \text{if } s_{ijt} < \bar{S}_j - 1 \text{ and purchase at chain } j \text{ in period } t; \\ 0 & \text{if } s_{ijt} = \bar{S}_j - 1 \text{ and purchase at chain } j \text{ in period } t; \\ s_{ijt} & \text{if purchase at chain } -j \text{ or no shopping in period } t, \end{cases} \tag{1}$$

where $-j$ stands for the stores other than j.

The parameters of the model are $\{\alpha_j, G_j\}_{j=1}^{J}$ and β. It is well-known that there is a one-to-one relationship between the solution of the dynamic optimisation problem and the following functional equation (i.e. Bellman equation). Notice that this is a stationary dynamic optimisation problem because conditioning on the value of the state variables, the optimal decisions of the consumers do not depend on t. Therefore, we drop the subscript t hereafter. Let θ be the vector of parameters and E_ϵ denote expectation w.r.t. ϵ. The integrated value function, which captures the expected maximum of the alternative-specific value functions, is given by:[6]

$$\begin{aligned} V(s_i; \theta) &\equiv E_\epsilon \max_{j \in \{0,1,\ldots,J\}} \{V_j(s_i; \theta) + \epsilon_{ij}\} \\ &= \ln\left[\sum_{j=0}^{J} \exp(V_j(s_i; \theta))\right], \end{aligned} \tag{2}$$

where the second equality follows from the type I extreme value assumption on ϵ. The *alternative-specific value functions* obey the Bellman equation (Bellman, 1957):

[6] The literature also refers to it as the Emax function.

$$V_j(s_{ij}, s_{i-j}; \theta) = \begin{cases} \alpha_j + \beta V(s_{ij} + 1, s_{i-j}; \theta) & \text{if } s_{ij} < \bar{S}_j - 1, \\ \alpha_j + G_j + \beta V(0, s_{i-j}; \theta) & \text{if } s_{ij} = \bar{S}_j - 1, \end{cases} \tag{3}$$

$$V_0(s_i; \theta) = \beta V(s_i; \theta). \tag{4}$$

Let Γ_θ be the Bellman operator corresponding to the value function defined by Equations (2)–(4), where the subscript θ indicates that the operator is specific to θ. Let $\bar{U}_j(s_i;\, \theta) \equiv U_{ij}(s_i,\, \epsilon_{ij}) - \epsilon_{ij}$, and s_i' denote the vector of the number of stamps next period. Then for any arbitrary function, f,

$$(\Gamma_\theta)f(s_i) = E_\epsilon \max_j \{\bar{U}_j(s_i; \theta) + \epsilon_{ij} + \beta f(s_i')\}. \tag{5}$$

Note that the value function, $V(s_i;\, \theta)$, is a fixed point of Γ_θ (i.e. $V = \Gamma_\theta V$). Moreover, it can be shown that Γ_θ is a contraction mapping. This result is very useful because it implies that: (i) there is a unique fixed point of Γ_θ; and (ii) if we start off with any arbitrary initial guess of the value function, V^0, and recursively apply the Bellman operator to it (i.e. $V^{n+1} = \Gamma_\theta V^n$), then $V^n \to V$ uniformly. This result leads to the popular *method of successive approximation* to solve for the Bellman equation and agents' optimal decisions numerically.

The alternative specific value functions can be used to form choice probabilities and the likelihood of observed choice data. This model can be estimated using either the classical (see e.g. Aguirregabiria and Mira, 2010) or the Bayesian approach (see e.g. Ching *et al.*, 2012).

4. Identification

4.1 An intuitive explanation

The main source of dynamics is the intertemporal trade-off created by the rewards programme. Suppose that a consumer is closer to the completion of the stamp card for chain 1 but the price is lower in chain 2 today (captured by a low ϵ_2). If the consumer chooses chain 2 based on the lower price, he or she will delay the completion of the stamp card for chain 1. If the consumer takes the future into account, the delay will lower the present discounted value of the reward. Thus he/she will have an incentive to keep buying at chain 1 even though the price at chain 2 (i.e. captured by ϵ_2) is lower. Moreover, such an incentive should depend on the value of the discount factor.

This dynamic trade-off suggests that the variation of the empirical choice frequency of visiting coffee shop chains across states (i.e. the number of stamps collected) should allow us to pin down the discount factor. To illustrate this point, we consider the case of $J = 1$ and simulate the choice probabilities across s for different discount factors by setting $\alpha = -2$, $G = 3$ and $\bar{S} = 5$. Figure 1 shows how the choice probability of visiting the chain changes across states for different discount factors ($\beta = 0$, 0.5, 0.75, 0.9, 0.999). In general, we see that the choice probabilities increase with s. When β is small, the choice probabilities are relatively flat for small s, but become much higher as s approaches $\bar{S} - 1$. In the extreme case of $\beta = 0$, consumers only care about the current period utility. As a result, the choice probability of visiting the chain is flat for $s = 0, 1, 2, 3$ and goes up only when $s = 4$ (because the current period utility

