

4-1-2013

Estimating Discount Factors For Public and Private Goods and Testing Competing Discounting Hypotheses

Andrew G. Meyer

Marquette University, andrew.g.meyer@marquette.edu

Accepted version. *Journal of Risk and Uncertainty*, Vol. 46, No. 2 (April 2013): 133-173. The final publication is available at Springer via [DOI](#). © 2013 Springer. Used with permission. Andrew Meyer was affiliated with Ohio Wesleyan University at the time of publication. [Shareable Link](#). Provided by the Springer Nature [SharedIt](#) content-sharing initiative.

Estimating Discount Factors for Public and Private Goods and Testing Competing Discounting Hypotheses*

Andrew Meyer[†]

January 22, 2013

*This work was supported by the STC program of the National Science Foundation via the National Center for Earth-surface Dynamics under the agreement Number EAR-0120914.

[†]Assistant Professor, Department of Economics, Ohio Wesleyan University, Delaware OH, 43015, agmeyer@owu.edu

Abstract

The observation of declining discount rates in experimental settings has led many to promote hyperbolic discounting over standard exponential discounting as the preferred descriptive model of intertemporal choice. I develop a new framework, consistent with the random utility model, which directly models the intertemporal utility function and produces explicit maximum likelihood estimates of discounting parameters. I apply this estimation method to a stated-preference survey of river basin cleanup options and revealed-preference lottery payment choices. Formal statistical tests fail to find evidence in support of hyperbolic or quasi-hyperbolic discounting. Annual discount rates range from ten to fourteen percent across the data sets and empirical specifications.

JEL Classification: D90, Q25, Q53, H43

Keywords: discounting, hyperbolic, random utility, intertemporal choice

Acknowledgements: I thank Nicholas Flores for helpful comments concerning the survey design and data collection, for allowing me to borrow liberally from his MRB description, and for guidance throughout my dissertation process. I thank Randy Walsh for helpful comments at the inception of this research. I thank two anonymous reviewers for multiple insightful comments that greatly enhanced the quality of this paper. Finally, I thank participants at the 2008 AERE Sessions at the Summer Meeting of the AAEA and at the 10th Occasional Workshop on Environmental and Resource Economics at UC Santa Barbara.

Every day we make decisions involving tradeoffs of benefits and costs over time. Would I rather spread my workload evenly over the next few days and distribute the pain or procrastinate and have an extremely painful task several days from now? Should I exercise regularly while I'm younger so that I can enjoy the health benefits when I'm older? Will I invest time and money in my education today so that I can have a better lifestyle later? Am I willing to give up some consumption today so that I and others can enjoy a better environment in the future? These intertemporal choices penetrate nearly every aspect of our behavior. Such decisions require weighing benefits and costs that are realized with differing temporal patterns. Typically, individuals discount future outcomes, but how much, and in what way?

To answer this question, I develop a new approach that facilitates estimating discount factors for monetary choices or for other choices that can be presented in a stated-preference framework. I directly model the intertemporal utility function associated with an intertemporal outcome, which produces explicit estimates of discounting parameters within a random utility framework. This empirical strategy allows direct testing of competing hypotheses of how people discount future benefits and costs in a unified statistical framework.

Some previous evidence suggests that individuals discount the future hyperbolically or quasi-hyperbolically. That is, some studies find that inferred discount rates decline over time. These results mostly stem from experimental settings with private goods. It is also important to understand how individuals discount future public good improvements, which has not received as much research attention. Public support for policies with immediate costs and delayed rewards can differ greatly under hyperbolic and exponential discounting. Furthermore, natural real-world intertemporal choices have not been extensively studied. Thus, I present an estimation strategy that is flexible enough to handle public goods, experimental data, and real-world monetary choices.

After presenting the statistical model for estimating discount factors, I estimate the model on three data sets. One source is a stated-preference survey on river basin improvements, one is comprised of choices that individuals make when they win state lottery jackpots, and the final source is stated-preference monetary choices from a previously published paper. The former data set represents a public good choice, and the other two represent private good choices. These three sources provide a diverse representation of discounting at the individual level. All data sets produce similar discounting results. For the sample sizes in the empirical evidence used in this paper, there seems to be no statistical evidence supporting hyperbolic or quasi-hyperbolic

models over the standard exponential model. Constant annual discount rates range from ten to fourteen percent over the three data sets and various specifications.

I explore a variety of specifications to model heterogeneity in preferences for the river basin improvements. I examine the case where only discount factors are allowed to randomly vary across individuals, the case where only parameters on basin improvement and net income are allowed to vary across individuals, and the case where all model parameters are allowed to vary across individuals. I find substantial evidence of heterogeneity in time preferences for both exponential and hyperbolic discounting. However, I find no evidence that this heterogeneity is explained by observable personal characteristics. Likewise, I find significant heterogeneity in the marginal utility of basin improvement and net income when utilizing random coefficients but virtually no evidence that observable personal characteristics can explain the heterogeneity. In the preferred specification, where all parameters are treated as random, I estimate a mean exponential discount rate of 12.8 percent. Due to the nature of the variation in the revealed preference lottery data, I am not able to get as good of a test on hyperbolic discounting as one would like. Nevertheless, I estimate a mean exponential discount rate of 12.99 percent. Hence, mean discount rates from Americans in different parts of the country and in different contexts are remarkably similar and in line with interest rates that we see in capital markets, such as those on personal credit cards.

1 Existing Literature

1.1 Historical Development of the Discounted Utility Model

Samuelson (1937) first developed the discounted utility (DU) model in an attempt to provide a general model of intertemporal choice. Commonly referred to as the exponential discounting model, the DU model simplified all discounting into a single parameter, the discount rate. A consumer's preferences over consumption bundles, (c_0, c_1, \dots, c_T) are represented by an intertemporal utility function, $U(c_0, c_1, \dots, c_T)$. Furthermore, the DU model assumes that the intertemporal utility function is described by

$$U(c_0, c_1, \dots, c_T) = \sum_{t=0}^T \psi_t u(c_t), \quad (1)$$

where the discount factor for year t is $\psi_t = \left[\frac{1}{1+\rho}\right]^t$ and ρ is the discount rate.

Samuelson's DU model was accepted almost immediately because of its analytic

simplicity and elegance. Interestingly, Samuelson did not endorse the DU model as a normative model of intertemporal choice or as a valid descriptive model. The DU model was never empirically verified but still became the standard model for intertemporal utility (Frederick et al., 2002).

1.2 Departures from the Discounted Utility Model

In the past several decades, research has uncovered many situations in which the DU model does not fit behavior.¹ One major departure from the DU model is that inferred discount rates often decline over time in experimental settings. This phenomenon is commonly termed hyperbolic discounting. This discounting gets its name because a hyperbolic functional form fits the data better than the traditional exponential functional form. Several functional forms have been suggested for hyperbolic discounting. The most popular of these takes the form of

$$\psi_t = (1 + \alpha t)^{-\beta/\alpha}, \text{ where } \alpha, \beta > 0 \quad (2)$$

(Loewenstein and Prelec, 1992). As α goes to 0, this hyperbolic discounting function becomes the exponential discounting function. To facilitate estimation, researchers typically simplify equation 2 to have only one parameter. Constraining α to be equal to one produces the model suggested by Harvey (1986). Harvey's single-parameter hyperbolic structure is given by

$$\psi_t^{Harvey} = (1 + t)^{-\mu}. \quad (3)$$

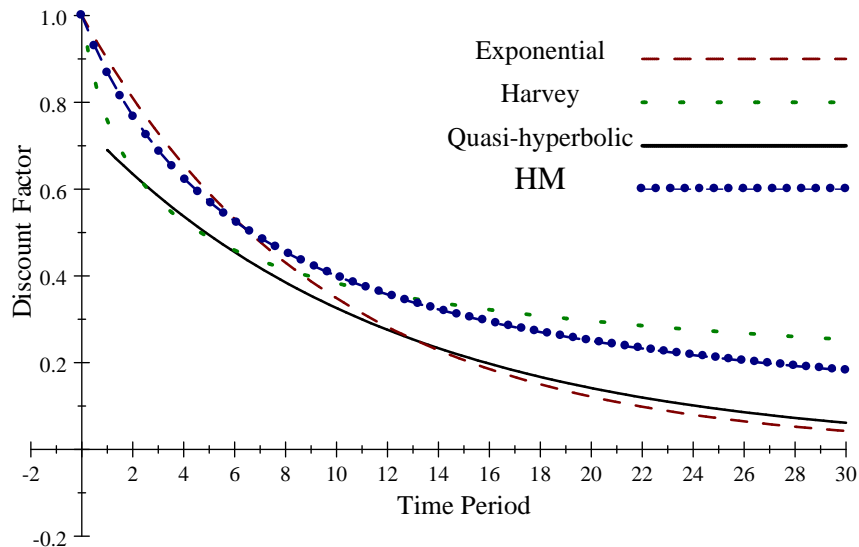
Alternatively, constraining the ratio of β/α to be equal to one results in the single-parameter model suggested by Herrnstein (1981) and Mazur (1987) (HM);

$$\psi_t^{HM} = (1 + \omega t)^{-1}. \quad (4)$$

In recent years, an alternative model of discounting that has received much attention is the quasi-hyperbolic (β, δ) discounting model. This model, developed by Laibson (1997), is also motivated by the observation of declining discount rates. The functional form was first introduced by Phelps and Pollak (1968) in the context of intergenerational altruism. The form of the quasi-hyperbolic discounting function is very simple and its contrast with the standard exponential discounting model is

¹See, for example, Cairns and van der Pol (2000, 1997)

Figure 1: Comparison of Discount Factors: Exponential (δ^t) with $\delta = .9$, Harvey Hyperbolic $((1 + t)^{-\mu})$ with $\mu = .4$, Quasi-hyperbolic $(1, \beta\delta^t)$ with $\beta = .75$, $\delta = .92$, and HM Hyperbolic $((1 + \omega t)^{-1})$ with $\omega = .15$.



readily apparent. The functional form is given by

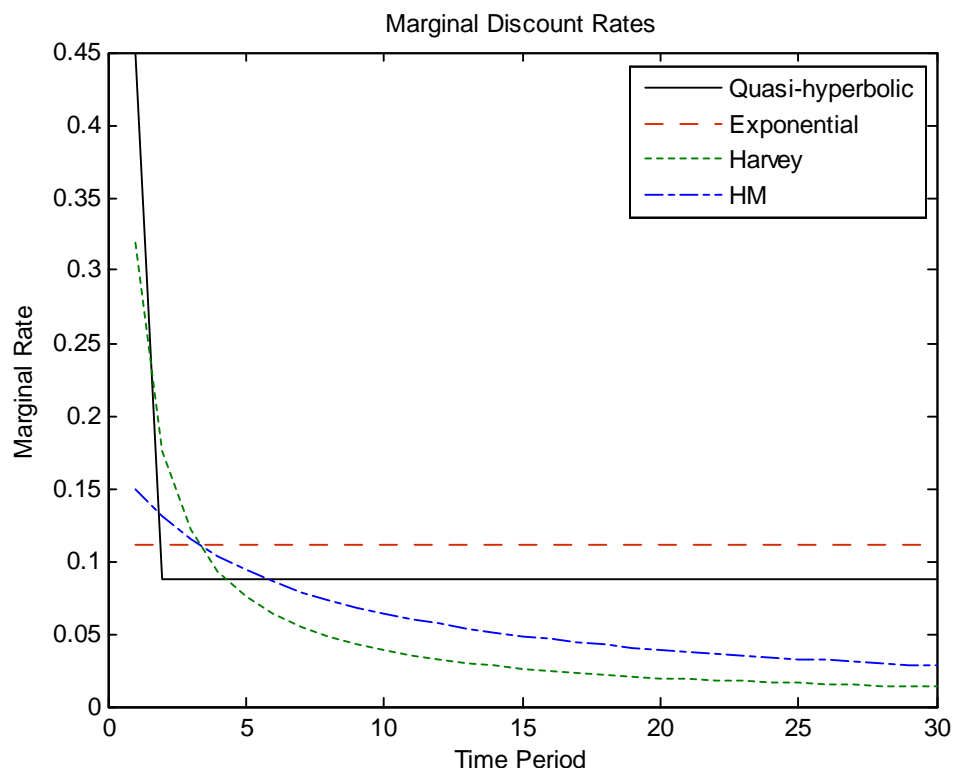
$$\psi_t = \begin{cases} 1 & \text{if } t = 0 \text{ and} \\ \beta\delta^t & \text{if } t > 0 \end{cases}, \text{ where } 0 < \beta < 1, \text{ and } \delta < 1. \quad (5)$$

Thus, the only difference between discount factors in the quasi-hyperbolic formulation and the exponential formulation is that all future time periods are discounted by the additional β factor in the quasi-hyperbolic model. Especially large importance is placed on immediate utility as compared to deferred utility. The (β, δ) discounting model is much easier to analyze than the true hyperbolic model, yet it retains many of the qualitative aspects of the more complicated model.

As shown in Figure 1², both hyperbolic and the quasi-hyperbolic discounting functions weight the near future less heavily than exponential discounting. However, for time periods far in the future, exponential discounters place less weight on the deferred utility than hyperbolic or quasi-hyperbolic discounters. Figure 2 shows the corresponding marginal discount rates for all four discounting functions. The point plotted for time period t is the marginal discount rate between time period $t - 1$ and

²The parameter values used for the exponential, Harvey hyperbolic, and HM hyperbolic models in these figures are consistent with those that I find from the data sets employed in this paper. The β chosen for the quasi-hyperbolic model is in the range of values discussed in the literature.

