Potential Methods of Structural Estimation of Loewenstein's Anticipatory Utility Model

Saul A. Rosen

I. Introduction

In his 1987 paper, Anticipation and Valuation of Delayed Consumption[1], George Loewenstein outlines a model of utility discounting to account for an anomalous result in experimental data. Contrary to expectations given the standard models of exponential and quasi-hyperbolic discounting, Loewenstein found that willingness to pay for pleasant experiences was increasing as the number of periods until the pleasant experiences increased in the short term, peaked after some number of periods, and decreased in the long term. Highest willingness to pay (WTP) was for an experience a few days in the future and WTP to avoid an unpleasant experience increased the further in the future the unpleasant experience was expected. To collect the data, he posed the question of how much subjects were willing to pay for pleasant experiences and to avoid unpleasant experiences at different points in the future. Both classical and standard behavioral models of discounting predict the maximum WTP for a pleasant experience and to avoid an unpleasant experience to be when the experience is expected in the same time period as WTP is elicited. Lowenstein posited that there must be some factor, accrued in each time period, between the realization that an event will happen and its occurrence, which affects the utility of gains from the experience and therefore one's WTP.

These results and this theory is quantifiable by adding a factor of accrued utility from anticipating the event to the standard exponential discounting model. Given a positive event some time in the future, the agent anticipates the event and

[1] Loewenstein, George. Anticipation and Valuation of Delayed Consumption, The Economic Journal, 97 (September I987), 666-684

gains some measure of utility simply from fantasizing about that experience. Loewenstein used, as his positive anticipated experience, a kiss from the subjects' favorite movie stars. This experience was chosen because there is very little planning necessary for the experience to be pleasant but, given time between the offer and the event, the subject would have time to pleasantly fantasize about the kiss. The results suggest that subjects were willing to pay for opportunity to realistically fantasize about the event absent the utility of the actual experience.

This paper examines the theoretical basis of Loewenstein's model, develops moment equations for observable population measures, then examines the potential for using minimum distance estimation to estimate the parameters of the model from experimental data. This paper proceeds as follows: Section II describes the quantified model, Section III discusses challenges in structural estimation of the parameters and develops equations for observable population moments, Section IV describes how one might be able to independently estimate the values of the model's parameters, Section V tests the feasibility of using the observable population moments to identify the model's parameters, Section VI describes a possible method of performing these estimates, Section VII examines further questions for research including means of experimentation, comparison, and potential applications

II. The Model

In a multi-period model, an agent's willingness to pay at time zero for a single event in the future is in some way a function of the size of the future reward, the time to that reward, and some set of other factors:

$$U_0 = f(u, t, \ldots)$$

Where: u= utility from event

t= number of periods until event

In classical exponential discounting models, the utility from the reward in period t is given by that utility times a discount factor (δ) raised to the number of periods until the reward is recieved i.e. t. In this model:

$$U_0 = f(u, t, \delta) = \delta^t(u)$$

Where $0 \leq \delta \leq 1$

This implies that U_0 will decrease with more periods between period 0 and the period in which u is received. In the standard behavioral model of quasi-hyperbolic discounting, preferences for utility are present-biased, meaning all future utility is discounted by a single factor (β)

$$U_0 = f(u, t, \delta, \beta) = u_0 + \beta(\delta^t(u_t))$$

Where $0 \leq \delta \leq 1$, $0 \leq \beta \leq 1$

u_0 is a reward given immediately

u_t is a reward given after t periods

As this model is meant to account for people putting off unpleasant experiences until the future in favor of immediate utility even if that procrastination is ultimately detrimental to total utility, it cannot explain the future-biased preferences exhibited by Loewenstein's subjects.

Focusing on the anticipated positive experience, there is some factor that gives the subject some positive utility in the short term which, as t increases, is overtaken by future discounting. This interaction leads to a short term increasing, long term decreasing U_0 curve with respect to t. Loewenstein suggests a model in which U_0 from the exponential discounting model is considered the discounted present value of the experience in terms of utility. In every period, the subject anticipates some measure of utility based on the ability to savor the positive experience to come.