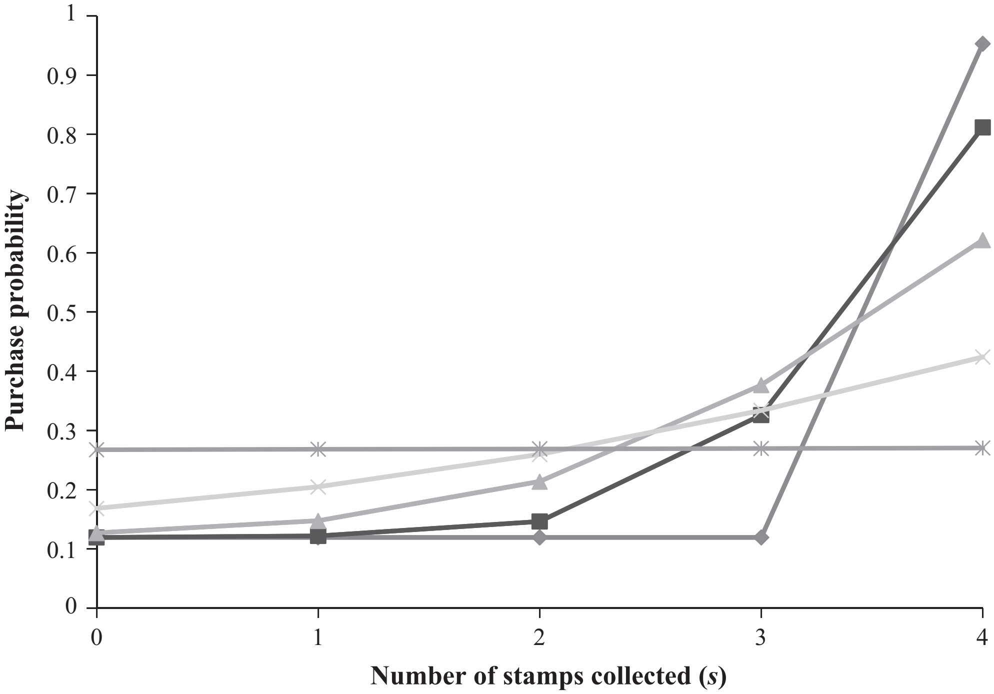

FIGURE 1. Choice probabilities across states for different discount factors. (—◆—) β = 0; (—■—) β = 0.5; (—▲—) β = 0.75; (—×—) β = 0.9; (—✳—) β = 0.999

of visiting the chain includes the value of the gift only when s = 4). As β increases, the choice probabilities become flatter as we move across s. As it approaches 1 (β = 0.999), the choice probabilities are essentially constant across states, and higher than those of β = 0 for s = 0, 1, 2, 3. This trend is due to the change of interactions of two counteracting forces: (i) the expected gain of obtaining an extra stamp today ($R(s, \beta)$); and (ii) the option value of waiting ($W(s, \beta)$). We define

$$R(s,\beta) \equiv \begin{cases} \beta V(s' = s+1) & \text{if } s < \bar{S} - 1, \\ G + \beta V(s' = 0) & \text{if } s = \bar{S} - 1. \end{cases}$$

$$W(s,\beta) \equiv \beta V(s' = s).$$

The choice probability of visiting a store is driven by the incentive of obtaining the gift sooner, which is measured by $R(s, \beta) - W(s, \beta)$. To understand the basic intuition, let's consider a case when β is small (say β = 0.5). When a consumer is very close to receiving a reward (i.e. s is close to $\bar{S} - 1$), $R(s, \beta)$ is much higher than $W(s, \beta)$ because waiting for an extra period reduces the expected discounted value of the gift significantly. However, when a consumer has very few stamps (i.e. s is close or equal to zero), both $R(s, \beta)$ and $W(s, \beta)$ are very small because the value of the gift is heavily discounted in both situations. Consequently, their difference also becomes smaller, and so is the incentive to obtain an extra stamp today. This explains why the increase in choice probability is relatively flat when s is small, but the choice probability increases sharply when it approaches $\bar{S}$.

Now let's consider the case of $\beta \to 1$. In this case, even when a consumer is close to receiving a reward, waiting for an extra period would hurt him/her very little, and, consequently, $R(s, \beta) - W(s, \beta)$ becomes smaller. This explains why the choice probability at $s = 4$ decreases as β increases. However, even if consumers are far away from receiving the reward, neither $R(s, \beta)$ nor $W(s, \beta)$ would be discounted much with β close to 1. As a result, $R(s, \beta) - W(s, \beta)$ remains fairly stable across s. This explains why the choice probabilities are essentially constant for all s. In Appendix I, we provide further explanations and a more formal argument for this result.

The above discussion suggests that unless the observed choice probabilities are flat across s, the overall shape of the choice probabilities across s will allow us to identify α, G and β. Two important aspects are: (i) changes in choice probability across s identify G and β; and (ii) the overall level of choice probabilities across s identifies α. If the choice probabilities are (almost) flat across s, then we could have either $G = 0$, or β is very close to 1. In practice, we expect that when β is close to 1, it could be quite hard to separately identify α and G because α shifts the choice probabilities equally across s, while G also shifts them almost equally across s.[7]

To illustrate the intuition that we discussed above further, Figure 2 shows how the choice probability of visiting the chain changes with β for any given s. In general, as we increase β, we have three observations: (i) when $s = 4$, the choice probability monotonically decreases; (ii) when $s < 3$, the choice probability monotonically increases; (iii) when $s = 3$, it first increases and then decreases. The discussion above has already explained observation (i). Observation (ii) indicates that $R(s, \beta) - W(s, \beta)$ always increases with β for small s. Observation (iii) shows an intermediate case: $R(s = 3, \beta)$ initially increases faster than $W(s = 3, \beta)$ as β increases; but when $\beta \to 1$, $W(s = 3, \beta)$ catches up. This explains why the choice probability of visiting the chain first increases and then decreases. Finally, $R(s, \beta)$ is always higher than $W(s, \beta)$ as long as $\beta > 0$.

Ching *et al.* (2012) use Monte Carlo experiments to provide evidence that even for a more complicated extension of the model presented here (they allow for unobserved heterogeneity in G_j, and explicitly incorporate the price variable and allow it to be serially correlated), all the structural parameters of the model, including the discount factor, can be recovered quite precisely for the case of $\beta = 0.6$ and 0.8.

4.2 Formal arguments for identification

The reason why we can identify the model parameters, especially the discount factor, β, is that the number of stamps, s, does not affect the current utility for $s < \bar{S} - 1$, but the expected future payoffs always depend on s as long as $G > 0$ and $\beta > 0$. Moreover, the expected future payoffs, obtained from the Bellman equation, are represented by a nonlinear function of β, α, G and s_i. Therefore, the variations of empirical choice frequencies of visiting a store and s_i allow us to pin down all the structural parameters, including β. Put another way, if $\beta = 0$ and the empirical choice frequencies vary systematically across s_i, for $s_i < S - 1$, it is not possible for the current per period utility alone to explain such a data pattern.

[7] Ching *et al.* (2012) provide Monte Carlo evidence that when β is large, it becomes harder to estimate α and G precisely (see p. 185 of their paper).

