

Published in final edited form as:

J Public Econ. 2012 June 1; 96(5-6): 465–473. doi:10.1016/j.jpubeco.2012.02.001.

Policy Effects in Hyperbolic vs. Exponential Models of Consumption and Retirement

Alan L. Gustman and

Department of Economics, Dartmouth College, Hanover, NH 03755-3514, and NBER,
Alan.L.Gustman@dartmouth.edu

Thomas L. Steinmeier

Department of Economics, Texas Tech University, Lubbock, TX 79409,
thomas.steinmeier@ttu.edu

Abstract

This paper constructs a structural retirement model with hyperbolic preferences and uses it to estimate the effect of several potential Social Security policy changes. Estimated effects of policies are compared using two models, one with hyperbolic preferences and one with standard exponential preferences. Sophisticated hyperbolic discounters may accumulate substantial amounts of wealth for retirement. We find it is frequently difficult to distinguish empirically between models with the two types of preferences on the basis of asset accumulation paths or consumption paths around the period of retirement. Simulations suggest that, despite the much higher initial time preference rate, individuals with hyperbolic preferences may actually value a real annuity more than individuals with exponential preferences who have accumulated roughly equal amounts of assets. This appears to be especially true for individuals with relatively high time preference rates or who have low assets for whatever reason. This affects the tradeoff between current benefits and future benefits on which many of the retirement incentives of the Social Security system rest.

Simulations involving increasing the early entitlement age and increasing the delayed retirement credit do not show a great deal of difference whether exponential or hyperbolic preferences are used, but simulations for eliminating the earnings test show a non-trivially greater effect when exponential preferences are used.

I. Introduction

The impending advance of the baby boom generation into their retirement years, and perhaps even more importantly, the decline in the birth rate, have created strains on the Social Security system (Goss, 2010). As a result, increasing attention is being paid to measures that might encourage workers to stay on the job longer, contribute more to the system, and hence relieve some of the strain¹. Some of the proposals which have been mentioned have been increases in the early entitlement age, increases in the normal retirement age (which amounts to a reduction in benefits), and elimination of the earnings

© 2012 Elsevier B.V. All rights reserved.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

¹The U.S. Senate's Special Committee on Aging (2010) discusses a number of options for Social Security reform.

test, among others. An important goal of current research in this field, including the present paper, is to gauge the likely success of such measures in increasing the retirement age.

This paper brings together two important strands of the recent literature on retirement and savings. One of these strands is the development of increasingly realistic structural models of retirement and saving. Empirical structural models of retirement began as relatively simple decisions as to the optimal time to retire.² Over time, advances in computational capacity have allowed these early models to evolve into far richer models involving more nuanced decision sets and more elements of uncertainty.³ The decision to retire now includes the possibility of partial retirement as well as the possibility of returning to work after a period of retirement.⁴ The stochastic environment includes not only mortality but also uncertainty as to the returns to assets and the degree to which an individual will find retirement enjoyable after the fact.⁵ Other authors introduce uncertainty in wages and unpredictable health and health care expenditures. The result has been that the structural models of retirement and saving have increasingly been able to encompass the major elements affecting the retirement decision.

A second strand of literature is the analysis of so-called “hyperbolic” preferences. These preferences have been introduced in recent years to reflect the fact that many individuals place a heavy weight on current consumption, but do not distinguish incremental years in the future quite so much. Proponents of hyperbolic preferences (Laibson, 1997) point to a wide range of phenomena which hyperbolic preferences help to explain, such as simultaneously having high-interest credit card debt and low-interest individual retirement accounts. The hyperbolic preferences used in this paper are actually what are called “quasi-hyperbolic” preferences, with a high discount rate between the current period and the next period and lower discount rates between successive future periods.

The literature on hyperbolic preferences focuses mainly on consumption and saving behavior. While saving for retirement may be an important element of this behavior, the actual retirement date is largely taken as fixed. On the other hand, the structural retirement models almost uniformly assume exponential preferences, where a uniform time preference rate, perhaps differing among individuals, is applied to all future periods.