Figure 2: Comparison of Marginal Discount Rates: Exponential (δ^t) with $\delta = .9$, Harvey Hyperbolic ($(1 + t)^{-\mu}$) with $\mu = .4$, Quasi-hyperbolic ($1, \beta\delta^t$) with $\beta = .75$, $\delta = .92$, and HM Hyperbolic ($(1 + \omega t)^{-1}$) with $\omega = .15$.



time period t .

Exponential discounters will always display time consistency because their marginal discount rate is constant over all time periods. Quasi-hyperbolic discounters have a large marginal discount rate between time period 0 (now) and time period 1 and a constant marginal discount rate thereafter. Thus, quasi-hyperbolic discounters are dynamically consistent for any choice that does not involve the present. Regardless, most interesting economic choices involve the present. Finally, hyperbolic discounters always have declining discount rates. Therefore, a hyperbolic discounter is subject to dynamic inconsistency for any time period. However, hyperbolic marginal discount rates change less for time periods farther in the future. That is, they will be less likely to be dynamically inconsistent for tradeoffs that occur far in the future than for tradeoffs that occur near to the present. Hyperbolic discounting makes individuals appear to be impatient for immediate tradeoffs, but sufficiently patient for tradeoffs occurring far enough in the future.

Table 1: Present Discounted Values of 100 Dollars Now vs. 120 Dollars 1 Year from Now for Exponential with $\delta = .9$, Harvey Hyperbolic with $\mu = .4$, HM Hyperbolic with $\omega = .15$, and Quasi-hyperbolic with $\beta = .75, \delta = .92$

Model	Discounted Value of \$100 Now	Discounted Value of \$120 1 Year from Now	Choice
Exponential	\$100.00	\$108.00	\$120 in 1 Year
Harvey Hyperbolic	\$100.00	\$90.95	\$100 Now
HM Hyperbolic	\$100.00	\$92.30	\$100 Now
Quasi-hyperbolic	\$100.00	\$82.80	\$100 Now

A simple example highlights the time inconsistency inherent in hyperbolic and quasi-hyperbolic discounting. Assuming parameter values that are in the range of those found in the literature, I analyze the choice between \$100 now and \$120 a year from now and compare this with the choice between \$100 five years from now and \$120 six years from now. The interval length between the options for each choice is one year so a dynamically consistent discounter should choose either the more proximate reward in both scenarios or the more distant reward in both scenarios. Table 1 presents the discounted values of \$100 now and \$120 one year from now for all four discounting models. Table 2 shows the discounted values of \$100 five years from now and \$120 six years from now. The exponential discounter remains consistent in their choice to take the deferred payoff. However, the hyperbolic and quasi-hyperbolic discounters choose the early reward for the immediate tradeoff and choose the more distant payoff for the future tradeoff.

It is desirable to be dynamically consistent from a normative standpoint. With free access to capital markets, individuals should equate the marginal rate of substitution between two time periods to one plus the interest rate. A third party planner could improve a hyperbolic or quasi-hyperbolic discounter's intertemporal utility by rearranging consumption between time periods. In contrast, the welfare of an exponential discounter that is trading off consumption between time periods at one plus the interest rate cannot be improved upon by a third party planner.³

³Note that in this paper I abstract from the notion of discounting the utility of others. While that is a fundamental question in itself, I only examine the behavior of an individual concerned with their own utility.

Table 2: Present Discounted Values of 100 Dollars 5 Years from Now vs. 120 Dollars 6 Years from Now for Exponential with $\delta = .9$, Harvey Hyperbolic with $\mu = .4$, HM Hyperbolic with $\omega = .15$, and Quasi-hyperbolic with $\beta = .75, \delta = .92$

Model	Discounted Value of \$100 5 Years from Now	Discounted Value of \$120 6 Years from Now	Choice
Exponential	\$59.05	\$63.77	\$120 6 Years from Now
Harvey Hyperbolic	\$48.84	\$55.10	\$120 6 Years from Now
HM Hyperbolic	\$57.14	\$63.16	\$120 6 Years from Now
Quasi-hyperbolic	\$49.43	\$54.58	\$120 6 Years from Now

1.3 Discounting Studies

I concentrate on several of the more recent contributions and note that a more extensive literature review on discounting is provided by Frederick et al. (2002). While three recent working papers use utility-theoretic models incorporating goods other than money, the majority of previous studies examine monetary trade-offs over time. I point out that much of the evidence supporting hyperbolic discounting can be recast in terms of confounding factors.

1.3.1 Estimation Methods

The most common method for gathering data on discounting is to elicit experimental responses to hypothetical or real monetary rewards. Two approaches are most widely utilized. Respondents are either asked to choose between two different sized rewards realized at different times in the future or to state the payoff today that would make them indifferent to a larger payoff in the future (or the payoff in the future that would make them indifferent to a smaller payoff today). Harrison et al. (2002) represents the former approach and Coller and Williams (1999) falls into the latter category. Harrison et al. find an overall individual discount rate in Denmark of 28.1 percent using money data and they observe significant heterogeneity in the data. One notable exception to the experimental emphasis is the revealed-preference study by Warner and Pleeter (2001). They examine the decisions of military personnel when faced with a downsizing. Personnel choices of whether to take a lump sum payment or an

annuity reveal information about their intertemporal preferences.

Several recent experimental papers have relaxed the assumption of linear utility. If individuals truly possess concave utility functions, the erroneous assumption of linear utility would lead to an overestimation of discount rates. Andersen et al. (2008) elicit risk preferences and discount rates of Danish individuals using experiments. They find an average discount rate of approximately 10 percent when allowing for a concave utility function as compared to approximately 25 percent when assuming linear utility. They are unable to test for quasi-hyperbolic discounting since they employ a front-end delay experimental design. Andreoni and Sprenger (2010a,b) design a set of experiments to generate simultaneous estimates of discounting and utility curvature parameters. They find average annual discount rates that range between 20 and 35 percent across specifications and studies, which are considerably lower than the rates from many previous experimental studies. Andreoni and Sprenger (2010a) do find evidence of concave utility but do not find evidence of hyperbolic discounting.

Various studies have examined discounting for health outcomes. This branch of the discounting literature appears to begin with Horowitz and Carson (1990). In these studies, respondents state how many lives saved in the future is equivalent to saving a certain number of lives today, or respondents choose between varying durations of illness experienced at different times in the future. van der Pol and Cairns (2001) use this second method and provide the first example of discrete choice experiments to address discounting for health outcomes.⁴

Two recent related papers use an empirical model that is similar to the model I propose. Bosworth et al. (2006) jointly estimate individual-specific discount rates and the demand for preventative public health policies. They utilize a conjoint survey design in which respondents make choices between policies that reduce the number of illnesses and deaths in their community. At the same time, they have individuals choose between a hypothetical lottery that provides a series of payments over several years and a lottery that provides a lump sum payment. This method is based on the identification strategy developed by Cameron and Gerdes (2003). The authors of both papers argue that the two distinct data sources allow improved joint estimation of the utility parameters and discount rates and that it is often not possible to identify discounting parameters out of a public goods choice.

I show that discounting parameters for public goods are identified in a stated-preference framework if the policy options are designed correctly. Bosworth et al.'s

⁴Multiple other health discounting studies exist. For example, see two papers by Johannesson and Johannesson (1997a,b).

(2006) empirical model uses a utility-theoretic structure for preferences and assumes i.i.d. extreme value errors for the policy choices. Bosworth et al. (2006) do not allow for discount rates to take forms other than the standard exponential and single parameter hyperbolic models. I extend the model to test for quasi-hyperbolic preferences.

Viscusi et al. (2008) design a study to infer discount rates for a publicly provided good. They utilize a stated preference survey concerning improvements in local water quality to identify individual rates of time preference. Using a random utility model, they find that the data fit better with the quasi-hyperbolic discounting model than with the exponential discounting model. They employ a two-stage empirical approach to generate parameter estimates but do not conduct hypothesis tests on the quasi-hyperbolic discounting parameters. I build upon the survey design from Viscusi et al. The random utility theoretic framework herein produces explicit standard errors for all discounting parameters so I am able to formally test quasi-hyperbolic discounting.

1.3.2 Confounding Factors in Discounting Studies

Although evidence in the literature suggests that individuals have hyperbolic discounting preferences, I propose that some of this evidence can be explained by confounding factors. As emphasized in the review article by Frederick et al. (2002), it is important to differentiate between pure rates of time preference and other reasons that cause individuals to care less about future outcomes. Pure time preference refers to "the preference for immediate utility over delayed utility" (Frederick et al., 2002). Confounding factors that cause individuals to care less about the future but should be considered separately from pure time preference include uncertainty about a future outcome, perceived future transaction costs, and the phenomenon of sub-additive discounting. In this section, I show how experimental designs that do not address these three confounding factors could make an exponential discounter appear as though they are a hyperbolic discounter.

Imagine an experimental setting in which an individual is choosing between a smaller immediate reward and a larger delayed reward. Uncertainty in the receipt of the future reward can be problematic for estimating discount rates in this scenario. Suppose that this individual is truly an exponential discounter but perceives only a 70 percent chance that the researcher will actually deliver the delayed reward at any time in the future and a 100 percent chance that the immediate reward will be delivered. Then, the results from the experiment would look exactly like the individual is a quasi-hyperbolic discounter with a β value of 0.7. Or, suppose that this individual is truly an exponential discounter with a constant discount factor of $\delta < 1$ but believes

with probability $p_0 = 1$ that they will receive an immediate reward, with probability $p_1 < 1$ that they will receive a delayed reward at $t = 1$, and with probability p_t , such that $p_{t+1} < p_t$ and $p_{t+1} - p_t > p_{t+2} - p_{t+1}$, that they will receive a delayed reward at time t . That is, the perceived probability of receiving a future reward declines at a decreasing rate. Then, observed discount factors including the confounding effect of uncertainty are given by $\{1, p_1\delta, p_2\delta^2, p_3\delta^3, p_4\delta^4, \dots\}$. Marginal observed discount rates are given by $\{1/p_1\delta - 1, p_1/\delta p_2 - 1, p_2/\delta p_3 - 1, p_3/\delta p_4 - 1, \dots\}$. These resulting observed discount rates are consistent with a hyperbolic functional form. To further illustrate with a numerical example, assume $\delta = .9, p_1 = .8, p_2 = .7, p_3 = .65, p_4 = .61$. This gives marginal discount rates of $\{38.9\%, 26.9\%, 19.7\%, 18.4\%\}$. However, when abstracting from the effects of uncertainty, true marginal rates of time preference are given by $\{1/\delta - 1, 1/\delta - 1, 1/\delta - 1, \dots\}$. Thus, it is important to minimize the effects of future uncertainty in a discounting study.

Next, suppose that within an experimental setting an individual perceives a transaction cost of c_t in order to collect a payment at time t in the future. Also suppose that this individual is an exponential discounter with a discount factor of δ^t . Then, in order to be indifferent between an immediate payment of $\$x_0$ and a delayed payment of $\$x_t$, it must be that $x_0 = \delta^t(x_t + c_t)$. If $c_{t+1} = c_t$ for all $t > 0$, observed marginal discount rates look like quasi-hyperbolic discount rates. If $c_{t+1} > c_t$ for all $t > 0$ observed marginal discount rates can look like hyperbolic discount rates.

To make ideas more concrete, consider the following example. Consider this individual indicating their indifference point between an immediate reward of \$100 and a larger delayed reward. Let the perceived future transaction costs $c = \{c_0, c_1, c_2, c_3, c_4\} = \{0, 10, 20, 30, 40\}$. Assume $\delta = .9$. Denote the marginal discount rate between time periods t and $t + 1$ as $r_{t,t+1}$. Let superscripts of *true* and *obs* denote the true (exponential) and observed values. Then $r_{t,t+1}^{true} = 11.1\%$ for all t . Denote the delayed reward at time period t as x_t . Next, ignoring the transaction cost, c_0 , it holds that $100 = .9 * x_1$. Solving, $x_1 = 111.11$ would make this individual indifferent in absence of transaction costs. Taking into account the effect of the transaction cost, $100 = \delta_1^{obs}(111.11 + 10)$. Solving, $\delta_1^{obs} = .8257$. Then, $r_{0,1}^{obs} = 1/\delta^{obs} - 1 = 21\%$. Again ignoring the transaction cost, c_1 , $100 = .81 * x_2$. Solving, $x_2 = 123.46$ would make this individual indifferent in absence of transaction costs. Taking into account the effect of the transaction cost, $100 = \delta_2^{obs^2}(123.46 + 20)$. Solving, $\delta_2^{obs^2} = .6971$. This implies $r_{1,2}^{obs} = \delta_2^{obs^2}/\delta_1^{obs} - 1 = 18.45\%$. Continuing with this pattern, I find $r_{2,3}^{obs} = 16.53\%$ and $r_{3,4}^{obs} = 15.10\%$. I observe declining marginal discount rates even though the true marginal discount rates are constant. The larger the transaction

cost relative to the size of the reward, the more pronounced this effect will be.