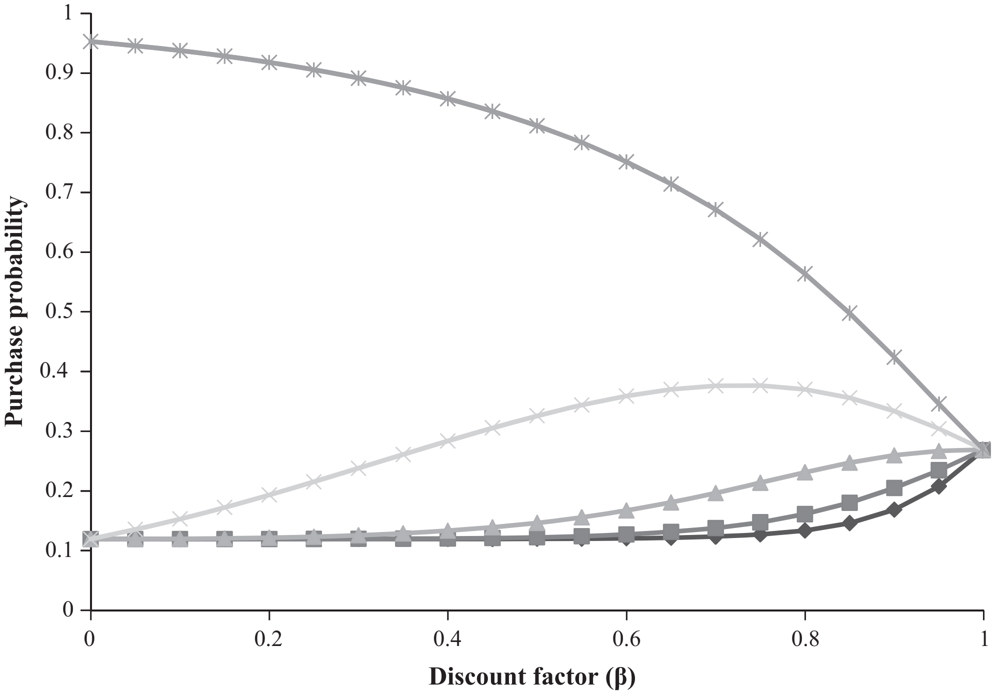

FIGURE 2. Choice probabilities for different discount factors across states. (◆) S = 0; (■) S = 1; (▲) S = 2; (×) S = 3; (✳) S = 4

This explains the intuition behind the exclusion restriction identification assumption. To formally show that the parameters are identified, we first note that our model satisfies the following identification assumptions, A1–A7. Following most previous research in identification, we assume A7 holds to avoid the issue of sampling error. A1–A5 and A7 are the same as those in Fang and Wang (2015):

A1. $U_{ijt}(s_{it}, \epsilon_{ijt})$ is additively separable in observable (i.e. $\bar{U}_j(s_{it})$) and unobservable components (i.e. ϵ_{ijt}).

A2. Stationarity (i.e. conditional on s_i, the optimal policy of the dynamic problem does not depend on t).

A3. Conditional Independence of ϵ (i.e. $p(s', \epsilon'|s, j, \epsilon) = q(\epsilon'|s')p(s'|s, j)$).

A4. ϵ follows Type I extreme value distribution.

A5. (Exclusion Restriction) There exist $s^1, s^2 \in \mathcal{S}$, with $s^1 \neq s^2$ such that: (i) for all j, $\bar{U}_j(s^1) = \bar{U}_j(s^2)$; and (ii) for some j, $f(s'|s^1, j) \neq f(s'|s^2, j)$, where $f(.|., j)$ is the transition density of s conditional on a consumer chooses j.

A6. $\bar{S} \geq 4$.

A7. There are an infinite number of observations for each state point, so that we can obtain the true choice probabilities for each state point from the data.

4.2.1 *Definition of identification*

We follow the definition of identification given by Magnac and Thesmar (2002) and Fang and Wang (2015). Let ϑ be the structure of the model given by

$$\vartheta = \{\beta, F, \{\bar{U}_j(s), V_j(s), j \in \mathcal{J}, s \in \mathcal{S}\}\},$$

where β is the discount factor, F is the distribution of ϵ_{ij}, $\bar{U}_j(s)$ is the mean of the per-period utility function for option j in state s, and $V_j(s)$ is the alternative-specific value function for option j in state s. Let Θ be the set of all possible structures that satisfy the model assumptions A1–A7. For any $\vartheta \in \Theta$, the predicted choice probability for option j in state s is given by

$$\hat{P}_j(s;\vartheta) = \Pr\{V_j(s) + \epsilon_j \geq V_k(s) + \epsilon_k, \forall k \neq j | s, \vartheta\}.$$

Two structures, $\vartheta, \vartheta' \in \Theta$, are said to be *observationally equivalent* if

$$\hat{P}_j(s;\vartheta) = \hat{P}_j(s;\vartheta') \ \forall j \in \mathcal{J}, s \in \mathcal{S}.$$

A model is said to be *identified* if and only if for any ϑ, $\vartheta' \in \Theta$, $\vartheta = \vartheta'$ if they are observationally equivalent.

4.2.2 Identification of ($\bar{U}_j(s)$, β)

Note that in our model, there are many s^l's that satisfy conditions A5(i) and A5(ii). For instance, just pick any s^l such that $s_j^l < \bar{S}_j - 1$, $\forall j$. Given that the transition from s to s' is deterministic in our model, it is clear that A5 is satisfied for all (s^1, s^2) with $s^1 \neq s^2$, and $s_j^l < \bar{S}_j - 1$ for all j. The fact that there are many combinations of state variables that satisfy the exclusion restriction implies that β is overidentified. One can then apply the results of subsection 3.2.1 of Fang and Wang (2015) to show that $\bar{U}_j$ is identified.