The purpose of this paper is to construct a structural retirement model that can encompass either exponential or hyperbolic preferences. We estimate the model twice, once using exponential preferences and once using hyperbolic preferences. Using these estimates, we can examine the differences in retirement and consumption outcomes between the two models, and estimate and compare the effects of several potential policy changes that would be implied by the two models. Such comparisons can indicate the sensitivity of the estimated structural models to assumptions regarding the type of preferences.

The next section looks at several properties of hyperbolic models and examines the degree to which it is possible to differentiate hyperbolic preferences from exponential preferences in the data. Section III specifies in more detail a structural model which can encompass either exponential or hyperbolic preferences, and Section IV discusses the stochastic specification and estimation of the model. Section V analyzes several simulations with the estimated model to examine what difference the choice of preferences makes to the

²See, for instance, Gustman and Steinmeier (1986) as an early example.

³See, for instance, Rust and Phelan (1997), French and Jones (2004), French (2005), Scholz, Seshandri, and Khitatrakun (2004) and Van der Klaauw and Wolpin (2006).

⁴As an example, see Gustman and Steinmeier (2008b).

⁵See Gustman and Steinmeier (2002).

estimated effects of potential policy changes. Concluding observations are contained in the last section.

II. Properties of Hyperbolic Models

In this section we will consider some of the properties of models with hyperbolic preferences and compare them to models with exponential preferences. With exponential preferences, expected utility at any period t can be expressed as

$$E_t U = E_t \left[\sum_{i=t}^T \left(\frac{1}{1+\rho} \right)^{i-t} u(C_i) \right]$$

where $u(C_i)$ is the usual increasing but with diminishing marginal returns within-period utility function, and ρ is the time preference rate. The term $\frac{1}{1+\rho}$ is the discount rate. For hyperbolic preferences, expected utility is

$$E_t U = u(C_t) + \frac{1}{1+\rho_1} E_t \left[\sum_{i=t+1}^T \left(\frac{1}{1+\rho_2} \right)^{i-(t+1)} u(C_i) \right]$$

In the hyperbolic model, the time preference rate between t and $t+1$ is ρ_1 , and between any two adjacent time periods after $t+1$ the time preference rate is ρ_2 . In the presumed case where $\rho_1 > \rho_2$, the individual weighs the current period more heavily than future periods, but does not distinguish quite as much among future periods. Exponential preferences may be viewed as a subset of hyperbolic preferences with $\rho_1 = \rho_2$.

Theorists distinguish among two variants of the hyperbolic model. In the first variant, the individual assumes that whatever consumption stream he chooses at time t will be followed in successive periods. This means that he thinks that at time $t+1$, the time preference rate between $t+1$ and $t+2$ will still be ρ_2 , despite the fact that at time t , the time preference rate between t and $t+1$ is ρ_1 . Since this is unlikely to be the case, such an individual is commonly called “naïve.” The alternative case is that the individual knows that although at time t the time preference rate between $t+1$ and $t+2$ is ρ_2 , when the individual reaches time $t+1$ the time preference rate between these same two periods will be ρ_1 . That is, he knows that his preferences are inconsistent across time and takes this into account. The literature refers to such individuals as “sophisticated.”⁶

In this paper we will deal with individuals with sophisticated hyperbolic preferences. Naïve hyperbolic individuals would not be able to accumulate any significant amount of wealth, contrary to what is frequently observed. Though they would recognize the need to save for retirement, they would always find it advantageous to wait until next year to start saving. Sophisticated hyperbolic individuals recognize this problem, and they are able to save significant amounts despite having a high time preference for the current period over future periods.