One other explanation for the observation of declining discount rates is the idea of subadditive discounting. That is, "discounting over a delay is greater when the delay is divided into subintervals than when it is left undivided" (Read, 2001). Many laboratory experiments look over days or months and confound the length of the delay with the length of the interval between choices. For example, a researcher will compare the discount rate inferred from a choice involving zero to six month delays to that from a choice involving zero to twelve month delays. When annualized, the discount rate will look larger from the choice involving zero to six month delays. Therefore, the discount rate looks like it declines over time. However, discount rates are declining because the length of the interval is increasing. Many researchers anchor all choices to a particular time and do not design choices to keep interval length independent from the length of delay. Typically, a shorter interval length necessarily means a shorter delay until the delayed outcome. Read (2001) uses experiments to verify the presence of subadditive discounting but finds no evidence of hyperbolic discounting.

1.3.3 Indirect Tests of Hyperbolic Discounting

Several studies have attempted to determine whether exponential or hyperbolic discounting is preferred. In this section, I summarize the studies that have indirectly tested for hyperbolic discounting. Also, I analyze how each study addresses uncertainty in a delayed reward, perceived future transaction costs, and subadditive discounting.

Kirby and Marakovic (1995) fit hyperbolic and exponential discount functions for each subject. They utilize nonlinear regression techniques on the continuous time equations for exponential and hyperbolic discounting. They find that, while both do a good job explaining subjects' responses, the hyperbolic model fits better in terms of R^2 for almost all of the subjects. Uncertainty in the payment of the delayed reward is present since delayed rewards were not to be delivered until the evening on the day that it came due. Transaction costs are especially relevant because the rewards are small (\$14.75-\$28.50 for delayed rewards). This study confounds length of delay until the delayed reward is received with the length of the interval between options since all choices are anchored to the present.

Slonim et al. (2007) conduct an experimental study in which they examine whether or not possession of the delayed reward affects subjects' discounting patterns. They find that discount rates decline over time in all cases. Possession of rewards sup-

ports quasi-hyperbolic discounting and no possession supports hyperbolic discounting. They do not find any evidence of exponential discounting. This study attempts to control for transaction costs in the best way possible by using possession of the reward as a control variable. Also, this study uses a common interval length of two months for all choices so interval length is not confounded with the length of delay until the receipt of the future reward. Uncertainty in future rewards is nullified in the cases where individuals choose between two future rewards if the perceived probability of receipt of the reward is constant over time. However, uncertainty in future rewards is still an issue if the probability of receipt of the reward declines with longer time delays. Also, for the choices anchored to the present, uncertainty in future rewards remains a confounding factor.

Cairns and van der Pol (2000) compare three hyperbolic models with the exponential model. For each individual and discounting model, they first estimate optimal parameter values using non-linear least squares. Second, they regress these parameter values on the period in years for which the benefit is delayed, claiming that delay should be insignificant for a correctly specified discounting model. Delay is insignificant only in the Loewenstein and Prelec model (2 parameter hyperbolic). They also note that the first stage regressions have the highest R^2 for the hyperbolic models. Since all choices are anchored to one year in the future, uncertainty in rewards is controlled for if the perceived probability of receiving the reward is constant over all time periods but not if the perceived probability of receipt declines with time. Transaction costs are minimized in the case of social financial benefits since the receipt of the reward does not require any work on part of the survey respondent. For private financial benefits, transaction costs likely get larger as the delayed reward moves farther into the future. If transaction costs are constant over all future time periods, they will have no influence in this study since all choices are anchored to one year from the present. However, because of this common anchor, the length of delay and length of interval are confounded. Subadditive discounting may explain any evidence for hyperbolic discounting.

Keller and Strazzera (2002) examine the predictive accuracy of the exponential and hyperbolic models in a simulated data set. Using Thaler's (1981) 1981 experimental data to calculate implicit monthly discount rates, the authors generate a simulated data set of predicted matching values, m_t , that would make a respondent indifferent to an immediate reward, m_0 . Comparing these predicted values with the actual matching values from Thaler's data set, they find that the hyperbolic model does a better job than the exponential model. Thus, indirect tests suggest that hyperbolic

discounting is preferred to exponential discounting. All choices are anchored to the present. This leaves open the possibility of confounding effects from uncertainty in future rewards, future transaction costs, and subadditive discounting.

As previously mentioned, Andreoni and Sprenger (2010a) do not find any evidence of hyperbolic discounting. They are careful to equate transaction costs between more proximate and delayed payments so that any transaction cost effect would be negated. They also take considerable action to reduce uncertainty in the receipt of a delayed reward. Interestingly, when Andreoni and Sprenger (2010b) reintroduce uncertainty in the reward payoff, they observe that experiment participants exhibit less utility concavity for certain consumption compared to uncertain consumption and disproportionately prefer certain consumption. The result is that present consumption would be disproportionately preferred to delayed consumption and individuals could appear to be quasi-hyperbolic discounters due to the differences in their evaluation of certain and uncertain consumption.

I build on these previous discounting studies by closely considering potential confounding factors. I select data sets that minimize uncertainty in (hypothetical and real) delayed rewards, decision-maker transaction costs, and subadditive discounting. Through jointly addressing these experimental concerns and developing a new empirical model that directly estimates the discounting parameters, I attempt to isolate pure rates of time preference for various models and test to find the statistically preferred model.

2 Empirical Strategy

2.1 Derivation of the General Model

Here I present the random utility model to analyze discrete choice data. This model analyzes choices over goods that are intertemporal in nature. In general, let the instantaneous utility for an individual i for choice j in year t be given by

$$u_{ijt} = v_{ijt} + \eta_{ijt}. \tag{6}$$

Here, v_{ijt} contains a vector of observed variables relating to choice j and a vector of fixed coefficients and η_{ijt} is the instantaneous error draw. It is important to note at this point that instantaneous utility is not at all observable. That is, the researcher only observes behavior at the choice level.

I make the usual assumption that intertemporal utility is additively separable over

time periods. Denote the final time period for which any of the alternatives in the choice set, s , have an impact on decision-maker utility as T_s . Then the utility for individual i that is associated with choice j defined through time period T_s is given by

$$U_{ij}(u_{ijt}, \psi_t) = \psi_0 u_{ij0} + \psi_1 u_{ij1} + \dots + \psi_{T_s} u_{ijT_s}, \quad (7)$$

where ψ_t is the discount factor for year t .⁵ Substituting equation 6 into equation 7 and rewriting in summation notation produces

$$U_{ij} = \sum_{t=0}^{T_s} \psi_t v_{ijt} + \epsilon_{ij}, \quad (8)$$

where $\epsilon_{ij} = \sum_{t=0}^{T_s} \psi_t \eta_{ijt}$ is the error for individual i associated with choice j . Thus, the intertemporal utility from a choice is essentially the weighted sum of all instantaneous utilities. Discount factors determine the weight placed on each time period. The specification of v_{ijt} will depend on the type of intertemporal choice that is being analyzed.

2.2 Structure of the Error Terms

Since a rational individual makes utility evaluations at the instantaneous level and discounts them back to the present, it is appropriate to assume the distribution of the instantaneous errors (η_{ijt}). However, the researcher observes choices at the alternative level so it is necessary to use the model structure to determine the alternative level error structure. This approach contrasts the Bosworth et al. assumption that alternative errors are i.i.d. extreme value. I show in the appendix that even i.i.d. error assumptions at the instantaneous level imply heteroskedastic errors at the alternative level.⁶

It is appropriate to build in correlation across utilities for a given individual in panel applications where decision makers face more than one choice occasion. One approach would be to assume that the η_{ijt} are correlated across time periods, alternatives, and choice occasions. I do not take this approach as the model quickly becomes intractable. Another approach when analyzing panel data is to specify random para-

⁵I first derive the basic model without allowing for individuals to have differing discount factors. Later, I show how the likelihood function generalizes to accommodate a random coefficients specification on the discount factors.

⁶In the applications contained herein, results do not substantially change if I ignore the heteroskedasticity. As further explained in the subsequent subsection, this could be a result of the properties of the survey design utilized.

meters. One option to model correlation across choices for a given decision maker is to assume that the discounting parameters are randomly distributed.⁷ Alternatively, one could assume v_{ijt} contains a vector of random coefficients to model an individual's correlated utilities.⁸ To ease exposition, I initially develop the model and likelihood function here assuming that v_{ijt} does not contain any random coefficients.

2.3 The Log-likelihood Equation

For each choice set, s , an individual chooses the alternative that provides the most utility. Therefore, the probability that individual i chooses alternative j from s is

$$P_{ij} = \Pr(U_{ij} > U_{ik} \forall k \neq j \in s). \quad (9)$$

The task is to determine the form of P_{ij} . Begin by substituting equation 8 into equation 9 to get

$$P_{ij} = \Pr\left(\sum_{t=0}^{T_s} \psi_t v_{ijt} + \epsilon_{ij} > \sum_{t=0}^{T_s} \psi_t v_{ikt} + \epsilon_{ik}\right) \quad (10)$$

$$= \Pr\left(\epsilon_{ik} - \epsilon_{ij} < \sum_{t=0}^{T_s} \psi_t v_{ijt} - \sum_{t=0}^{T_s} \psi_t v_{ikt}\right) \quad (11)$$

Next, denote the alternative error-difference term as $\tilde{\epsilon}_{ikj} = \epsilon_{ik} - \epsilon_{ij}$. Recall that $\epsilon_{ij} = \sum_{t=0}^{T_s} \psi_t \eta_{ijt}$ and note (shown in Appendix) that the variance of the alternative error-difference term, $\tilde{\epsilon}_{ikj}$, is

$$V(\tilde{\epsilon}_{ikj}) = 2 * \sum_{t=0}^{T_s} \psi_t^2. \quad (12)$$

Therefore, for any choice set, the variance of the alternative error-difference term will be larger when one or more alternatives have longer durations. Ignoring this in the likelihood function can lead to inconsistent parameter estimates and biased standard error estimates. Returning to equation 10 and using the definition of the c.d.f. (F)

⁷Again, in this initial model derivation, I assume discounting parameters are non-stochastic since two of the three applications herein do not involve panel data. I later show how the likelihood function generalizes to allow for random discounting parameters.

⁸I derive the simulated log likelihood equation for the case of random coefficients within v_{ijt} in the appendix.

of a normal random variable, I have

$$P_{ij} = F \left(\frac{\sum_{t=0}^{T_s} \psi_t v_{ijt} - \sum_{t=0}^{T_s} \psi_t v_{ikt}}{\sqrt{2 * \sum_{t=0}^{T_s} \psi_t^2}} \right). \quad (13)$$

The log-likelihood equation is then

$$LL = \sum_i \sum_j y_{ij} \ln P_{ij}, \quad (14)$$

where $y_{ij} = 1$ if i chose alternative j and zero otherwise.

Note that observations from choice sets with alternatives having longer durations are weighted less heavily than observations from choice sets with alternatives having shorter durations, which corrects for heteroskedasticity in the data. Depending on the correlation between the duration of choice alternatives and the other observed variables, this heteroskedasticity correction may or may not prevent significant bias in parameter estimates. If, for example, the number of time-periods across which future conditions will differ between alternatives is determined independently from other attributes of the alternatives, then the heteroskedasticity is not likely to create much bias. Utility differences from choices with longer time horizons will be divided by a larger denominator than choices from shorter time horizons, but this doesn't affect parameter estimates substantially because the survey was designed to have low correlation between survey attributes. If, however, data come from contexts where we would expect correlation between choice alternative characteristics, then it could be important to correct for this heteroskedasticity. For example, there may be correlation between policy characteristics in revealed preference data from real world referenda. Voters often choose between low levels of public goods provided at low cost and high levels of public goods provided at high cost. Suppose further that a low level of the public good would have a shorter time horizon and a high level of the public good would have a longer time horizon. Then, cost and quantity of the public good are correlated with policy duration and a researcher may find that ignoring the heteroskedasticity leads to biased estimates.⁹ This heteroskedasticity argument also relates to prior literature on discrete choice experiments. For example, DeShazo and Fermo (2002) show that it is important to control for heteroskedasticity that arises from characteristics of the choice sets.

⁹This is analogous to having a non-representative sample. If such selected sampling is uncorrelated with the dependent variable in a regression, the slope estimates will not differ substantially from those for a representative sample.

2.4 Application to a Public Good Choice

This model is particularly well suited to analyze discrete choice experiment data. Discrete choice experiments allow the researcher to specify several attribute dimensions of the intertemporal choices. Thus, the researcher can specify when the benefits and costs of an intertemporal choice are to be realized so that it is possible to identify the discount factors from respondents' choices. Public goods policies are a good example of choices that receive benefits and costs at differing points in time. For example, it is common to pay taxes today for a public good that will deliver benefits years into the future. In this section I develop the model for discrete choice experiments in the context of public goods choices.¹⁰

At any time the utility an individual receives from a simple public good policy depends on the level of benefit provided and the cost incurred. Specify the deterministic portion of instantaneous utility as

$$v_{ijt} = \alpha q_{ijt} + \gamma(I_{it} - c_{ijt}), \quad (15)$$

where q_{ijt} is the level of benefits from the public good, I_{it} is income, and c_{ijt} is the cost of the public good for individual i for policy j in year t .¹¹ In this specification, α is the marginal utility of the public good benefit and γ is the marginal utility of money. Let T_j denote the last year for which there are non-zero costs or benefits for policy j . Substituting equation 15 into equation 8 results in

$$U_{ij} = \sum_{t=0}^{T_j} \psi_t [\alpha q_{ijt} + \gamma(I_{it} - c_{ijt})] + \epsilon_{ij}. \quad (16)$$

This equation is the foundation of my econometric model.