To illustrate the identification power of exclusion restrictions, we will now show that β can be expressed as a closed-form function of the observed choice probabilities. To keep the algebra simple, we will demonstrate it in a simplified version of our model with $J = 1$:

$$U_{i1}(s_{it}, \epsilon_{i1t}) = \begin{cases} \alpha_1 + \epsilon_{i1t} & \text{if } s_{it} < \bar{S} - 1 \\ \alpha_1 + G + \epsilon_{i1t} & \text{if } s_{it} = \bar{S} - 1 \end{cases}$$

$$U_{i0}(s_{it}, \epsilon_{i0t}) = \epsilon_{i0t}.$$

The formulation above satisfies A1 (i.e. the utility function is additively separable in observable and unobservable components). Given the state s_{it}, the next period state, s_{it+1}, is determined as follows:

$$s_{it+1} = \begin{cases} s_{it} + 1 & \text{if } s_{it} < \bar{S} - 1 \text{ and purchase in period } t; \\ 0 & \text{if } s_{it} = \bar{S} - 1 \text{ and purchase in period } t; \\ s_{it} & \text{no purchase in period } t. \end{cases} \quad (6)$$

This evolution process of the state s_{it} clearly satisfies A3 (Conditional Independence). By A2 (Stationarity), we can drop the time subscript when we write down the value fuction. The alternative specific value function becomes:

$$V_j(s) = \bar{U}_j(s) + \beta V(s'), \tag{7}$$

where $\bar{U}_1(s) = \alpha_1$ if $s < \bar{S} - 1$ and $\bar{U}_1(\bar{S} - 1) = \alpha_1 + G$; $\bar{U}_0(s) = 0, \forall s$, and by A4,

$$\begin{aligned} V(s) &= E_\epsilon \max_j \{V_j(s) + \epsilon_j\} \\ &= \ln\left[\sum_{j=0}^{1} \exp(V_j(s))\right] \\ &= V_0(s) + \ln(1 + \exp(V_1(s) - V_0(s))). \end{aligned} \tag{8}$$

Let $P_j(s)$ be the choice probability for alternative j at state s:

$$P_j(s) = Prob\{V_j(s) + \epsilon_j \geq V_k(s) + \epsilon_k, k = 0, 1\}.$$

Because we assume that ϵ_{ijt} is *i.i.d.* type I extreme value distributed (A4),

$$\begin{aligned} P_j(s) &= \frac{\exp(V_j(s))}{\exp(V_0(s)) + \exp(V_1(s))}, \\ \ln\left(\frac{P_1(s)}{P_0(s)}\right) &= V_1(s) - V_0(s). \end{aligned} \tag{9}$$

This equation illustrates the inversion theorem by Hotz and Miller (1993). The theorem shows that under fairly general conditions, one can map the choice probabilities from a dynamic model to the difference in alternative specific value functions. More specifically, let $j \in \{0, \ldots, J\}$ be the set of alternatives, and $P(s) = (P_0(s), \ldots, P_J(s))$. Then the inverse theorem states that $f_j(P(s)) = V_j(s) - V_0(s)$, where 0 is the reference alternative.

Recall that $J = 1$ in our example. Substituting Equation (9) into Equation (8) gives:

$$V(s) = V_0(s) + \ln\left(1 + \frac{P_1(s)}{P_0(s)}\right). \tag{10}$$

Note that if we know $V_0(s)$, $V(s)$ is identified. Substituting Equation (10) into Equation (7) for $j = 0$, we obtain (recall that $s' = s$ if a consumer chooses no purchase),

$$\begin{aligned} V_0(s) &= \beta V(s) \\ &= \beta V_0(s) + \beta \ln\left(1 + \frac{P_1(s)}{P_0(s)}\right) \\ &= \frac{\beta}{1 - \beta} \ln\left(1 + \frac{P_1(s)}{P_0(s)}\right). \end{aligned} \tag{11}$$

Substituting Equation (11) into Equation (9), we obtain

$$V_1(s) = \ln\left(\frac{P_1(s)}{P_0(s)}\right) + \frac{\beta}{1-\beta}\ln\left(1+\frac{P_1(s)}{P_0(s)}\right). \tag{12}$$

Substituting Equation (11) into Equation (10), we obtain

$$\begin{aligned} V(s) &= \frac{\beta}{1-\beta}\ln\left(1+\frac{P_1(s)}{P_0(s)}\right) + \ln\left(1+\frac{P_1(s)}{P_0(s)}\right) \\ &= \frac{1}{1-\beta}\ln\left(1+\frac{P_1(s)}{P_0(s)}\right). \end{aligned} \tag{13}$$

Now, let's rewrite Equation (7) for $j = 1$ as

$$\bar{U}_1(s) = V_1(s) - \beta V(s'). \tag{14}$$

By A6 ($\bar{S} \geq 4$). It follows that

$$\bar{U}_1(1) = V_1(1) - \beta V(2), \tag{15}$$

$$\bar{U}_1(2) = V_1(2) - \beta V(3). \tag{16}$$

By A5 (Exclusion Restriction), $\bar{U}_1(s=1) = \bar{U}_1(s=2)$ (this is one exclusion restriction). Hence, the two equations above imply

$$V_1(1) - V_1(2) = \beta(V(2) - V(3)). \tag{17}$$

It follows from Equations (12) and (13) that

$$\begin{aligned} &\left[\ln\left(\frac{P_1(1)}{P_0(1)}\right) - \ln\left(\frac{P_1(2)}{P_0(2)}\right)\right] + \frac{\beta}{1-\beta}\left[\ln\left(1+\frac{P_1(1)}{P_0(1)}\right) - \ln\left(1+\frac{P_1(2)}{P_0(2)}\right)\right] \\ &\qquad = \frac{\beta}{1-\beta}\left[\ln\left(1+\frac{P_1(2)}{P_0(2)}\right) - \ln\left(1+\frac{P_1(3)}{P_0(3)}\right)\right]. \end{aligned} \tag{18}$$

Let $a \equiv \ln(\frac{P_1(1)}{P_0(1)}) - \ln(\frac{P_1(2)}{P_0(2)})$, $b \equiv \ln(1+\frac{P_1(1)}{P_0(1)}) - \ln(1+\frac{P_1(2)}{P_0(2)})$, $c \equiv \ln(1+\frac{P_1(2)}{P_0(2)}) - \ln(1+\frac{P_1(3)}{P_0(3)})$. Then, we rewrite the above equation as:

$$\begin{aligned} a + \frac{\beta}{1-\beta}b &= \frac{\beta}{1-\beta}c \\ \beta &= \frac{a}{a+c-b}. \end{aligned} \tag{19}$$

Because a, b and c are all functions of choice probabilities $P_j(s)$, which we can obtain (by A7) based on empirical choice frequencies at $s = 1, 2, 3$, it follows from Equation (19) that β is identified.

With β, we can now use Equation (15) to identify α_1 because it identifies $\bar{U}_1(1) = \alpha_1$. (Recall that $V_1(s)$ and $V(s)$ can be identified from choice probabilities; see Equations (12) and (13).)