Theoretically, this means that there is utility accrued in every period between period 0 and t. In period (τ) 0 through t-1 , the subject gets the anticipatory factor (α) times $\delta^{t-\tau}u$ discounted to period 0 i.e. $\alpha\, \delta^{\tau}\delta^{t-\tau}u=\alpha\delta^{t}u$. In period t, the subject will receive the full value of u. As these periods are each in the future when U_0 is elicited, they are each discounted relative to the present period. Summed across all periods:

$$U_0 = \sum_{\tau=0}^{\tau=t-1} \alpha\delta^{t}\, u + \delta^{t}(u)$$

$$\text{simplified: } U_0=(1+ \alpha t)\, \delta^{t}(u)$$

III. Challenges in Structural Estimation

Within this model there are two main observables and two independent variables. The first observable is U_0. Assuming quasi-linear utility of money ($u(y)=y$), one can solicit a person's WTP in dollars for the reward (u) t periods in the future and it will equal U_0. Therefore:

$$WTP = U_0 = (1 + \alpha t)\ \delta^t (u)$$

The second main observable is the optimal t (t*), that level of t at which WTP is the highest. Setting ∂WTP/∂t equal to zero and solving for t:

$$t^* = -\ (1/\log(\delta)) - (1/\alpha)$$

Setting these two equations equal to each other does not eliminate the problem of α and δ being functions of one another. WTP=t* Solved for α:

$$\alpha = (\delta^{-t}(\sqrt{((\delta^t \log(\delta)+1)^2 - 4t\ \delta^t \log^2(\delta))} + \delta^t(-\log(\delta))-1))/(2t \log(\delta))$$

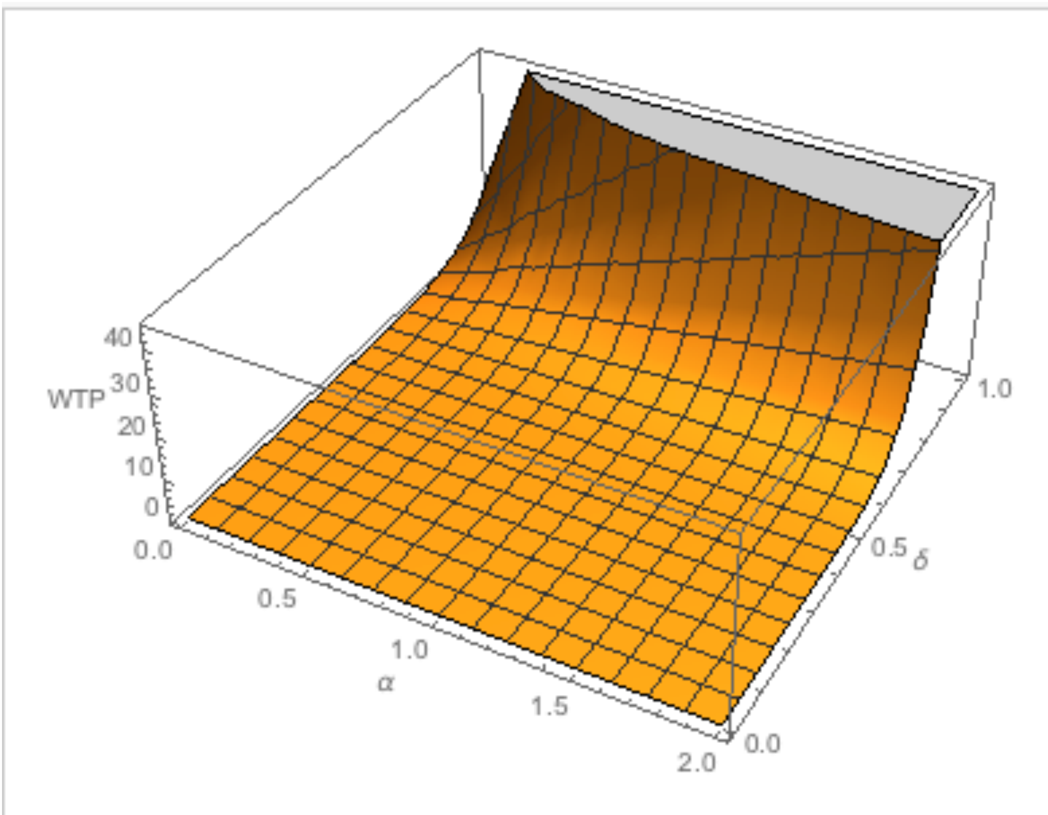
A third population moment appears to be required if one is to hope to attempt structural estimation, though there may still be unique values of α and δ which make this equation true. The exogenously variable parameters are t and u. The experimenter can vary the size of the reward and the number of periods until the reward is given. Given these adjustable and observable parameters the only other apparent observable population moment is how WTP changes with respect to time while u is kept constant i.e.:

$$\partial WTP/\partial t = u(\log(\delta)(1 + \alpha t)\delta^t + \delta^t\alpha)$$