Because only differences in utility matter in the random utility model (RUM), any personal characteristic on its own such as I_{it} drops out of the analysis. Personal characteristics can enter through interactions with policy characteristics. Since $\psi_0 = 1$ by economic theory, there are $T_j + 2$ parameters to estimate in this model. The α parameter is identified through contemporaneous variation in the level of the public good benefit. Similarly, the γ parameter is identified through contemporaneous variation in the level of cost of the policy. That is, α and γ can be identified without consid-

¹⁰Viscusi et al. (2008) provide the first example of a study designed to infer discount rates for public goods.

¹¹I am imposing risk neutrality here on preferences because instantaneous utility is linear and additively separable in income.

ering the discounting. Then, the discounting parameters (ψ_t) are identified through variation over time. If there is not enough variation in the data to identify each ψ_t individually, structure can be placed on the type of discounting.¹² For example, with quasi-hyperbolic discounting, there are only two discounting parameters (β, δ). Exponential discounting imposes the restriction that $\beta = 1$ in equation 5. A likelihood ratio test on the constrained and unconstrained models determines whether I reject the null hypothesis that $\beta = 1$. This is an improvement over previous studies which tend to assume a specific functional form for discounting.

Until now I have ignored any potential correlation across utilities for a given individual. However, each individual usually evaluates multiple choice scenarios in stated preference surveys. A more realistic model should therefore build in correlations among the utilities for different alternatives faced by an individual. Both random coefficients and error components specifications can accomplish this and are formally equivalent (Train, 2003). As previously stated, there are several ways to incorporate random parameters. I investigate three different random parameter approaches for the Minnesota River Basin application; one treats only discounting parameters as random, another specifies random coefficients on the public good benefit and public good cost, and one assumes all model parameters (ψ , α , and γ) are random. The simulated log likelihood functions for these random parameters specifications are in the appendix.

2.5 Application to a Monetary Choice

Suppose that an intertemporal monetary choice, j , describes a real or hypothetical amount of money, m_{ijt} , that will be paid to or collected from individual i , in time period t . During any time period, an individual receives utility from their non-experimental income, I_{it} , and the money from the experiment, m_{ijt} . Specify the deterministic portion of instantaneous utility as

$$v_{ijt} = \gamma(I_{it} + m_{ijt}), \tag{17}$$

¹²It is difficult to imagine a revealed preference data source that would provide sufficient intertemporal variation to identify each discount factor separately. One could imagine a clever stated preference survey in which respondents make a huge number of choices with different time horizons that would facilitate such an estimation of each time period's discount factor. I do not take this approach in my survey so in this paper I restrict attention to exponential, hyperbolic, and quasi-hyperbolic functional forms.

where γ is the marginal utility of money. Substituting equation 17 into equation 8 results in

$$U_{ij} = \sum_{t=0}^{T_j} \psi_t[\gamma(I_{it} + m_{ijt})] + \epsilon_{ij}. \quad (18)$$

The choice probabilities and likelihood equation are then derived with straightforward substitutions into equations 13 and 14.

3 Data

3.1 The Minnesota River Basin Survey

I administered a survey to 250 Minnesota residents in January of 2008 to gain information about individuals' preferences for immediate and delayed environmental improvements. The specific environmental improvement utilized for the survey is a proposal to improve water quality in the Minnesota River Basin (MRB). The Minnesota River Basin¹³ encompasses an area surrounding the Minnesota River of approximately 16,770 square miles, or roughly ten million acres. The Minnesota River flows from Big Stone Lake on the Minnesota / South Dakota border until it joins the Mississippi River at Fort Snelling near St. Paul, Minnesota (see Figure 3). As Minnesota's largest tributary to the Mississippi River, the Minnesota River doubles the volume of the Mississippi at their confluence. Largely rural, 92 percent of the land around the Minnesota River Basin is used for agricultural purposes. With such a high percentage of agricultural use, nonpoint source water pollution is a major problem for the Minnesota River Basin. According to the Minnesota River Data Center at Minnesota State University, Mankato (2007), the Minnesota River is one of the most seriously polluted rivers in the state of the Minnesota and the United States. The first portion of the survey familiarizes respondents with the study area and provides basic information about the sub-basins within the MRB. This portion of the survey also gathers some information about the respondents' prior knowledge and use of the study area. Next, the survey explains the current environmental situation of the MRB and informs respondents that the Minnesota Pollution Control Agency is currently developing pollution limits for impaired areas in the MRB. Additionally, respondents receive information about the specific actions that the Minnesota Pollution Control Agency is considering for improving water quality in the MRB.

¹³The following description borrows heavily from the description developed by Nicholas Flores (2008).

Figure 3: The Minnesota River Basin



The subsequent portion of the survey presents discrete choice experiment questions to respondents. Discrete choice experiments have several desirable characteristics for valuation purposes. One can define multiple attributes of an environmental policy and estimate the partworths of each attribute instead of valuing the policy as a whole. By including cost as an attribute, one can calculate the value of other attributes. In the context of this paper, I can calculate the willingness to pay for river basin improvements by including an attribute for the percentage of the MRB that is cleaned and an attribute for cost of the policy per year. Finally, attribute-based questions simulate choice scenarios that individuals make on a daily basis. Respondents are presented with a choice set and must select their most preferred option.

In order to introduce the time dimension of this study, the survey explains that while water quality improvements may sometimes happen now there may be situations where improvements will not occur until some time in the future. I ask respondents to consider policies that would temporarily increase water quality in the Minnesota River Basin. This survey design follows the strategy utilized by Viscusi et al. (2008). There are two alternatives available to the respondent for each choice scenario and each policy is defined by three attributes:

- **Percentage of Basin Cleaned.** This is the percentage of the Minnesota River Basin's surface waters having water quality high enough after the cleanup is fulfilled to: 1) maintain healthy, diverse and successfully reproducing populations

of aquatic organisms, including invertebrates such as crayfish and snails and vertebrates such as fish, AND, 2) be suitable for swimming and other forms of water recreation. Once the cleanup is fulfilled, the improvement lasts for a total of five years. After five years of improved water quality, funding for the policy ends and water quality returns to its previous level.

- **Cost of the Policy Per Year.** This is the amount of money that a household would have to contribute per year in the form of increased state income taxes. For each policy, your household would make five yearly payments of increased taxes, beginning immediately this year.
- **Time When Cleanup is Fulfilled.** This is the year in which the cleanup is fulfilled to the level of the policy. The water quality improvement ends five years after the cleanup is fulfilled.

"Percentage of Basin Cleaned" has three levels and ranges from fifty to seventy percent. "Cost of the Policy Per Year" has three levels and ranges from \$100 to \$300 per year. "Time When Cleanup is Fulfilled" has six levels: zero, one, two, three, four, and five years from now.

The survey goes on to reinforce a few key components of the policies. Respondents are instructed that cleanups that happen in the future are just as certain as cleanups that happen today. While I don't explicitly survey respondents' perceptions about their confidence of future rewards, I argue that my assertion about the relative certainty of future cleanups is plausible because once a public policy is passed, a law that stipulates a future cleanup is just as binding as a law that calls for an immediate cleanup. I establish a common reference point for respondents by stating that zero percent of the basin cleaned means that none of the surface waters of the MRB are cleaned up. To remove any potential heterogeneity in prior beliefs, I state that zero percent cleaned is the current situation. Similarly, one hundred percent of the basin cleaned means that all of the surface waters of the MRB are cleaned to levels high enough to support fish and recreation. I also reiterate the timing of the costs and benefits of the proposed policies. Pretests indicated that respondents understood all of these clarifying points.

Survey participants from Survey Sampling International (SSI) completed the questionnaire on the internet. SSI maintains a large pool of internet respondents in Minnesota, the United States, and across 72 countries. Individuals voluntarily join SSI's panel with participation incentives. The survey was hosted on a University of Colorado server, and SSI notified their panel of Minnesota residents by email of the

online survey instrument. Interested participants were then redirected from SSI's website with a unique identification number to the survey instrument. Because of the unique identification number, I could avoid duplicate responses and allow respondents to save and continue at a later time. After removing partial responses from some individuals that did not complete all required questions, I was left with a sample of 237 individuals. Each respondent faced a series of eight attribute-based stated preference questions. All together, this yielded a total of 1819 choice occasions.

Table 3 presents demographic information for the study sample as well as the Minnesota population as a whole. The information on the Minnesota population comes from the U.S. Census Bureau, 2005-2007 American Community Survey (2007). Some of the sample characteristics match up well with Minnesota demographics. For example, educational and race measures are similar between the survey sample and the Minnesota population. On the other hand, the study sample is clearly older and weighted more heavily toward females than the Minnesota population.¹⁴ The median income of the survey sample is lower than that of the Minnesota population, but I have an imprecise measure of income from the survey.¹⁵ While this is not a representative sample, there is enough variation to test for differences based upon personal characteristics. Also, this sample is much more diverse than the average sample of undergraduate students utilized for experimental work.

3.1.1 Survey Design

To identify the discounting parameters in this model it is essential that there is enough intertemporal variation. The survey design must provide enough variation while still remaining plausible and comprehensible to the survey respondents. That is, one must consider the real world decision so that policy options make sense. In this section I explore how the survey design can affect the ability to identify discount factors.

For illustrative purposes, first consider the two extremes of intertemporal variation. On one extreme, if benefits and costs vary across policies, individuals, and all time periods, model parameters are overidentified. At the other extreme, if all policies have the same time horizon and costs and benefits do not vary across time periods within policies, discounting parameters are not identified. This follows from the general property of the RUM that parameters that only affect scale of utility are

¹⁴As shown later in Table 10, neither age nor income is significantly related to model parameters.

¹⁵Participants reported income on the survey by selecting one of the 10 bracketed ranges. The lower bound on the highest bracket was \$150,000. As a result, all participants that fell in the highest category were coded with an income of \$150,000. In actuality, many of these individuals likely earned more than this amount.

Table 3: Demographics for the Minnesota River Basin Survey and the State of Minnesota

<i>Characteristic</i>	MRB Survey Sample	Minnesota Population
<i>Median Age</i>	51.0	36.9
<i>(Standard Deviation)</i>	15.16	
<i>Percent Male</i>	21.1	49.8
<i>Median Household Income</i>	\$ 42,500	\$ 55,616
<i>(Standard Deviation)</i>	\$ 34,324	
<i>Percent High School Degree or Higher</i>	98.3	90.7
<i>Percent Bachelor's Degree or Higher</i>	34.2	30.6
<i>Percent White</i>	92.4	88.0
<i>Percent Black or African American</i>	1.3	4.3
<i>Percent American Indian or Alaska Native</i>	2.1	1.0
<i>Percent Asian</i>	2.1	3.4

not identified. As the discounting parameters get larger the scale of utility increases but choice behavior is not at all impacted because all policies get more attractive at the same rate.

When considering the design of the survey it is important to make the policy choices as close to a real life situation as possible. In the case of public goods, I believe that it is most realistic to have costs uniformly start today and benefits start with a delay of zero to Q years, with Q selected such that respondents still believe that the policy will affect them. It is common for taxes to begin now and continue with a specific duration at the same cost per year and benefits to arrive at different times in the future at the same level of benefits per year.

Huber and Zwerina (1996) identify four principles in designing survey questions for choice experiments to make the choice design more efficient. Level balance means that each level of an attribute should occur an equal number of times in the survey. Orthogonality means that attributes should not be correlated in the design stage. Minimal overlap stipulates that attribute levels should be repeated within choice sets as little as possible. Utility balance attempts to balance the utility of the alternatives within a choice set. It is not generally possible to simultaneously uphold all four of these design principles. One popular quantitative measure of design efficiency is D-error = $|\Sigma|^{1/k}$, where Σ is the covariance matrix of the maximum likelihood estimator in the conditional logit model and k is the number of parameters in the model. By minimizing D-error, the researcher can approximately satisfy the four design principles.

Clearly, utility balance can only be achieved when the researcher has some a priori information about the parameters to be estimated. Huber and Zwerina (1996) show that even when the parameter estimates are incorrect there are efficiency gains from using them in the survey design. The SAS *choiceff* macro directly minimizes D-error to generate efficient choice designs for the conditional logit model, allowing the researcher to use a priori estimates on model parameters. To my knowledge, no research exists on design efficiency for more complicated models like the one proposed here. However, meeting the design principles for the simple conditional logit model should provide a good design for the more complicated model. Applying this reasoning, I use the conditional logit results from Viscusi et al. (2008) as the parameter estimate inputs for the *choiceff* macro to create my survey design.

There is little consensus on how many choice sets to create or how many choice sets each individual should face. Respondents can become fatigued when faced with too many choice sets. Additional choice complexity has been shown to increase inconsistencies in respondents' choices as in DeShazo and Fermo (2002). Including too few choice sets can lead to an inability to estimate the desired parameters. In my design, I generate 32 choice sets and divide them into four versions so that each respondent answers eight choice questions. I identify discount factors by varying the level of the River Basin cleanup and the number of years until the cleanup is fulfilled. The cost attribute facilitates estimation of the per-year willingness to pay but does not affect the discount rate because costs have the same time dimension over all alternatives. Simulation analysis confirms that this survey design is sufficient to identify the discounting parameters.¹⁶

¹⁶Simulation results are not presented in this paper but are available from the author.