Similarly, we can identify $\bar{U}_1(\bar{S} - 1)$ by using Equation (20):

$$\bar{U}_1(\bar{S} - 1) = V_1(\bar{S} - 1) - \beta V(0). \tag{20}$$

Note that $\bar{U}_1(\bar{S} - 1) = \alpha_1 + G$, and we have just shown how to identify α_1. Hence, G is also identified.

In fact, β could be overidentified. For instance, if $\bar{S} \geq 5$, we can use $s = 2$ and $s = 3$ to derive another exclusion restriction. We can go through the same arguments to derive another equation similar to Equation (19) with the distinction that it depends on the choice probabilities at $s = 2, 3, 4$. In general, the β's derived from these two equations may not be the same, and, therefore, so long as we have more than one exclusion restriction, β is overidentified.

5. Estimation

This section briefly describes how to estimate the model in Section 3 using the maximum likelihood estimation. Suppose we observe agents' choice and their corresponding state for T periods, $\{\mathsf{a}_i, \mathsf{s}_i\}_{i=1}^{I}$, where $\mathsf{a}_i = \{\mathsf{a}_{it}\}_{t=1}^{T}$, and $\mathsf{s}_i = \{\mathsf{s}_{it}\}_{t=1}^{T}$. To estimate the vector of model parameters, $\theta \equiv (\alpha_j, G_j, \beta)$, we need to construct the likelihood of the observed data. The likelihood will take the form of products of choice probabilities, $P_j(\mathsf{s}_{it}; \theta)$. The probability of observing agent i choosing j given the observed state, s_{it}, is given by

$$\begin{aligned} P_j(\mathsf{s}_{it}; \theta) &= P(\epsilon : V_j(\mathsf{s}_{it}, \epsilon_{ijt}; \theta) \geq V_l(\mathsf{s}_{it}, \epsilon_{ilt}; \theta), \forall l) \\ &= \frac{\exp(V_j(\mathsf{s}_{it}; \theta))}{\sum_{l=0}^{J} \exp(V_l(\mathsf{s}_{it}; \theta))}. \end{aligned}$$

Then the likelihood increment for agent i at time t is:

$$L_{it}(\mathsf{a}_{it} | \mathsf{s}_{it}; \theta) = P_j(\mathsf{s}_{it}; \theta) \text{ if } \mathsf{a}_{it} = j.$$

The likelihood of the observed data is

$$L(\mathsf{a} | \mathsf{s}; \theta) = \prod_i \prod_t L_{it}(\mathsf{a}_{it} | \mathsf{s}_{it}; \theta),$$

where $\mathsf{a} = (\mathsf{a}_1, \ldots, \mathsf{a}_I), \mathsf{s} = (\mathsf{s}_1 \ldots, \mathsf{s}_I)$. In the maximum likelihood estimation, we search for θ that maximises $\ln L(\mathsf{a} | \mathsf{s}; \theta)$.

The above likelihood construction requires the computation of the (integrated) value functions in Equation (2). In the standard nested fixed-point approach (Rust, 1987), we use the method of successive approximation for solving the value functions; that is, we iteratively apply the Bellman operator in Equation (5) until convergence. However, the method of successive approximation is known to be slow and several approaches for reducing the computational burden have been proposed (see e.g. Aguirregabiria and Mira, 2010). Furthermore, the above model can be estimated using Bayesian estimation. Ching *et al.* (2012) provide detailed procedures for estimating the above reward programme model using the Bayesian Markov chain Monte Carlo algorithm proposed by Imai *et al.* (2009).

6. Goal-gradient hypothesis and state-dependent discount factors

The well-known goal-gradient hypothesis, first proposed by Hull (1932), has been tested using animals and humans (e.g. Hull, 1934; Kivetz *et al.*, 2006). So far, the literature makes use of the testable implication that the subject has the tendency to make choices to reach the goal faster, as they get closer to it. In particular, Kivetz *et al.* (2006) find that: (i) participants in a real cafe reward programme purchase coffee more frequently the closer they are to earning a free coffee; and (ii) Internet users who rate songs in return for reward certificates visit the rating website more often and rate more songs per visit as they approach the cutoff of earning the certificate. These two results are qualitatively consistent with the prediction of our rational dynamic model with a moderate fixed discount factor, as shown above.[8] As far as we know, this is the first paper that explicitly uses a model to show the connection between the rational forward-looking consumer theory in economics and the goal-driven theory in consumer psychology.[9]

It is important to point out that the goal-gradient theory emphasises the role of motivation. In our model, the parameter that corresponds to motivation well is the discount factor. In particular, we feel that a more direct interpretation of the goal-gradient hypothesis is that an agent becomes more impatient when they are closer to receiving a reward (or reaching a goal). The identification argument that we discussed above can be extended to conduct such a “direct” test of the goal-gradient hypothesis, which is more powerful and has not been done before. Suppose that $\bar{S} = 10$. Then we can use the above arguments to construct 8 exclusion restrictions associated with: (i) $(s^1 = 0, s^2 = 1, s^3 = 2)$, (ii) $(s^1 = 1, s^2 = 2, s^3 = 3)$, (iii) $(s^1 = 2, s^2 = 3, s^3 = 4)$, ..., (viii) $(s^1 = 7, s^2 = 8, s^3 = 9)$. Each exclusion restriction will produce one estimate of β. One caveat of the above identification approach is that we need to assume that the discount factor associated with two state points is the same. Therefore, without loss of generality, we denote $\beta(s^1)$. One can apply this identification strategy to the data, and if we find that $\beta(s^1)$ decreases with s^1, this provides support for Hull’s goal-gradient hypothesis. More recent research also finds evidence that the motivation to engage in goal-consistent behaviour can be higher when subjects are either far from or close to the goal/reward and lower when they are halfway (Bonezzi *et al.*, 2011). If we find that $\beta(s^1)$ is inverted U-shaped, this will provide support to the psychophysical model proposed by Bonezzi *et al.* (2011). The pattern of $\beta(s^1)$’s can potentially help us to improve our understanding of how goals/rewards motivate consumer behaviour.