⁶Hyperbolic discounting has been used to explain the failure to save and anomalies in saving behavior. An extension by Diamond and Koszegi (2003) jointly analyzes retirement and saving behavior. This work focuses on sophisticated hyperbolic discounters and the actions taken by early selves to influence the behavior of later selves. Diamond and Koszegi discuss cases where, to augment own welfare in later years, one might either undersave to discourage later selves from too early a retirement, or may subsidize an undesired early retirement through additional early savings, even though early selves favor later retirement.

Exponential vs. Hyperbolic Models for a Representative Individual

Consider next the difficulties when attempting to distinguish exponential vs. hyperbolic preferences among individuals in household surveys such as the Health and Retirement Study. Here we consider a multiperiod model which begins at age 25. The model incorporates survival tables in calculating future expected utilities, up to a maximum age of 100. This representative individual is assumed to have a steady income from earnings of \$25,000 (in 1992 dollars) from age 25 until he retires at age 62. The wife, who is two years younger, is assumed to have a steady income of \$15,000 from age 25 until age 62, at which point she also retires. The husband's PIA is assumed to be around \$10,000 and the wife's PIA is assumed to be around \$7,000. To keep things simple, neither spouse is assumed to have a pension. These numbers are fairly close to the median amounts for the original HRS cohort.

In Figure 1 we look at the asset trajectory of this individual. The solid line in Figure 1 details the levels of assets for the individual if he had hyperbolic preferences with a discount factor of 0.67 ($\rho_1 = 0.50$) for the first period and an additional discount factor of 0.98 ($\rho_2 = 0.02$) for each period beyond the first period. As has been previously noticed in the literature (Harris and Laibson, 2002), individuals with hyperbolic preferences may accumulate a substantial amount of wealth as long as those preferences are "sophisticated," meaning that in each time interval the individual realizes that his future preferences, and in particular the low discount rate for periods subsequent to the next period, will not be the same as his current preferences. The dashed line in the figure details the levels of assets for the individual if he had exponential preferences with a discount factor of 0.958 ($\rho = 0.041$), which would yield approximately the same amount of assets at age 62 as would the hyperbolic preferences.

For individuals in the age range 50 to 65, which is typical of the age range for which we observe individuals in the HRS, the graph suggests that in practice it will be difficult to distinguish between exponential and hyperbolic preferences, or to distinguish the two parameters of hyperbolic preferences. Between ages 50 and 65 the assets of an individual with exponential preferences grow by 63 percent, or at an average rate of 3.2 percent per year, while the assets of an individual with hyperbolic preferences grow by 39 percent, or at an average rate of 2.2 percent per year. Given the substantial reporting errors for the asset levels of individual respondents in the HRS, especially those who reported asset levels in brackets, and the uncertainty over the returns that a particular individual would experience because of variations in asset mixes and in idiosyncratic returns to particular assets, it does not seem prudent to infer the values of the two discount factors for the hyperbolic model, or whether these two discount factors are equal, from the growth rate of assets over the period typically observed for HRS respondents.

One might ask whether it is possible to distinguish between the two types of preferences on the basis of consumption. The short answer is that the consumption data probably provides even less of a means to determine exponential vs. hyperbolic preferences than does the asset data. Figure 2 plots the consumption amounts for the same individual as before using exponential vs. hyperbolic preferences. Throughout most of the decade when the individual is in his 50's, the consumption differences between exponential and hyperbolic preferences is no more than a couple of thousand dollars, an amount that is undoubtedly within the measurement error of any kind of household survey. In the early 60's, the consumption difference widens somewhat, to the four to five thousand dollar range. Even this represents only a 10 percent difference between the two consumption profiles, a difference that would still be quite difficult to measure reliably for individual households in the HRS.

The long and short of the story, then, is that it is probably difficult to tell exponential preferences from hyperbolic preferences for individual households, or to tell the degree of hyperbolic preferences for those who have hyperbolic preferences, on the basis of either the asset data or the consumption data in the HRS. However, it does seem reasonable to infer from the overall levels of those assets whether the discount rates are high or low, with individuals with low levels of assets probably having high discount factors and those with high levels of assets probably having lower discount factors.