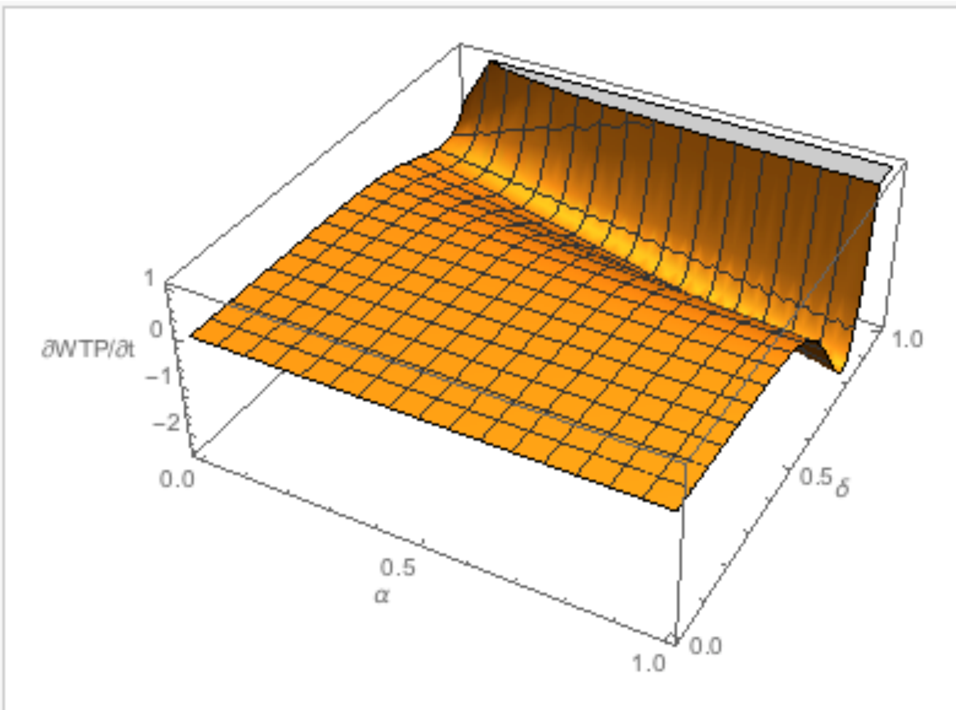
With these three observable moments one is able to at least test the feasibility of structural estimation.

IV. Independently identifying α and δ

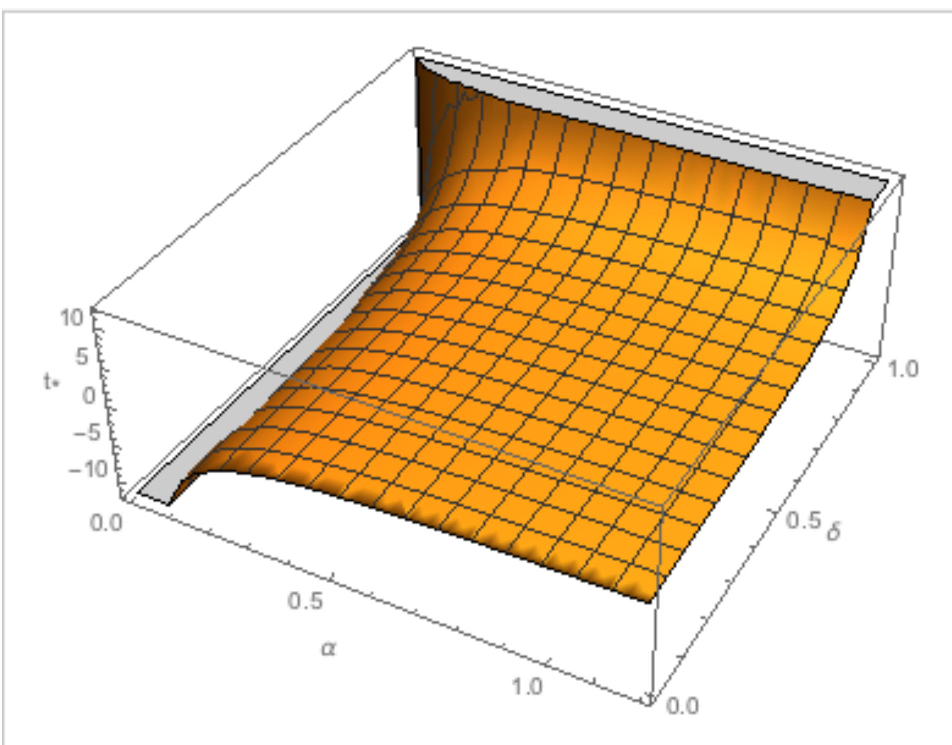
Given: the observable population moments and the goal of independently identifying unique values for α and δ given a data set. Therefore, One must examine what information observations for WTP, ∂WTP/∂t, and t* would impart based on the functions implied by the model and whether or not this information would allow us to possibly isolate unique values for α and δ. Each of the population's moments can be graphed as three-dimensional functions of α and δ given fixed values for u and t:*WTP(α, δ, u=15, t=10)*