---

8 Kivetz *et al.* (2006) also provide a third result to support the illusionary goal progress hypothesis by giving subjects two bonus stamps on a stamp card and requiring two more stamps to receive a free coffee. They find that the treatment group finishes the stamp card faster than the control group, who did not receive bonus stamps, and require the same number of absolute stamps to complete the card. This prediction is consistent with an alternative explanation that bonus stamps build brand loyalty, and this will cause the brand intercept term to shift up in our model.

9 Kivetz *et al.* (2006) recognise that these two implications may be consistent with a rational model, but they do not explicitly derive such a model.

7. Conclusion

This paper shows that when consumers are rational and forward-looking, the "frequent-buyer" type of rewards programme environment can allow us to recover some fundamental structural parameters that capture consumer preferences. In particular, we show that the discount factor is also identified, which is not always the case for DDP models. Therefore, by observing consumer choice in this environment, it is possible to measure how consumers make intertemporal trade-offs.

Given that rewards programmes are commonly used, the consumer databases maintained by these companies provide a unique opportunity for us to study consumer intertemporal choice.[10] However, we should also note that there are limitations of our identification arguments. In a more realistic model, the error term of the model could be serially correlated. For instance, if consumer choice is influenced by state dependence and such choice dynamics is not modelled properly,[11] this could lead to serial correlation in the error term, and, hence, the conditional independence assumption (A3) would be violated. If consumer preferences change in a hidden-Markov fashion, this would also lead to violation of conditional independence. Under those circumstances, it is likely that we need additional exclusion restrictions for identification. We leave it for future research to investigate the identification in such an environment.

Finally, in light of our identification results, it should be highlighted that empirical researchers may not want to simply fix the discount factor according to the interest rate. Instead, the extent to which consumers discount the future payoffs is an empirical research question, which could be answered by the data if the dynamic problem at hand has natural exclusion restrictions. This is an important research question because if consumers do not discount the future payoffs based on the prevailing interest rate, incorrectly normalising the discount factor using this traditional approach may very well lead to biased estimates for other preference parameters. Being able to measure consumers' structural parameters of intertemporal trade-offs is crucial for marketing management. This goes beyond the rewards programme environment analysed here. Other examples of economics and marketing applications include: (i) stockpiling when a storable product is on sale; (ii) experimenting with a new product with uncertain quality; and (iii) durable good or new technology adoption. We hope more future research will make use of exclusion restrictions naturally arising from the institutional setting to estimate dynamic models.[12]

---

10 Rossi (2017) applies our identification argument to estimate the discount factor of consumers who participate in a retail gasoline chain's reward programme, which is similar to the one we consider here.

11 One type of state dependence is consumer learning behaviour; see Ching *et al.* (2013, 2017) for recent surveys.

12 Other approaches have also been used. Dubé *et al.* (2014) and Yang and Ching (2014) use the stated preference approach to estimate and calibrate the discount factor, respectively. Yao *et al.* (2012) use data from a static environment to estimate the current payoffs first, and then use data from a dynamic environment to estimate the discount factor; their approach is in similar spirit to Geweke and Keane (2000).

Appendix I

We provide more formal arguments about the behaviour of $R(s, \beta) - W(s, \beta)$ using the simple dynamic store choice model with only one store presented in Section 3.

To illustrate how $R(s, \beta) - W(s, \beta)$ changes across s, we first consider consumers' purchase decision today under the assumption that they will visit the store in every period from tomorrow on. Under this assumption, the incentive to earn an extra stamp today, $R(s, \beta) - W(s, \beta)$, will have a simple and intuitive expression.

Suppose that $s = 0$. Then if he/she visits the store today, the present discounted value, $R(s = 0, \beta)$, will be

$$\begin{aligned}R(s=0,\beta) =&\alpha+\beta\alpha+\beta^2\alpha+\beta^3\alpha+\beta^4(\alpha+G)+\beta^5\alpha+\beta^6\alpha+\beta^7\alpha+\beta^8\alpha\\&+\beta^9(\alpha+G)+\cdots.\end{aligned}$$

If he/she does not visit the store today, then the present discounted value, $W(s = 0, \beta)$, will be

$$\begin{aligned}W(s=0,\beta) =&0+\beta\alpha+\beta^2\alpha+\beta^3\alpha+\beta^4\alpha+\beta^5(\alpha+G)+\beta^6\alpha+\beta^7\alpha+\beta^8\alpha+\beta^9\alpha\\&+\beta^{10}(\alpha+G)+\cdots.\end{aligned}$$

Thus,

$$\begin{aligned}R(s=0,\beta)-W(s=0,\beta)&=\alpha+\beta^4(1-\beta)G+\beta^9(1-\beta)G+\cdots\\&=\alpha+\beta^4(1-\beta)(1+\beta^5+\beta^{10}+\cdots)G\\&=\alpha+\frac{\beta^4(1-\beta)}{(1-\beta)(1+\beta+\beta^2+\beta^3+\beta^4)}G\\&=\alpha+\frac{\beta^4}{1+\beta+\beta^2+\beta^3+\beta^4}G.\end{aligned}$$

It is easy to verify that for any s,

$$R(s,\beta)-W(s,\beta)=\alpha+\frac{\beta^{\bar{S}-1-s}}{\sum_{k=0}^{\bar{S}-1}\beta^k}G. \tag{21}$$

This equation implies that:

Result 1: *When* $0 < \beta < 1$, *(i)* $R(s, \beta) - W(s, \beta)$ *increases with* s; *(ii)* $\frac{\partial(R(s,\beta)-W(s,\beta))}{\partial s}$ *increases with* s.

It follows from (i) that the choice probability increases with s. It follows from (ii) that the increase in choice probability is relatively flat when s is small, but increases sharply when s approaches $\bar{S}$. The expression above also demonstrates why the shape of the choice probabilities across s would identify β and G. As long as $G > 0$ and $\beta > 0$, it is not possible to find two different combinations of (β, G) that give the same value of $R(s, \beta) - W(s, \beta)$, $\forall s$.

Finally, Equation (21) implies that

Result 2: *As $\beta \to 1$, we have*

$$R(s, \beta) - W(s, \beta) \to \alpha + \frac{G}{\bar{S}}, \ \forall s.$$

Thus, when β approaches one, we observe a flat choice probability across s. In addition, it is clear that it becomes hard to separately identify α and G as β approaches one.

Now we provide a formal proof for the convergence of $R(s, \beta) - W(s, \beta)$ under the extreme value distribution assumption on the unobserved state variable.