As previously stated, each choice set contains two alternatives, each with three attributes: "percentage of basin cleaned," "cost of the policy per year," and "time when cleanup is fulfilled." The first two attributes each have three levels, while the third attribute has six levels. Table 4 shows a sample discrete choice question from one of the survey versions.¹⁷

Importantly, note that this survey design keeps the length of delay before the more delayed cleanup independent from the length of the interval between cleanup alternatives. Also, since this is a public good choice, transaction costs are not a factor. Once a respondent indicates their preferred policy, they no longer have a role in the execution of the policy. The effort required of the respondent is no different whether the cleanup happens today or years from now. Finally, uncertainty in the receipt of a future reward is also minimized as much as possible in this survey. Explicit instructions are repeated in the survey that there is no difference in the probability of cleanup for a policy with immediate benefits versus one with delayed benefits, and it is hoped that respondents believe these assertions.

I do not have an experimental design in this survey that allows the testing of the assumption of linear utility while simultaneously estimating discounting parameters because I am using the tradeoff between the percentage of the basin cleaned and the timing of the improvement to identify discount rates. However, I can specify a model where I represent the levels of basin improvement and cost with indicator variables and use researcher imposed discounting parameters to test for linear changes from level to level. For basin improvement, and assuming an exponential discount rate of eleven percent, the ratio of the difference between the utilities of the first two

¹⁷See the online Appendix for the attribute levels in all versions. The survey in its entirety is available from the author.

Table 4: (Sample Question) If you had to choose, would you prefer Policy A or Policy B?

Characteristic	Policy A	Policy B
Percentage of Basin Cleaned	50%	70%
Cost of Policy Per Year	\$100	\$300
Time When Cleanup is Fulfilled	2 Years From Now (2010)	Now (2008)
<i>Check the box of the policy you prefer</i>	<input type="checkbox"/>	<input type="checkbox"/>
	I prefer Policy A	I prefer Policy B

levels to the difference between the utilities of the second two levels is 1.06. The analogous ratio for cost is 0.85. A ratio of 1 would indicate utility that is linear in that particular attribute so I confirm that utility is approximately linear in the amount of basin improvement and cost.

In any case, the results of Andersen et al. (2008) and Andreoni and Sprenger (2010a) suggest that the discount rates (factors) estimated from a linear utility specification would be upper (lower) bounds on the discounting parameters from a concave utility function. Also, if uncertainty did in fact still remain regarding the delivery of future improvements, this would serve to make survey participants appear more present-biased and hyperbolic in their discounting. That is, uncertainty would tend to make participants look less like exponential discounters. Thus, the finding of exponential discount rates would be strengthened by the argument that there is residual decision-maker uncertainty.

3.2 Italian Money Data

I also have money choice data from Alberini and Chiabai's (2007) survey of 776 Italian residents.¹⁸ In this survey, respondents choose between a hypothetical immediate lump sum payment and a hypothetical stream of constant payments over 10 years. The lump sum payment option is always €10,000 received now. The stream of constant payments option is varied with annual payments of €1150, €1500, or €1650. The respondents also have a third option of being indifferent between the lump sum and the annuity. In this analysis, to match the analysis performed in Alberini and Chiabai, I omit the observations for which the respondent is older than 59. This leaves 505 observations.

3.3 State Lottery Lump Sum vs. Annuity Choice Data

In addition to the two stated-preference data sources already presented, I introduce a data set containing choices that lottery jackpot winners have made between lump sum and annuity payment options. Many states offer winners the option between a smaller lump sum payout and a larger sum of annual payments—the annuity option. Winners make choices over huge sums of money, providing a rich source of revealed-preference data. I have gathered data from three different state lotteries: Colorado Lotto, Texas Lotto, and Florida Lotto. These three states have open records laws

¹⁸I am grateful to the authors for generously sharing these data.

which facilitated collection of the data.¹⁹

Annuity options are defined by two variables; the number of annual payments and the dollar amount of each payment. Comparing the stream of payments option to the lump sum option, one can calculate the implicit interest rate of the annuity. The implicit interest rate is the rate that equates the present value of the annuity stream to the lump sum option. An individual prefers the lump sum payment over the annuity payments if the lump sum value exceeds their own internal present value of the annuity. Equivalently, an individual prefers the lump sum payment over the annuity payments if their internal (exponential) discount rate is higher than the implicit interest rate offered in the annuity. The less patient the individual, the more likely they will be to take the lump sum option. By observing the choices that winners make between the two options at multiple implicit interest rates, I am able to identify the average discount rate for lottery winners.

All three of these state lotteries advertise the dollar amount of the annuity option. Colorado and Florida allow winners to choose whether they want the lump sum or the annuity option after winning when claiming the prize. However, Texas requires winners to select their payment option when purchasing the ticket. Texas provides information to lottery players about the estimated lump sum payment for a given drawing. Therefore, I use the actual lump sum and annuity options offered to winners for Colorado and Florida but rely on the advertised lump sum and annuity options available to Texas lottery winners at the time of ticket purchase. Currently, the Colorado Lotto stipulates that the lump sum option is 50 percent of the annuity option. Prior to November of 2003, Colorado's lump sum option was equal to 40 percent of the annuity option. Alternatively, Texas and Florida's lump sum option varies as a percentage of the annuity option. Therefore, I get variation in the implicit annuity interest rate over the lotteries, which aids identification of the discount rate. Florida Lotto 30(20)-yr offers 30(20) annual payments for the annuity option. The payments are all large enough that Federal and State marginal tax rates for the last dollars on the lump sum and annuity payments are the same so tax rates should not bias the results.

Table 5 summarizes the data for the three state lotteries. As expected, the lottery with the highest implicit annuity interest rate (Colorado Lotto 40 percent) has the lowest percentage of winners choosing the lump sum option. The lotteries with

¹⁹Colorado's lotto data was publicly available on the internet at <http://www.coloradolottery.com>. Texas and Florida's lottery agencies responded to my requests for data. Unfortunately, no information on personal characteristics of winners is available.

Table 5: Summary of State Lotteries. Lump Sum Vs. Annuity Options.

<i>Lottery</i>	Date Range	N	Lump Sum / Annuity	Implicit Annuity Interest Rate	% Choosing Lump Sum
Colorado Lotto 40%	08/20/1994– 10/25/2003	177	40%	9.98%	60.45%
Colorado Lotto 50%	11/12/2003– 1/05/2008	37	50%	6.97%	86.49%
Texas Lotto	10/27/2001– 12/08/2007	74	54.7% to 64%	5.89% to 4.16%	82.43%
Florida Lotto 30-yr	11/28/1998– 12/22/2007	343	42.5% to 70.3%	7.45% to 2.64%	91.80%
Florida Lotto 20-yr	10/24/1998– 11/14/1998	5	64.5% to 64.7%	5.2% to 5.15%	60.00%
Total		636			81.43%

Note: The implicit annual interest rate is the interest rate that equates the present value of the sum of annuity payments to the lump sum option.

Table 6: Frequency of Lump Sum Choices For Texas and Florida Lottery Winners by Implicit Annuity Interest Rate.

<i>Lottery</i>	N	Lump Sum/Annuity	Implicit Annuity Interest Rate	% Choosing Lump Sum
Texas Lotto	17	54.7% to 56.99%	5.89% to 5.43%	58.82%
	16	57% to 58.99%	5.4% to 5.04%	81.25%
	23	59% to 60.49%	5.02% to 4.76%	91.30%
	18	60.5% to 64%	4.75% to 4.16%	94.44%
Florida Lotto 30-yr	86	42.5% to 50.64%	7.45% to 5.6%	84.90%
	84	50.68% to 54.64%	5.59% to 4.85%	94.00%
	85	54.7% to 57.5%	4.83% to 4.38%	94.10%
	88	57.49% to 70.3%	4.37% to 2.6%	94.30%

the lowest implicit annuity interest rates (Texas and Florida 30-yr) have the lowest percentage of winners choosing the lump sum option. Within Colorado, moving from an implicit annuity interest rate of 9.98 percent to 6.97 percent results in a jump in the percentage of winners choosing the lump sum option from 60.45 percent to 86.49 percent.

Table 6 illustrates the frequency of the lump sum choices by Texas and Florida lottery winners broken down by implicit annuity interest rate. For Texas, higher implicit annuity interest rates correlate with a lower percentage of winners choosing the lump sum option. Almost half of the Texas winners faced with implicit annuity interest rates higher than 5.43 percent choose the annuity option whereas only about 5.5 percent of the Texas winners facing implicit annuity interest rates lower than 4.75 percent choose the annuity option. There is less variation in the percentage of winners choosing the lump sum option in the Florida Lotto. However, it holds that fewer winners choose the lump sum option with a higher implicit annuity interest rate. Clearly, winners are considering the implicit annuity interest rate.

One expects decision makers to perceive more credibility in the receipt of a future reward for an official state lottery than for a laboratory experiment. Transaction costs are likely to be minimal for the receipt of future lottery payments because

Table 7: Magnitude of the State Lottery Lump Sum Options.

<i>Lottery</i>	Mean Lump Sum Option	Median Lump Sum Option	Standard Deviation Lump Sum Option	N
Colorado Lotto 40%	1,996,676.76	1,600,000.00	1,401,166.07	177
Colorado Lotto 50%	1,691,205.00	1,416,067.00	1,121,074.35	37
Texas Lotto	12,366,727.35	7,993,148.73	13,263,747.56	74
Florida Lotto 30-yr	6,813,203.19	4,374,972.36	6,457,104.32	343
Florida Lotto 20-yr	4,567,324.61	4,866,004.40	408,984.15	5
All Data	5,803,285.40	3,483,569.56	7,341,957.70	636

payments are spelled out explicitly in the annuity agreement. Also, transaction costs will be much less significant as a percentage of the huge sums of money at stake here compared to the small rewards in laboratory experiments. As shown in Table 7, the average size of the lump sum option throughout the data set is almost six million dollars. Finally, the nature of the lottery choice is different from most laboratory choices. Instead of comparing various one time payments to an anchor time period, a stream of annual payments is compared to a lump sum. Thus, there is no confounding of length of delay with interval length so subadditive discounting would not be a concern.

4 Estimation Results

4.1 Results for the Minnesota River Basin Survey

In Specifications I and II, I ignore the panel nature of the data. That is, I assume that utilities are not correlated across choice occasions for a given individual.²⁰ In Specification I, no interactions are assumed and I have $U_{ij} = \sum_{t=0}^{T_j} \psi_t[\alpha q_{ijt} + \gamma(Y_{it} - c_{ijt})] + \epsilon_{ij}$. Results²¹ for Specification I are shown in Table 8. In Model I.a I

²⁰This is not a realistic assumption so these are not my preferred specifications. I include them to serve as a baseline. Also, the random coefficients specifications require simulated maximum likelihood estimation methods, so I present basic maximum likelihood estimates first.

²¹I utilize the unconstrained minimization routine in Matlab's (2006) Optimization Toolbox V3.0.4 to minimize the negative of the log likelihood function as in Equations 13 and 14. The asymptotic

assume exponential discounting. Model I.b assumes HM hyperbolic discounting, Model I.c assumes Harvey hyperbolic discounting, and Model I.d assumes quasi-hyperbolic discounting. All coefficients are highly significant for each estimation, with the exception of β in Model I.d. In Model I.a, I estimate a constant discount factor of $\widehat{\delta} = .908$. This is equivalent to an estimated discount rate of $\widehat{r} = 10.2$ percent. In other words, individuals discount the future at a constant rate of 10.2 percent per year. In Model I.b, for the HM model, $\widehat{\omega} = 0.148$. Recall that as ω goes to zero the HM discounting model becomes the exponential discounting model. Thus, $\widehat{\omega} = 0.148$ suggests that the best fitting HM hyperbolic model is close to an exponential model. For the Harvey Model in Model I.c, $\widehat{\mu} = 0.389$. This is in line with estimates from previous studies. In Model I.d, for quasi-hyperbolic discounting, $\widehat{\beta} = .909$. Recall that the β parameter measures the extent of the departure from the exponential discounting model. Quantitatively, $\widehat{\beta} = .909$ does not represent a large deviation from the exponential assumption and is statistically not different from 1.

In specification I, the exponential discounting model has the second highest log likelihood value. Thus, I do not find evidence that either of the single parameter hyperbolic models are preferred to the exponential model. Models I.a and I.d are nested; Model I.c is a special case of Model I.d with the restriction that $\beta = 1$. Therefore, I can perform a likelihood ratio test with the null hypothesis that $\beta = 1$. From Table 8, the likelihood ratio test statistic is equal to 0.4393.²² Hence, I fail to reject the null hypothesis that $\beta = 1$. There is no evidence in this first specification to support quasi-hyperbolic discounting over the standard exponential model.