∂WTP/∂t(α, δ, u=15, t=1)



t(α, δ)*



Assuming every individual has unique values of α and δ, each individual should be identified by a point on the α, δ plane and a height in each of these three graphs. Additionally, for any given value of WTP(α, δ, u, t), ∂WTP/∂t(α, δ, u, t), or t*(α, δ) (i.e. height on their respective graphs), there is a locus of all possible values of α and δ acquired by solving α as a function of δ and setting the observable parameter to its respective observed value for each of the population moments (e.g. α(WTP=30, δ, u=10, t=10) is a locus of (α,δ) pairs). Each of these loci can then be plotted concurrently on an α, δ plane. These three two-dimensional curves display all possible values of α and δ for given values of each of the observables. Any intersection points between all three of these curves specify a possible α, δ pair for that subject that would generate the specific set of observed values of WTP, ∂WTP/∂t, and t*. If the three curves intersect at more than one point, then it is not possible to uniquely identify one α, δ pair using these three observable variables. If there is only one intersection point, and this holds true over a reasonable range of α and δ, then experimental data can potentially be used to identify α and δ.

V. Testing feasibility of independent identification assuming a perfectly informed individual

Given: a fixed u and t and assuming an individual who can perfectly identify their WTP for any given t. This individual can, by extension, perfectly identify their ∂WTP/∂t and t* given constant u and varying values of t. With these three values, the three curves mentioned in the previous section can be graphed and their intersection behavior assessed.

Assuming an individual with relatively realistic α=0.5, δ=0.9, and fixing t = 8, u = 15:

$$WTP(\alpha, \delta, t, u) = WTP(0.5, 0.9, 8, 15) = 32.285$$

$$\partial WTP/\partial t(\alpha, \delta, t, u) = \partial WTP/\partial t(0.5, 0.9, 8, 15) = 0.185658$$

$$t^*(\alpha, \delta) = t^*(0.5,0.9) = 7.49122$$

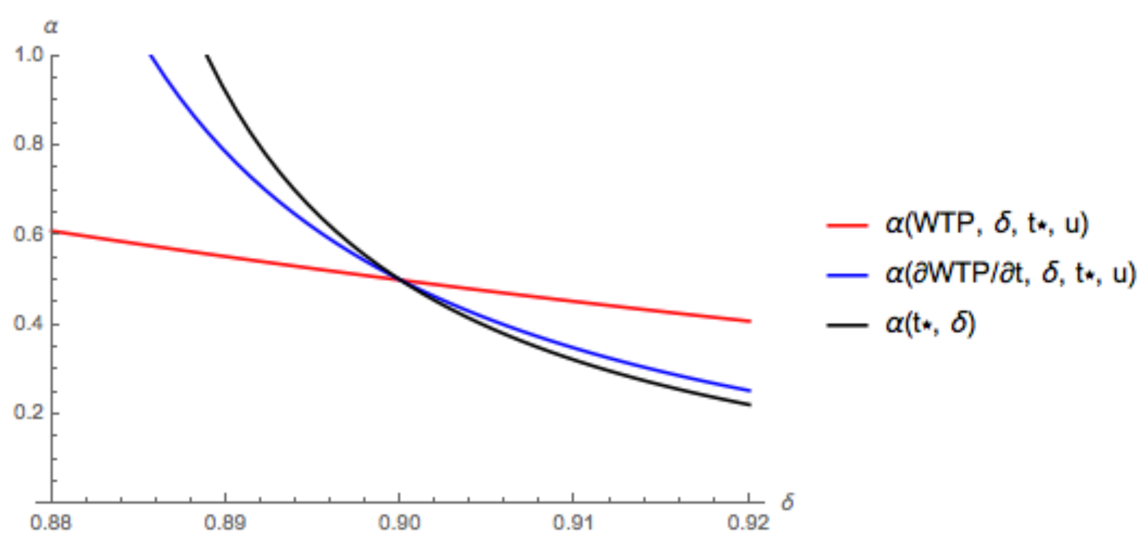
In order to check for identification of the α and δ parameters, one must solve for α as a function of δ and other variables:

$$\alpha(WTP, \delta, t, u) = (WTP - u\ (\delta^t))/(t\ u\ (\delta^t))$$

$$\alpha(\partial WTP/\partial t, \delta, t, u) = (\partial WTP/\partial t - u\ Log(\delta)\ (\delta^t))/((u\ Log(\delta)\ (\delta^t)\ t) + ((\delta^t)\ u))$$

$$\alpha(\delta, t^*) = -Log(\delta)/(t^*\ Log(\delta) + 1)$$

Plotting α(WTP, δ, t, u) = α(32.285, δ, 8, 15), α(∂WTP/∂t, δ, t, u) = α(0.185658, δ, 8, 15), and α(δ, t*)= α(δ, 7.49122):



Immediately apparent is the clear, unique intersection of all three curves at the values of α and δ specified at the beginning of this thought experiment. Additionally this process also appears to demonstrate that α and δ can be uniquely identified with just WTP and t* observations, but not with just WTP and ∂WTP/∂t observations as, with error added, they could be parallel or concurrent. After verifying these results for several other values of α and δ, this demonstrates that in ideal conditions, one can perfectly identify values of α and δ by eliciting a subjects WTP, ∂WTP/∂t, and t*.

VI. A proposed method of estimating α and δ values for a population accounting for human error

For each subject in this proposed study, multiple values of WTP would be elicited for different values of t and u. This within-subjects data would then be used to determine values of ∂WTP/∂t, t* and by extension α and δ. Relaxing the assumptions of a perfectly introspective subject, the results of soliciting WTP, ∂WTP/∂t, and t* will now have some error such that for subject i:

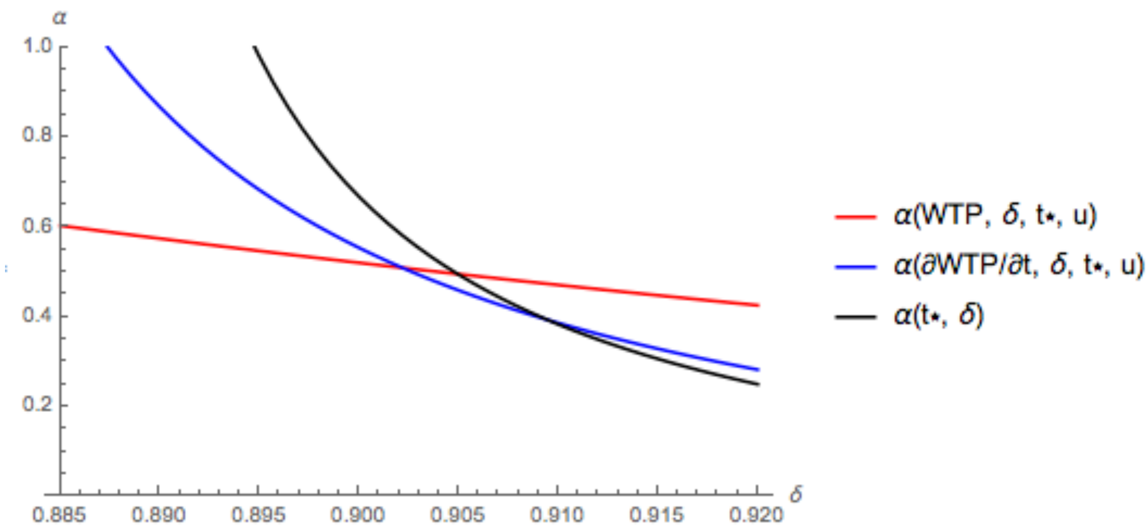
$$WTP_i = (1+ \alpha_i t)\, \delta_i t(u) + \varepsilon$$