Result 3: *Assume that the unobserved state variable, ϵ, follows the extreme value distribution, and $\bar{S} = 2$. Then, as $\beta \to 1$, $(R(s, \beta) - W(s, \beta))$ converges to $\alpha + \frac{G}{2}$, $\forall s$.*

Proof. For $s = 0$, the Bellman equation is given by

$$V(s=0) = E_\epsilon \max\{V_0(s=0) + \epsilon_0, V_1(s=0) + \epsilon_1\},$$

where

$$\begin{aligned} V_1(s=0) &= \alpha + \beta V(s=1) \\ V_0(s=0) &= \beta V(s=0). \end{aligned}$$

For s=1, the Bellman equation is given by

$$V(s=1) = E \max\{V_0(s=1) + \epsilon_0, V_1(s=1) + \epsilon_1\},$$

where

$$\begin{aligned} V_1(s=1) &= \alpha + G + \beta V(s=0) \\ V_0(s=1) &= \beta V(s=1). \end{aligned}$$

Note first that

$$\begin{aligned} R(s=0, \beta) - W(s=0, \beta) &= V_1(s=0) - V_0(s=0) \\ &= \alpha + \beta(V(s=1) - V(s=0)). \\ R(s=1, \beta) - W(s=1, \beta) &= V_1(s=1) - V_0(s=1) \\ &= \alpha + G - \beta(V(s=1) - V(s=0)). \end{aligned}$$

Define $\Delta \equiv V(s=1) - V(s=0)$. If we assume that ϵ follows the extreme value distribution, then we have

$$\begin{aligned}
&\Delta = \ln(\exp(V_0(s=1)) + \exp(V_1(s=1))) - \ln(\exp(V_0(s=0)) + \exp(V_1(s=0))) \\
\Leftrightarrow\ &\Delta = \ln(1 + \exp(V_1(s=1) - V_0(s=1))) + V_0(s=1) \\
&\qquad - \ln(1 + \exp(V_1(s=0) - V_0(s=0))) - V_0(s=0) \\
\Leftrightarrow\ &\Delta = \ln(1 + \exp(\alpha + G - \beta\Delta)) - \ln(1 + \exp(\alpha + \beta\Delta)) + \beta\Delta \\
\Leftrightarrow\ &\exp((1-\beta)\Delta) = \frac{1 + \exp(\alpha + G - \beta\Delta)}{1 + \exp(\alpha + \beta\Delta)}.
\end{aligned}$$

Now note that when $\beta \to 1$, the left-hand side will approach one. Thus, we have $G - 2\beta\Delta \to 0$, or $\Delta \to \frac{G}{2}$. Therefore, as $\beta \to 1$, $R(s, \beta) - W(s, \beta)$ converges to $\alpha + \frac{G}{2}$ for all s. ■

Final version accepted 11 January 2018.

REFERENCES

Aguirregabiria, V. and P. Mira (2010) "Dynamic Discrete Choice Structural Models: A Survey", *Journal of Econometrics*, Vol. 156, No. 1, pp. 38–67.

Banerjee, A. V. and L. H. Summers (1987) "On Frequent Flyer Programs and Other Loyalty-Inducing Economic Arrangements", Discussion Paper No. 1337, Harvard Institute of Economic Research, Harvard University, Cambridge.

Bellman, R. (1957) *Dynamic Programming*, Princeton, NJ: Princeton University Press.

Bonezzi, A., C. M. Brendl and M. D. Angelis (2011) "Stuck in the Middle: The Psychophysics of Goal Pursuit", *Psychological Science*, Vol. 22, No. 5, pp. 607–612.

Brendl, C. M. and E. T. Higgins (1996) "Principles of Judging Valence: What Makes Events Positive or Negative?", *Advances in Experimental Social Psychology*, Vol. 28, pp. 95–160.

Breugelmans, E., T. H. Bijmolt, J. Zhang, L. J. Basso, M. Dorotic, P. Kopalle, A. Minnema, W. J. Mijnlieff and N. V. Wünderlich (2015) "Advancing Research on Loyalty Programs: A Future Research Agenda", *Marketing Letters*, Vol. 26, No. 2, pp. 127–139.

Brown, J. S. (1948) "Gradients of Approach and Avoidance Responses and Their Relation to Level of Motivation", *Journal of Comparative and Physiological Psychology*, Vol. 41, No. 6, pp. 450–465.

Caminal, R. (2012) "The Design and Efficiency of Loyalty Rewards", *Journal of Economics & Management Strategy*, Vol. 21, No. 2, pp. 339–371.

——— and A. Claici (2007) "Are Loyalty-Rewarding Pricing Schemes Anti-Competitive?", *International Journal of Industrial Organization*, Vol. 25, No. 4, pp. 657–674.

——— and C. Mautes (1990) "Endogenous Switching Costs in a Duopoly Model", *International Journal of Industrial Organization*, Vol. 8, No. 3, pp. 353–373.

Chen, Y. and J. Pearcy (2010) "Dynamic Pricing: When to Entice Brand Switching and When to Reward Customer Loyalty", *RAND Journal of Economics*, Vol. 41, No. 4, pp. 674–685.

Chevalier, J. and A. Goolsbee (2009) "Are Durable Goods Consumers Forward-Looking? Evidence from College Textbooks", *Quarterly Journal of Economics*, Vol. 124, No. 4, pp. 1853–1884.

Ching, A. T. and M. Osborne (2015) "Identification and Estimation of Forward-looking Behavior: The Case of Consumer Stockpiling", Working Paper, Rotman School of Management, University of Toronto, Toronto, ON.

———, S. Imai, M. Ishihara and N. Jain (2012) "A Practitioner's Guide to Bayesian Estimation of Discrete Choice Dynamic Programming Models", *Quantitative Marketing and Economics*, Vol. 10, No. 2, pp. 151–196.

———, T. Erdem and M. P. Keane (2013) "Learning Models: An Assessment of Progress, Challenges, and New Developments", *Marketing Science*, Vol. 32, No. 6, pp. 913–938.

———, ——— and ——— (2014) "A Simple Method to Estimate the Roles of Learning, Inventories and Category Consideration in Consumer Choice", *Journal of Choice Modelling*, Vol. 13, pp. 60–72.