In specification II, I model heterogeneity by assuming that α and γ depend upon observed personal characteristics; $\alpha = \alpha_0 + \lambda x_{it}$ and $\gamma = \gamma_0 + \xi x_{it}$, where x_{it} is a vector of personal characteristics for individual i at time t . Gathered personal characteristics that could potentially influence utility include age, income, sex, education level, and whether the respondent resides within the Minnesota River Basin. Table 9 reports results for this specification.^{23,24}

standard errors for the maximum likelihood parameter estimates, $\widehat{\beta}$, are estimated with the diagonal entries of $\sqrt{H^{-1}}$, where H is the Hessian matrix of second derivatives = $\frac{\partial^2 LL(\widehat{\beta})}{\partial \beta \partial \beta}$ evaluated at the optimum. The Hessian is calculated by the BFGS method.

²²The test statistic is equal to twice the difference of the log likelihoods and is distributed chi-square with one degree of freedom.

²³I also examine specifications treating the discounting parameters as linear functions of observable characteristics and α and γ as random coefficients for all discounting models. However, all observable personal characteristics are insignificant in these specifications.

²⁴The positive and significant coefficient on *income*/1000000 in the γ estimation implies that the marginal utility of income increases as income increases. This is counterintuitive so I run a specification where γ is still a function of personal characteristics but α is not. In this case, the

Table 8: Minnesota River Basin Maximum Likelihood Results: Specification I

<i>Variable / Parameter</i>	I.a	I.b	I.c	I.d
Basin Improvement/100	2.620 *** (0.272)	2.658 *** (0.27)	2.644 *** (0.273)	2.662 *** (0.277)
Net Income/100	0.321 *** (0.0235)	0.320 *** (0.0234)	0.317 *** (0.0233)	0.322 *** (0.0234)
Exponential (δ)	0.908 *** (0.0088)			0.912 *** (0.0104)
HM (ω)		0.148 *** (0.0211)		
Harvey (μ)			0.389 *** (0.0366)	
Quasi-Hyperbolic (β)				0.909 (0.119)
Log L	-1145.89	-1146.81	-1150.09	-1145.64

Note: Asymptotic Standard Errors are given in parentheses.
 * significant at 10%, **significant at 5%, *** significant at 1%
 The t-tests for β and δ are against 1. All others are tested against 0.

Table 9: Minnesota River Basin Maximum Likelihood Results: Specification II

<i>Variable / Parameter</i>	II.a	II.b	II.c	II.d
Exponential (δ)	0.906 *** (0.009)			0.912 *** (0.01)
HM (ω)		0.152 *** (0.021)		
Harvey (μ)			0.396 (0.035)	
Q-H (β)				0.883 (0.113)
<i>α (M.U. Public Good) Estimation</i>				
Intercept*100	3.702 *** (1.096)	3.561 *** (1.116)	3.045 *** (1.091)	3.611 *** (1.118)
Age/1000	-0.218 (0.157)	-0.213 (0.16)	-0.192 (0.156)	-0.217 (0.16)
Income/1000000	0.252 *** (0.079)	0.266 *** (0.0807)	0.275 *** (0.0808)	0.266 *** (0.0832)
Male/100	-0.395 (0.577)	-0.366 (0.59)	-0.258 (0.585)	-0.354 (0.593)
Education/1000	-1.841 (1.6)	-1.585 (1.634)	-0.101 (0.161)	-1.623 (1.639)
MRB Resident/100	-0.610 (0.463)	-0.685 (0.472)	-0.763 (0.464)	-0.671 (0.477)
<i>γ (M.U. Net Income) Estimation</i>				
Intercept*100	0.3055 *** (0.106)	0.2941 *** (0.106)	0.2573 *** (0.102)	0.2980 *** (0.107)
Age/1000	-0.0044 (0.016)	-0.0049 (0.016)	0.5847 (1.488)	-0.1248 (0.16)
Income/1000000	0.0190 *** (0.0075)	0.0193 *** (0.0075)	0.0189 *** (0.0072)	0.0194 *** (0.0076)
Male/100	-0.0319 (0.0551)	-0.0302 (0.0552)	0.2576 (0.5854)	-0.0293 (0.0555)
Education/1000	-0.1391 (0.159)	-0.1210 (0.159)	0.1007 (0.161)	-0.0047 (0.0161)
MRB Resident/100	-0.0511 (0.0451)	-0.0549 (0.0452)	-0.0581 (0.0437)	-0.0544 (0.0455)
Log L	-1135.72	-1136.43	-1140.19	-1135.28

Note: Asymptotic Standard Errors are given in parentheses.

* significant at 10%, ** significant at 5%, *** significant at 1%

The t-tests for β and δ are against 1. All others are tested against 0.

Results for the discounting parameters in specifications II.a-d are similar to results from specifications I.a-d. Again, the exponential discounting model fits the data slightly better than the two single-parameter hyperbolic models, although non-nested model preference is somewhat indefinite. Viewing Model II.a as a restricted version of Model II.d I can again perform a likelihood ratio test. The test statistic is equal to .879 so I fail to reject the null hypothesis that $\beta = 1$. There is no evidence in this interactions model in support of quasi-hyperbolic discounting.

In specification III, I assume that discounting parameters are random coefficients. Specifically, I assume that discounting parameters vary over people but are constant over choice situations for each person. In Model III.a, I assume exponential discounting with a discount factor, δ_i , that is distributed normally with mean $\bar{\delta}$ and variance z_δ^2 . In Model III.b, I assume that the single parameter for HM hyperbolic discounting, ω_i , is distributed lognormally, where $\ln(\omega) \sim N(\bar{\omega}, z_\omega^2)$.²⁵ Model III.c assumes hyperbolic discounting with the single parameter, μ_i , being distributed normally with mean $\bar{\mu}$ and variance z_μ^2 . Model III.d assumes quasi-hyperbolic discounting with a constant β factor and a δ_{Hi} factor that is distributed normally with mean $\bar{\delta}_H$ and variance $z_{\delta_H}^2$. Attempts to treat both the β factor and the δ_H factors as random fail to converge when I assume normal distributions. Thus, I reparameterize assuming that the β and δ_H factors are distributed lognormally for Model III.e.²⁶

Table 10 gives results for the random discount factors specification. The maximized value of the simulated log likelihood equation is greater in the exponential specification (Model III.a) than in either of the single parameter hyperbolic specifications (Models III.b and III.c). Freeing up the additional β parameter in Model III.d leads to only a miniscule improvement in the simulated log likelihood at convergence compared to Model III.a with β restricted to one. Finally, both the mean and standard deviation of the log of the quasi-hyperbolic β parameter in Model III.e are statistically insignificant and model fit improves only slightly. Once again, I fail to reject the null hypothesis of exponential discounting.

As seen in Table 10, significant heterogeneity exists in the discounting parameters throughout all four discounting models. There are especially wide distributions in the single parameter hyperbolic specifications. When accounting for heterogeneity in

significance on the *income/1000000* variable in the γ estimation disappears.

²⁵When assuming a normal distribution for ω , large draws from the normal distribution imply large negative discount factors, which are theoretically impossible and cause problems for maximization of the simulated log likelihood equation. Thus, I assume a lognormal distribution for ω .

²⁶This reparameterization comes at a cost; it makes the nested comparison of the exponential and quasi-hyperbolic models less obvious.

Table 10: Minnesota River Basin Simulated Maximum Likelihood Results: Specification III

<i>Variable / Parameter</i>	III.a	III.b	III.c	III.d	III.e
Basin Improvement/100	3.408 *** (0.274)	3.598 *** (0.294)	3.360 *** (0.28)	3.404 *** (0.282)	3.429 *** (0.3)
Net Income/100	0.388 *** (0.027)	0.382 *** (0.0269)	0.372 *** (0.0262)	0.388 *** (0.0269)	0.391 *** (0.0283)
Exponential (δ)					
<i>Mean</i>	0.899 *** (0.0109)			0.899 *** (0.012)	
<i>S.D.</i>	0.107 *** (0.0128)			0.107 *** (0.013)	
ln(ω) (HM)					
<i>Mean</i>		-2.248 *** (0.205)			
<i>S.D.</i>		1.860 *** (0.306)			
Harvey (μ)					
<i>Mean</i>			0.418 *** (0.045)		
<i>S.D.</i>			0.411 *** (0.054)		
Quasi-Hyperbolic (β)				1.006 (0.116)	
ln (β)					
<i>Mean</i>					0.0234 (0.133)
<i>S.D.</i>					-0.374 (0.289)
ln(δ)					
<i>Mean</i>					-0.111 *** (0.014)
<i>S.D.</i>					0.114 *** (0.015)
Simulated Log L	-1111.22	-1127.43	-1126.38	-1111.22	1110.46

Note: Asymptotic Standard Errors are given in parentheses.

* significant at 10%, **significant at 5%, *** significant at 1%

The t-tests for β and δ are against 1. All others are tested against 0.

Table 11: Point Estimates of Mean, Median, and Standard Deviation of (ω) and (β) and (δ) Parameters from Specification III

<i>Parameter</i>	Median	Mean	Std. Dev.
HM (ω)	0.106	0.595	3.301
Quasi-Hyperbolic (β)	1.024	1.098	0.426
Quasi-Hyperbolic (δ)	0.8948	0.901	0.103

discount factors, the average exponential discount factor is 0.899, which corresponds to an average annual exponential discount rate of 11.2 percent. Point estimates of the median, mean, and standard deviation for logged discounting parameters in specification III are shown in Table 11.

In specification IV, I assume that discount factors do not vary over people but coefficients α and γ do vary over individuals. Specifically, I assume for each of the four discounting models that α_i and γ_i are distributed normally across individuals with respective means $\bar{\alpha}$ and $\bar{\gamma}$ and variances z_α^2 and z_γ^2 . Table 12 presents results for specification IV.

Examining the maximized values of the simulated log-likelihood functions in Table 12, the exponential model IV.a fits the data better than either of the single parameter hyperbolic models. A simulated likelihood ratio test fails to reject the null hypothesis that $\beta = 1$.²⁷ Also, a $\hat{\beta} = 1.183$ does not indicate any present bias. In Model IV.a, a $\hat{\delta} = .876$ implies a constant discount rate of $\hat{r} = 0.141$.

Finally, in specification V, I assume that all model parameters are random. Table 13 presents results for specification V and Table 14 shows point estimates for the logged parameters. The same pattern from the previous specifications holds in that the exponential model (V.a) achieves a higher maximized log likelihood value than does either of the single parameter hyperbolic models. Comparing Model V.d to Model V.a we once again fail to see significant model improvement with the un-

²⁷Lee (1999) shows that a simulated likelihood ratio test statistic equal to twice the difference between the maximized values of the unconstrained and constrained simulated log likelihood functions is asymptotically chi-square distributed. The asymptotic chi-square distribution of the test statistic has a non-centrality parameter k and v degrees of freedom, where v equals the number of constraints. Lee (1999) also finds that the non-centrality parameter k is negligible when the number of simulation draws is large enough relative to the sample size. In a Monte Carlo example, Lee finds that 100 draws is sufficient for a sample size of 200 to reasonably ignore the non-centrality parameter.

Table 12: Minnesota River Basin Simulated Maximum Likelihood Results: Specification IV

<i>Variable / Parameter</i>	IV.a	IV.b	IV.c	IV.d
Basin Improvement/100				
<i>Mean</i>	2.888 *** (0.402)	3.104 *** (0.396)	3.077 *** (0.395)	2.760 *** (0.413)
<i>S.D.</i>	3.429 *** (0.468)	3.487 *** (0.496)	3.220 *** (0.511)	3.368 *** (0.461)
Net Income/100				
<i>Mean</i>	0.395 *** (0.0405)	0.397 *** (0.0405)	0.391 *** (0.0403)	0.391 *** (0.0405)
<i>S.D.</i>	0.440 *** (0.0456)	0.435 *** (0.0453)	0.434 *** (0.0446)	0.440 *** (0.0456)
Exponential (δ)	0.876 *** (0.0125)			0.867 *** (0.0155)
HM (ω)		0.207 *** (0.0301)		
Harvey (μ)			0.471 *** (0.0449)	
Quasi-Hyperbolic (β)				1.183 (0.1652)
Simulated Log L	-1047.74	-1052.33	-1061.91	-1046.90

Note: Asymptotic Standard Errors are given in parentheses.
 * significant at 10%, **significant at 5%, *** significant at 1%
 The t-tests for β and δ are against 1. All others are tested against 0.

Table 13: Minnesota River Basin Simulated Maximum Likelihood Results: Specification V

<i>Variable / Parameter</i>	V.a	V.b	V.c	V.d
Basin Improvement/100				
<i>Mean</i>	4.490 *** (0.516)	4.834 *** (0.545)	4.464 *** (0.514)	4.469 *** (0.556)
<i>S.D.</i>	4.876 *** (0.687)	5.235 *** (0.69)	4.850 *** (0.735)	4.845 *** (0.687)
Net Income/100				
<i>Mean</i>	0.524 *** (0.0507)	0.511 *** (0.0496)	0.496 *** (0.0491)	0.531 *** (0.0548)
<i>S.D.</i>	0.492 *** (0.0567)	0.493 *** (0.0552)	0.503 *** (0.0554)	0.500 *** (0.0601)
Exponential (δ)				
<i>Mean</i>	0.887 *** (0.012)			
<i>S.D.</i>	0.0966 *** (0.012)			
ln(ω) (HM)				
<i>Mean</i>		-1.964 *** (0.189)		
<i>S.D.</i>		1.552 *** (0.233)		
Harvey (μ)				
<i>Mean</i>			0.449 *** (0.0451)	
<i>S.D.</i>			0.371 *** (0.047)	
ln (β) (Q-H)				
<i>Mean</i>				0.150 (0.13)
<i>S.D.</i>				0.307 (0.306)
ln(δ) (Q-H)				
<i>Mean</i>				-0.129 *** (0.015)
<i>S.D.</i>				0.107 *** (0.014)
Simulated Log L	-1008.50	-1024.53	-1030.44	-1008.16

Note: Asymptotic Standard Errors are given in parentheses.