$$\partial WTP/\partial t_i = u(\log(\delta_i)(1+ \alpha_i t)\, \delta_i^t + \delta^t \alpha_i) + \theta$$

$$t^*_i = - (1/\log(\delta_i)) - (1/\alpha_i) + \psi$$

where ε, θ & ψ are i.i.d error terms

This means for that for any given subject the curves resulting from the estimation method described above might not intersect at a single α, δ pair as in ideal conditions. In the following graph, compared to the α and δ graph in part V, WTP has been increased by 1, ∂WTP/∂t has been increased by 0.1, and t* has been decreased by 0.5, all of which are arbitrary values and add a hypothetical amount of error to the values used in the prior example.



The curves no longer intersect at a single point and none pass through the true value of $(\delta, \alpha) = (0.9, 0.5)$.

Eliciting values of the three observables from a population, one would require a method of dealing with this margin of error, namely minimum distance estimation. One would need to write a program to assess the most likely point at which the three curves would intersect in the absence of error. The specific design of this program is beyond the scope of this research. Given the outputs of $t^*_{i,}$ $\partial WTP/\partial t_i$, and WTP_i, for subjects $i=(1,\ldots,n)$, n being the number of subjects in the study, one would get $\alpha(WTP_i, \delta, t , u)$, $\alpha(\partial WTP/\partial t_i, \delta, t , u)$, and $\alpha(t^*_I, \delta)$, and perform a minimum distance estimation of the most likely (δ_i, α_i) for each value of i. The average of these estimates of α and δ would be an estimate of α and δ for the population as a whole.

VII. Further questions and research

Understanding the values of α and δ is important to accounting for decisions in which an agent forgoes utility in the short term in favor of apparently large rewards in the long term. Given that α and δ can theoretically be independently identified in Loewenstein's model, further research would focus on getting population estimates for α and δ, examining differences between different populations' α's and δ's, and exploring the finer mechanics of how this toy model relates to the real world.

The first step in conducting any applied research would be to write a program for minimum distance estimation. A thorough exploration of theoretical and applied uses of minimum distance estimates in scholarly literature would hopefully uncover a method to use statistical analysis software to take in values of all parameters for subjects in a study and return an estimated value of α and δ. Once this method was established, the next logical step is to design an experiment to rigorously attempt to observe WTP and t* for different values of u and t across a population. The development of this specific method and experiment are beyond the scope of this research.

Given a baseline estimate of α and δ, one could perform cross population analysis on the ability and willingness to delay gratification. As Loewenstein says, "…it requires little effort to think of examples of behaviour in which negative discounting is apparent. The pleasurable deferral of a vacation, the speeding up of a dental appointment, the prolonged storage of a bottle of expensive champagne are all instances of this phenomenon," (Loewenstein 1987)[2]. Understanding the strength of anticipation for different professions, education levels, nationalities, and in relation to other determinates of consumption behavior would give economists better insight into the effects of discounting and anticipation behavior have on one's quality of life.

To design further experiments, beyond the work of Loewenstein, we must answer questions of how people understand and interpret values of u and t. In the model as we currently understand it, the reward, u, is a scaling variable. WTP is linear with respect to u. However, this assumption may not hold in practice. The relationship should be tested with respect to reference dependence, diminishing marginal utility, and any other theories that would question how a person interprets the utility of a given reward. Likewise, the model does not give a simple explanation of what periods to use in measuring t. Loewenstein's original paper used a non-linear scale of hours, days, years, up to a decade. A better understanding of how people interpret delays in gratification is important to getting accurate estimates from the proposed procedure. Simply varying time without a better understanding of what people consider a period of time greatly decreases the validity of the estimates.

Given extensive study of a large sample, the process for estimation should produce a robust estimate of α and δ. However, further research is needed to understand discounting behavior and frames of reference. There is opportunity for improvement of current models of discounting behavior which is one of the most

---

[2] ibid

important areas of study for understanding consumption behavior given time delays between planning to consume, paying to consume, and actual consumption.