———, ——— and ——— (2017) "Empirical Models of Learning Dynamics: A Survey of Recent Developments", in B. Wierenga and R. van der Lans, eds, *Handbook of Marketing Decision Models*, Cham, Switzerland: Springer, pp. 223–257.

Chung, D. J., T. Steenburgh and K. Surdhir (2014) "Do Bonuses Enhance Sales Productivity? A Dynamic Structural Analysis of Bonus-Based Compensation Plans", *Marketing Science*, Vol. 33, No. 2, pp. 165–187.

Dubé, J.-P., G. Hitsch and P. Jindal (2014) "The Joint Identification of Utility and Discount Functions from stated choice data: An application to durable goods adoption", *Quantitative Marketing and Economics*, Vol. 12, No. 4, pp. 331–377.

Fang, H. and Y. Wang (2015) "Estimating Dynamic Discrete Choice Models with Hyperbolic Discounting, with an Application to Mammography Decisions", *International Economic Review*, Vol. 56, No. 2, pp. 565–596.

Fong, Y-F. and Q. Liu (2011) "Loyalty Rewards Facilitate Tacit Collusion", *Journal of Economics & Management Strategy*, Vol. 20, No. 3, pp. 739–775

Geweke, J. F. and M. P. Keane (2000) "Bayesian Inference for Dynamic Discrete Choice Models Without the Need for Dynamic Programming", in R. Mariano, T. Schuermann and M. J. Weeks, eds, *Simulation Based Inference and Econometrics: Methods and Applications*, Cambridge, UK: Cambridge University Press, pp. 100–131.

Hartmann, W. R. and V. B. Viard (2008) "Do Frequency Reward Programs Create Switching Costs? A Dynamic Structural Analysis of Demand in a Reward Program", *Quantitative Marketing and Economics*, Vol. 6, No. 2, pp. 109–137.

Hotz, J. V. and R. Miller (1993) "Conditional Choice Probabilities and the Estimation of Dynamic Models", *Review of Economic Studies*, Vol. 60, No. 3, pp. 497–529.

Hull, C. L. (1932) "The Goal-Gradient Hypothesis and Maze Learning", *Psychological Review*, Vol. 39, No. 1, pp. 25–43.

——— (1934) "The Rats' Speed of Locomotion Gradient in the Approach to Food", *Journal of Comparative Psychology*, Vol. 17, No. 3, pp. 393–422.

Imai, S., N. Jain and A. Ching (2009) "Bayesian Estimation of Dynamic Discrete Choice Models", *Econometrica*, Vol. 77, No. 6, pp. 1865–1899.

Ishihara, M. and A. T. Ching (2016) "Dynamic Demand for New and Used Durable Goods without Physical Depreciation: The Case of Japanese Video Games", Working Paper, Stern School of Business, New York University, New York.

Kim, B.-D., M. Shi and K. Srinivasan (2001) "Reward Programs and Tacit Collusion", *Marketing Science*, Vol. 20, No. 2, pp. 99–120.

Kivetz, R., O. Urminsky and Y. Zheng (2006) "The Goal-Gradient Hypothesis Resurrected: Purchase Acceleration, Illusionary Goal Progress, and Customer Retention", *Journal of Marketing Research*, Vol. 43, No. 1, pp. 39–58.

Kopalle, P. K., S. A. Neslin, B. Sun, Y. Sun and V. Swaminathan (2012) "The Joint Sales Impact of Frequency Reward and Customer Tier Components of Loyalty Programs", *Marketing Science*, Vol. 31, No. 2, pp. 216–235.

Lal, R. and D. E. Bell (2003) "The Impact of Frequent Shopper Programs in Grocery Retailing", *Quantitative Marketing and Economics*, Vol. 1, No. 2, pp. 179–202.

Lee, R. S. (2013) "Vertical Integration and Exclusivity in Platform and Two-Sided Markets", *American Economic Review*, Vol. 103, No. 7, pp. 2960–3000.

Lewin, K. (1938) *The Conceptual Representation and the Measurement of Psychological Forces*, Durham, NC: Duke University Press.

Lewis, M. (2004) "The Influence of Loyalty Programs and Short-Term Promotions on Customer Retention", *Journal of Marketing Research*, Vol. 41, No. 3, pp. 281–292.

Liu, Y. (2007) "The Long-Term Impact of Loyalty Programs on Consumer Purchase Behavior and Loyalty", *Journal of Marketing*, Vol. 71, No. 4, pp. 19–35.

Losco, J. and S. Epstein (1977) "Relative Steepness of Approach and Avoidance Gradients as a Function of Magnitude and Valence of Incentive", *Journal of Abnormal Psychology*, Vol. 86, No. 4, pp. 360–368.

Louro, M. J., R. Pieters and M. Zeelenberg (2007) "Dynamics of Multiple-Goal Pursuit", *Journal of Personality and Social Psychology*, Vol. 93, No. 2, pp. 174–193.

Magnac, T. and D. Thesmar (2002) "Identifying Dynamic Discrete Decision Processes", *Econometrica*, Vol. 70, No. 2, pp. 801–816.

Rossi, F. (2017) "Lower Price or Higher Reward? Measuring the Effect of Consumers' Preferences on Reward Programs", *Management Science*, https://doi.org/10.1287/mnsc.2017.2821 (forthcoming)

Rust, J. (1987) "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher", *Econometrica*, Vol. 55, No. 5, pp. 999–1033.

——— (1994) "Structural Estimation of Markov Decision Processes", in R. Engle and D. McFadden, eds, *Handbook of Econometrics*, Amsterdam: North Holland, pp. 3081–3143.

Wang, Y., M. Lewis, C. Cryder and J. Sprigg (2016) "Enduring Effects of Goal Achievement and Failure Within Customer Loyalty Programs: A Large-Scale Field Experiment", *Marketing Science*, Vol. 35, No. 4, pp. 565–575.

Yang, B. and A. T. Ching (2014) "Dynamics of Consumer Adoption of Financial Innovation: The Case of ATM Cards", *Management Science*, Vol. 60, No. 4, pp. 903–922.

Yao, S., C. F. Mela, J. Chiang and Y. Chen (2012) "Determining Consumers' Discount Rates with Field Studies", *Journal of Marketing Research*, Vol. 49, No. 6, pp. 822–841.