* significant at 10%, ** significant at 5%, *** significant at 1%

The t-tests for β and δ are against 1. All others are tested against 0.

Table 14: Point Estimates of Mean, Median, and Standard Deviation of (ω) and (β) and (δ) Parameters from Specification V

<i>Parameter</i>	Median	Mean	Std. Dev.
HM (ω)	0.14	0.468	1.489
Quasi-Hyperbolic (β)	1.162	1.218	0.382
Quasi-Hyperbolic (δ)	0.879	0.884	0.095

restricted quasi-hyperbolic β parameter. Furthermore, the β parameter enters the model insignificantly. The point estimate of the median value of $\hat{\beta} = 1.162$ does not support quasi-hyperbolic discounting either. Appropriately modeling heterogeneity in hyperbolic discount rates in specification V.d makes the quasi-hyperbolic model an even less appealing descriptive model for discounting behavior in this study. The point estimate of the discount rate in Model V.a is $\hat{r} = 0.128$.

There is clearly heterogeneity in the discounting parameters for all discounting models as evidenced by the statistically significant estimates of discounting parameter standard deviations and the maximized values of the simulated log likelihood equations in specifications V.a through V.d. However, the point estimates of the means of discounting parameters in Model V.a through Model V.d are quite similar to those from specifications IV.a through IV.d where I did not account for heterogeneity in time preferences. Furthermore, the point estimates of the means of discounting parameters from the single parameter hyperbolic models in Model V.b and Model V.c are quite similar to those from Models I. and II. (b and c) where I do not account for correlation across utilities or the panel nature of the data. The exponential discount factor (rate) is slightly overestimated (underestimated) in Model I.a and Model II.a compared to the mean from the more appropriate specification V.a. The mean exponential discount rate from Model V.a is about 2.6 percentage points higher than the point estimate from Model I.a.

In summary, all specifications lead us to fail to reject the null hypothesis that $\beta = 1$. In other words, freeing up the additional β parameter in the quasi-hyperbolic framework does not significantly improve model fit over the standard exponential model. Finally, the maximized value of the log likelihood function is greater in the exponential cases (a) than in the hyperbolic cases (b and c) for all specifications. I conclude that there is not any evidence to prefer hyperbolic discounting over expo-

ponential discounting for any of the specifications.

4.2 Results for the Italian Money Data

With the two money data sets, I can estimate the discounting model assuming either standard exponential or single parameter hyperbolic functional forms. It is not possible to uniquely identify quasi-hyperbolic discount factors because I never observe choices that individuals make between two alternatives with future outcomes and different time profiles. That is, annuity and lump sum options are anchored to the present in all choices. Recall that the quasi-hyperbolic β parameter measures the extent of present bias, which we cannot hope to measure when all choices are anchored to the present.

I make some slight modifications to the empirical model for this data set because it includes observations on individuals who are indifferent between a lump sum payment and an annuity option. Denote the utility from the lump sum option as U_{LS} and the utility from the annuity option as U_A , where utility is defined as in equation 18. Then, let $U_{diff} = U_{LS} - U_A$. An individual prefers the lump sum option when $U_{diff} > k$, prefers the annuity option when $U_{diff} < -k$, and is indifferent when $-k < U_{diff} < k$. The probability that an individual chooses the lump sum is $1 - F(k - U_{diff})$, the probability that an individual chooses the annuity is $F(-k - U_{diff})$, and the probability that an individual is indifferent is $F(k - U_{diff}) - F(-k - U_{diff})$, where $F(\cdot)$ is the cumulative standard normal distribution. I estimate all model parameters including k and these ordered-probit results are given in Table 15.

Assuming exponential discounting, maximum likelihood estimation gives $\hat{\delta} = 0.906$. Since $\delta = 1/(1+r)$, this implies $\hat{r} = 0.103$. That is, individuals discount with a constant rate of 10.3 percent. For the Harvey hyperbolic model, I find $\hat{\mu} = .293$ and for the HM hyperbolic model, I estimate $\hat{\omega} = 0.135$. The maximized value of the log-likelihood function is identical under all three specifications. This data set does not provide enough information to prefer one discounting specification over the others.

4.3 Results for the State Lottery Data

Here, I apply equations 13, 14 and 18 to the state lottery data. Table 16 summarizes maximum likelihood results for exponential and hyperbolic specifications. Assuming an exponential discounting form leads to an estimate of 0.885 for the constant discount factor, which is equivalent to a constant discount rate of 12.99 percent. Returning to

Table 15: Results for Italian Money Data

<i>Discounting Model</i>	ψ	γ (<i>M.U. Money</i>)	k	<i>Log L</i>
<i>Exponential</i>	0.906*** (0.016)	0.0881*** (0.0388)	0.107*** (0.0158)	-465.750
<i>HM Hyperbolic</i>	0.135*** (0.0031)	0.0882*** (0.0390)	0.107*** (0.0158)	-465.750
<i>Harvey Hyperbolic</i>	0.293 *** (0.0501)	0.0882*** (0.0389)	0.107*** (0.0158)	-465.750

Note: Asymptotic Standard Errors are given in parentheses.

* significant at 10%, **significant at 5%, *** significant at 1%

t-test for exponential discount factor tests for δ not equal to 1. All other t-tests are against 0.

Table 5, we recall that over eighty percent of lottery winners in this sample chose the lump sum option. Among the lotteries, the Colorado Lotto 40% and the Florida Lotto 20-yr had the highest percentage of winners choosing the annuity option. Focusing on the Colorado Lotto 40%, we see that less than 40 percent of winners chose the annuity option when offered an implicit annuity interest rate of 9.98 percent. Thus, it makes sense that the discount rate for this data is estimated higher than the 9.98 percent of the Colorado Lotto 40%. From the variation we observe in this data, the model estimates that the average individual would switch from preferring the lump sum option to preferring the annuity option at an implicit annuity interest rate of approximately thirteen percent. The magnitude of the exponential discount rate is within the range of interest rates found in capital markets; it is in line with interest rates on personal credit cards for example. This implies that these individuals do roughly equate the marginal rate of substitution between two years to one plus the interest rate.

Discounting parameters in the two hyperbolic specifications are not statistically significant. Recall that, in this lottery data, I never observe any choices that involve trading off a future reward in one time period for a future reward in another time period. I only observe a lump sum versus an annuity that is paid out over a twenty to thirty year time period. Marginal discount rates decline over time in the Harvey and

Table 16: Results for State Lottery Data: n=636

<i>Discounting Model</i>	ψ	γ (<i>M.U. Money</i>)	<i>Log L</i>
Exponential	0.885*** (0.0377)	0.0265 *** (0.00846)	-322.896
HM Hyperbolic	0.484 (0.704)	0.0205** (0.011)	-323.914
Harvey Hyperbolic	0.966 (0.826)	0.172*** (0.00735)	-324.002

Note: Asymptotic Standard Errors are given in parentheses.

* significant at 10%, **significant at 5%, *** significant at 1%

The t-test for exponential discount factor is for δ not equal to 1. All other t-tests are against 0.

HM hyperbolic models making the hyperbolic discounter present-biased. Looking at Figure 2, we cannot hope to precisely determine the curvature of the marginal discount rate graphs for the hyperbolic models when we do not observe future tradeoffs, hence the large standard errors on the Harvey and HM discounting parameters. Conversely, we can precisely estimate exponential discounting parameters from this data because the discount rate remains constant for the duration of the annuity, hence the small standard error on δ . Note that there was not a problem precisely estimating hyperbolic discounting parameters in the Italian money application because the time-frame was much shorter. Examining Figure 1, discount factors track rather closely together for the first 10 years after which time there is a divergence.

Comparing the maximized values of the log likelihood functions, the HM hyperbolic model fits the data only slightly better than the Harvey hyperbolic model. The exponential specification is cautiously preferred to both of the hyperbolic specifications on the basis of non-nested model selection criteria keeping in mind that I do not have enough intertemporal variation here to precisely estimate the hyperbolic parameters. I can imagine two types of intertemporal variation that would facilitate the estimation of hyperbolic parameters for a relatively long time horizon such as the 20 to 30 year durations of the annuity options in this lottery data. One option would be if we could observe some individuals making the choice between a lump sum option and an annuity option with a much shorter time horizon of perhaps several years, observe some other individuals making the choice between a lump sum option

and an annuity option with a time horizon of perhaps ten years, and some others making the longer term annuity choices such as those which I observe in this lottery data. Another option would be if we could observe some individuals make the choice between lump sums of different magnitudes to be received at different points in the future. A third option would be if we could observe individuals choosing between immediate lump sums and annuities to be received beginning at different points in the future. I am not aware of any revealed preference data sources that resemble the second two options, so perhaps the best hope for estimating hyperbolic discounting parameters from revealed preference data would be to pool several data sources of lump sum versus annuity choices containing substantially different annuity durations.

One concern is that evidence suggests that lottery players may not be representative of the population. For example, the 2006 and 2007 Demographic Surveys of Texas Lottery Players (2007) find a statistically significant differences in participation due to income and employment status. And, the 2008 Demographic Survey of Texas Lottery Players (2008) finds a statistically significant difference in participation due to employment status only. Thus, it likely that the individuals in the lottery data set have a lower income on average than the general population. However, Clotfelter et al. (1999) find "that the differences among groups are much greater with respect to amount played than with respect to participation rate. Indeed, with few exceptions there is remarkable uniformity in participation" (Clotfelter et al., 1999). Therefore, I would not claim that this is the representative discount rate of the general population. However, according to previous research, there is significant heterogeneity in lottery players and participation rates may not vary tremendously across demographics.

Thirty-eight percent of survey respondents in 2007 and thirty-nine percent of survey respondents in 2008 said that they participated in Texas lottery games within the last year (University of Houston Center for Public Policy, 2007, 2008). Clotfelter et al. (1999) find that approximately half of the adult population in State Lottery states plays the lottery in any given year. Thus, lottery players do make up a substantial subgroup of the population so it is relevant to examine their average rates of time preference. The results are still informative for public policies that pertain to lottery players. For example, local, state, and federal tax policies apply to lottery winnings. It would be an interesting future extension to find a lump sum versus annuity choice situation where individual demographic information is available to see how much the lottery players differ in time preferences from the general population.

5 Conclusion

The empirical strategy introduced in this paper provides a method of estimating discounting factors that is consistent with utility-maximization theory. I structurally model intertemporal choices to produce explicit estimates of discounting parameters and then formally test several hypothesized discounting functions. This general estimation framework can be applied to private or public goods choices.

I apply the empirical model to three data sources which represent private and public goods, and stated-preference and revealed-preference choices. Estimation results from all of the data sources suggest that the standard exponential discounting model is at least as preferred as single-parameter specifications of the hyperbolic discounting model. Likelihood ratio tests of the quasi-hyperbolic model for the public goods data fail to reject the null hypothesis of standard exponential discounting. Estimates of the constant exponential discount rates range from approximately eight to fourteen percent throughout the three data sets.

Because of the nature of the data sets employed in this paper, confounding factors that are commonly part of experimental studies are minimized. Specifically, the data sets minimize perceived uncertainty in the receipt of future rewards, perceived future transaction costs, and the correlation between the length of delay before a future outcome and the length of the interval between two outcomes. This may be one contributing factor for the absence of hyperbolic discounting in this study.

I find no evidence that individuals fail to behave rationally when making intertemporal decisions. They are dynamically consistent in their choices and do not appear to be present-biased. The range of discount rates estimated here falls below the discount rates commonly found in the experimental literature but is consistent with interest rates that we see in capital markets, as we would expect from theory. This information is useful for conducting positive cost-benefit analysis of intertemporal policies; that is calculating the net present value of projects as individuals view them. Moreover, these results have implications for describing behavior in a variety of contexts including personal savings decisions, participation in preventative health programs, the formation of human capital, and environmental sustainability. Whether hyperbolic discounting should be used for the discounting of policies with long time horizons is a normative question and is not addressed in this research.

References

- Alberini, A. and Chiabai, A. (2007). Discount rates in risk versus money and money versus money tradeoffs. *Risk Analysis*, 27(2):483–498. M3: doi:10.1111/j.1539-6924.2007.00899.x.
- Andersen, S., Harrison, G. W., Lau, M. I., and Rutstrom, E. E. (2008). Eliciting risk and time preferences. *Econometrica*, 76(3):583 – 618.
- Andreoni, J. and Sprenger, C. (2010a). Estimating time preferences from convex budgets. Working Paper.
- Andreoni, J. and Sprenger, C. (2010b). Risk preferences are not time preferences: Discounted expected utility with a disproportionate preference for certainty. Working Paper.
- Bosworth, R., Cameron, T. A., and DeShazo, J. R. (2006). Preferences for preventative public health policies with jointly estimated rates of time preference.
- Cairns, J. and van der Pol, M. (1997). Saving future lives. A comparison of three discounting models. *Health Economics*, 6(4):341–350.
- Cairns, J. and van der Pol, M. (2000). Valuing future private and social benefits: The discounted utility model versus hyperbolic discounting models. *Journal of Economic Psychology*, 21(2):191–205.
- Cameron, T. A. and Gerdes, G. R. (2003). Eliciting individual-specific discount rates. Working Paper.
- Clotfelter, C. T., Cook, P. J., Edell, J. A., and Moore, M. (1999). State lotteries at the turn of the century: report to the national gambling impact study commission.
- Coller, M. and Williams, M. B. (1999). Eliciting individual discount rates. *Experimental Economics*, 2(2):107–127.
- DeShazo, J. R. and Fermo, G. (2002). Designing choice sets for stated preference methods: The effects of complexity on choice consistency. *Journal of Environmental Economics and Management*, 44(1):123 – 143.
- Frederick, S., Loewenstein, G., and O’Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, 40(2):351–401.

- Harrison, G. W., Lau, M. I., and Williams, M. B. (2002). Estimating individual discount rates in Denmark: A field experiment. *American Economic Review*, 92(5):1606–1617.
- Harvey, C. M. (1986). Value functions for infinite-period planning. *Management Science*, 32(9):1123–1139.
- Herrnstein, R. (1981). Self-control as response strength. In Bradshaw, C. M., Szabadi, E., and Lowe, C. F., editors, *Quantification of Steady-State Operant Behavior*, pages 3–20. Elsevier/North-Holland, New York.
- Horowitz, J. K. and Carson, R. T. (1990). Discounting statistical lives. *Journal of Risk and Uncertainty*, 3(4):403–413.
- Huber, J. and Zwerina, K. (1996). The importance of utility balance in efficient choice designs. *Journal of Marketing Research*, 33(3):307–317.
- Johannesson, M. and Johannsson, P.-O. (1997a). Quality of life and the WTP for an increased life expectancy at an advanced age. *Journal of Public Economics*, 65(2):219–228.
- Johannesson, M. and Johannsson, P.-O. (1997b). The value of life extension and the marginal rate of time preference: A pilot study. *Applied Economics Letters*, 4:53.
- Keller, L. R. and Strazzera, E. (2002). Examining predictive accuracy among discounting models. *Journal of Risk and Uncertainty*, 24(2):143–160.
- Kirby, K. N. and Marakovic, N. N. (1995). Modeling myopic decisions: Evidence for hyperbolic delay-discounting within subjects and amounts. *Organizational Behavior and Human Decision Processes*, 64(1):22–30.
- Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics*, 112(2):443–477.
- Lee, L.-F. (1999). Statistical inference with simulated likelihood functions. *Econometric Theory*, 15(3):337 – 360.
- Loewenstein, G. and Prelec, D. (1992). Anomalies in intertemporal choice: Evidence and an interpretation. *Quarterly Journal of Economics*, 107(2):573–597.
- Mathworks, T. (2006). Matlab.

- Mazur, J. E. (1987). An adjustment procedure for studying delayed reinforcement. In Commons, M. L., Mazur, J. E., Nevin, J. A., and Rachlin, H., editors, *Quantitative analysis of behaviour: The effect of delay and intervening events on reinforcement value*. Erlbaum, Hillsdale, NJ.
- Minnesota River Basin Data Center (2007). Minnesota river basin data center. available at <http://mrbdc.mnsu.edu/>. (Accessed: November 15, 2007).
- Phelps, E. S. and Pollak, R. A. (1968). On second-best national saving and game-equilibrium growth. *The Review of Economic Studies*, 35(2):185–199.
- Read, D. (2001). Is time-discounting hyperbolic or subadditive? *Journal of Risk and Uncertainty*, 23(1):5–32.
- Samuelson, P. A. (1937). A note on measurement of utility. *The Review of Economic Studies*, 4(2):155–161.
- Slonim, R., Carlson, J., and Bettinger, E. (2007). Possession and discounting behavior. *Economics Letters*, 97(3):215–221.
- Thaler, R. (1981). Some empirical evidence on dynamic inconsistency. *Economics Letters*, 8(3):201–207.
- Train, K. E. (2003). *Discrete Choice Methods with Simulation*. Cambridge University Press, New York, NY.
- United States Census Bureau (2007). American factfinder: 2007 american community survey 3-year estimates.
- University of Houston Center for Public Policy (2007). Demographic survey of Texas lottery players 2007. available at <http://www.uh.edu/hcpp/txlottery.pdf>. (Accessed: October 1, 2010).
- University of Houston Center for Public Policy (2008). Demographic survey of Texas lottery players 2008. available at <http://www.uh.edu/hcpp/txlottery2008.pdf>. (Accessed: October 1, 2010).
- van der Pol, M. and Cairns, J. (2001). Estimating time preferences for health using discrete choice experiments. *Social Science and Medicine*, 52(9):1459–1470.

Viscusi, W. K., Huber, J., and Bell, J. (2008). Estimating discount rates for environmental quality from utility-based choice experiments. *Journal of Risk and Uncertainty*, 37(2-3):199–220.

Warner, J. T. and Pleeter, S. (2001). The personal discount rate: Evidence from military downsizing programs. *American Economic Review*, 91(1):33–53.

Appendices

A Alternative Error-Difference Variances

Denote the alternative error-difference terms as $\tilde{\epsilon}_{ikj} = \epsilon_{ik} - \epsilon_{ij}$. Recalling that, for choice set s , $\epsilon_{ij} = \sum_{t=0}^{T_s} \psi_t \eta_{ijt}$, I have

$$\tilde{\epsilon}_{ikj} = \sum_{t=0}^{T_s} \psi_t \eta_{ikt} - \sum_{t=0}^{T_s} \psi_t \eta_{ijt}. \quad (19)$$

Then, note that $\tilde{\epsilon}_{ikj}$ is heteroskedastic because the number of terms in the summations is determined by the length of the intertemporal alternative. $\tilde{\epsilon}_{ikj}$ is a normal error term with mean zero and variance given by

$$V(\tilde{\epsilon}_{ikj}) = V\left(\sum_{t=0}^{T_s} \psi_t \eta_{ikt} - \sum_{t=0}^{T_s} \psi_t \eta_{ijt}\right). \quad (20)$$

$$= \psi_0^2 V(\eta_{ik0}) + \psi_1^2 V(\eta_{ik1}) + \dots + \psi_{T_s}^2 V(\eta_{ikT_s}) + \psi_0^2 V(\eta_{ij0}) + \psi_1^2 V(\eta_{ij1}) + \dots + \psi_{T_s}^2 V(\eta_{ijT_s}) \quad (21)$$

since the instantaneous errors are independent. With the assumption that η_{ijt} *i.i.d* $N(0, \sigma_\eta)$, this leads to

$$V(\tilde{\epsilon}_{ikj}) = \sum_{t=0}^{T_s} \psi_t^2 \sigma_\eta + \sum_{t=0}^{T_s} \psi_t^2 \sigma_\eta. \quad (22)$$

It is well known that a probit model needs to be normalized for scale so set $\sigma_\eta = 1$ and I have

$$V(\tilde{\epsilon}_{ikj}) = \sum_{t=0}^{T_s} \psi_t^2 + \sum_{t=0}^{T_s} \psi_t^2 = 2 * \sum_{t=0}^{T_s} \psi_t^2 \quad (23)$$

B Random Discount Factors Simulated Log Likelihood Equation

Here, I develop the simulated log likelihood equation for the random discount factors specification. For clarity, I present the exponential discounting case. All other discounting models are easily derived with a few substitutions. This section loosely follows the exposition of Train (2003).

Recall the probability of a single choice for the non-stochastic discounting parameters case, $P_{ij} = F\left(\frac{\sum_{t=0}^{T_s} \delta^t v_{ijt} - \sum_{t=0}^{T_s} \delta^t v_{ikt}}{\sqrt{2 * \sum_{t=0}^{T_s} \delta^{2t}}}\right)$. In the case of random discounting parameters, I focus on the sequence of choices by individual i . Denote the choice situation as h and a sequence of alternatives as $\mathbf{j} = \{j_1, \dots, j_H\}$. Then, conditional on δ , the probability that individual i makes a sequence of choices is the product over all h of the single choice probabilities. I have

$$\mathbf{P}_{ij}(\delta) = \prod_{h=1}^H F\left(\frac{\sum_{t=0}^{T_s, h} \delta^t v_{ijth} - \sum_{t=0}^{T_s, h} \delta^t v_{ikth}}{\sqrt{2 * \sum_{t=0}^{T_s, h} \delta^{2t}}}\right). \quad (24)$$

Since the δ are random, I integrate out over all values of δ to get the unconditional choice probability

$$P_{ij} = \int \mathbf{P}_{ij}(\delta) f(\delta) d\delta. \quad (25)$$

I draw R values of δ from $f(\delta)$ and denote them δ_r . The simulated choice probability is $\tilde{P}_{ij} = \frac{1}{R} \sum_{r=1}^R \mathbf{P}_{ij}(\delta_r)$. In this application, I set $R = 200$. Finally, I insert these simulated choice probabilities into the log-likelihood function to get the simulated log likelihood (SLL)

$$\text{SLL} = \sum_i \sum_{\mathbf{j}} y_{ij} \ln \tilde{P}_{ij}, \quad (26)$$

where $y_{ij} = 1$ if i chose sequence \mathbf{j} and zero otherwise.

C Random Coefficients (α and γ) Simulated Log Likelihood Equation

Recall the probability of a single choice for the non-stochastic discounting parameters case, $P_{ij} = F\left(\frac{\sum_{t=0}^{T_s} \psi_t v_{ijt} - \sum_{t=0}^{T_s} \psi_t v_{ikt}}{\sqrt{2 * \sum_{t=0}^{T_s} \psi_t^2}}\right)$. In the case of random benefit and cost

coefficients, $v_{ijt} = \alpha_i q_{ijt} + \gamma_i (Y_{it} - c_{ijt})$. Denote the vector for individual i containing both α_i and γ_i as θ_i . θ_i is fixed for an individual across choice occasions, but varies across individuals. Assume θ_i is normally distributed in the population with mean H and covariance W : $\theta_i \sim N(H, W)$. I focus on the sequence of choices by individual i . Denote the choice situation as h and a sequence of alternatives as $\mathbf{j} = \{j_1, \dots, j_H\}$. Then, conditional on θ , the probability that individual i makes a sequence of choices is the product over all h of the single choice probabilities. I have

$$\mathbf{P}_{ij}(\theta) = \prod_{h=1}^H F \left(\frac{\sum_{t=0}^{T_{s,h}} \psi_t v_{ijth} - \sum_{t=0}^{T_{s,h}} \psi_t v_{ikth}}{\sqrt{2 * \sum_{t=0}^{T_s} \psi_t^2}} \right). \quad (27)$$

Since the θ are random, I integrate out over all values of θ to get the unconditional choice probability

$$P_{ij} = \int \mathbf{P}_{ij}(\theta) f(\theta) d\theta. \quad (28)$$

I draw R values of θ from $f(\theta)$ and denote them θ_r . The simulated choice probability is $\tilde{P}_{ij} = \frac{1}{R} \sum_{r=1}^R \mathbf{P}_{ij}(\theta_r)$. In this application, I set $R = 200$. Finally, I insert these simulated choice probabilities into the log-likelihood function to get the simulated log likelihood (SLL)

$$\text{SLL} = \sum_i \sum_{\mathbf{j}} y_{ij} \ln \tilde{P}_{ij}, \quad (29)$$

where $y_{ij} = 1$ if i chose sequence \mathbf{j} and zero otherwise.

D Random Coefficients (Discount Factors and α and γ) Simulated Log Likelihood Equation

Here, I show the simulated log likelihood equation when discount factors as well as benefit and cost coefficients are assumed to vary among individuals. Again, in the interest of clarity, I present the exponential discounting case. Retaining the notation and distributional assumptions from the two preceding sections of the appendix, denote the vector for individual i containing both δ_i and θ_i as ζ_i . Conditional on ζ , the probability that individual i makes a sequence of choices is the product over all h of

the single choice probabilities. I have

$$\mathbf{P}_{ij}(\zeta) = \prod_{s=1}^H F \left(\frac{\sum_{t=0}^{T_{s,h}} \delta_i^t v_{ijth} - \sum_{t=0}^{T_{s,h}} \delta_i^t v_{ikth}}{\sqrt{2 * \sum_{t=0}^{T_{s,h}} \delta^{2t}}} \right) \quad (30)$$

with $v_{ijt} = \alpha_i q_{ijt} + \gamma_i (Y_{it} - c_{ijt})$.

Since the ζ are random, I integrate out over all values of ζ to get the unconditional choice probability

$$P_{ij} = \int \mathbf{P}_{ij}(\zeta) f(\zeta) d(\zeta). \quad (31)$$

I draw R values of ζ from $f(\zeta)$ and denote them ζ_r . The simulated choice probability is $\tilde{P}_{ij} = \frac{1}{R} \sum_{r=1}^R \mathbf{P}_{ij}(\zeta_r)$. In this application, I set $R = 200$. Finally, I insert these simulated choice probabilities into the log-likelihood function to get the simulated log likelihood (SLL)

$$\text{SLL} = \sum_i \sum_j y_{ij} \ln \tilde{P}_{ij}, \quad (32)$$

where $y_{ij} = 1$ if i chose sequence \mathbf{j} and zero otherwise.