The lifetime asset and consumption trajectories of Figures 1 and 2 use a utility function based approximately on log consumption. In the full model considered in the next section, utility that is approximately log consumption is required in order to insure that distribution of individuals who retire at various ages is similar for both individuals with low lifetime earnings and those with high lifetime earnings. In models with only consumption, however, a more general model of consumption can be considered. One commonly used is

$u(C_t) = \frac{1}{1-\alpha} C_t^{1-\alpha}$. If $\alpha = 1$, this model reduces to log consumption, but this value of α is frequently considered too low. When we look at the asset and consumption trajectories with higher values of α , however, the two trajectories (hyperbolic vs. exponential preferences) are even closer together than in Figures 1 and 2 (Gustman and Steinmeier, 2010). Thus, the closeness of the asset and consumption trajectories of hyperbolic vs. exponential preferences does not seem to be driven by a low value of α .

Compensating Variations

When analyzing potential changes in Social Security, one important issue is the tradeoff between current benefits and future benefits. Social Security before the normal retirement age is frequently characterized as being roughly actuarially fair in the sense that if a person waits to collect benefits, the present value of future increases in those benefits, evaluated at a commonly used interest rate, is approximately equal to the value of the benefits currently given up. From the perspective of retirement models, however, what counts is not so much whether the increases in future benefits are actuarially fair in an accounting sense, but whether the perceived utility value of the future benefits, discounted at the time preference rate and not by an interest rate, is equal to the utility value of the benefits currently foregone.

Table 1 presents information related to the magnitude of these tradeoffs. Specifically, it reports the amount of money that would have to be given to an individual at age 62 to make him just as well off, in utility terms, as he would be if he were given a \$100 real annuity which would last the longer of his lifetime or his spouse's lifetime. This amount is frequently called a compensating variation. The top part of the table applies to the same individual that we have considered earlier in this section, whose behavior is depicted in Figures 1 and 2. The middle column gives the compensating variations for the individual with exponential preferences ($\rho = 0.041$), while the right hand column gives the variations for the individual with hyperbolic preferences ($\rho_1 = 0.5$, $\rho_2 = 0.02$). As noted previously, these two individuals would have roughly the same level of assets at age 62.

Since the value of the compensating variation depends on the level of assets, one can ask what would be the variation if these two individuals had alternative levels of assets at age 62. For initial levels of assets above around \$25,000, the compensating variations for the hyperbolic individuals are higher, but not by much. At asset levels below \$10,000 (which are not likely, given the relatively low time preference rates for these individuals), the compensating variation of the hyperbolic individual is considerably higher.

The lower part of the table gives compensating variations for individuals with higher time preferences. In this part of the table, the middle column refers to an exponential consumer

with a time preference rate of 0.18, while the right column refers to a hyperbolic consumer with a first period time preference rate of 0.5 and subsequent rates of 0.1. Both individuals will have accumulated the same amount of assets (a little over \$20,000) at age 62. For the higher time preference individuals, the compensating variations are around \$300 higher if the individual has hyperbolic preferences than if he has exponential preferences, and this difference holds up at almost any level of initial assets at age 62.

The general conclusions of this exercise is that, for individuals who have achieved a given level of wealth, they are more likely to be willing to trade off current assets for a stream of future assets if they have hyperbolic preferences than if they have exponential preferences. This appears to be especially true for individuals whose overall time preference is relatively high or whose asset levels are relatively low for whatever reason.

III. A Hyperbolic Model of Consumption and Retirement

As in a standard life cycle model, in each period individuals maximize expected utility subject to an asset evolution constraint. Consumption and leisure in that period are the choice variables for that period. Stochastic variables include the returns to assets, mortality outcomes, and retirement preferences. Potential wages and health are treated as exogenous and non-stochastic. Preferences for retirement may change unexpectedly once a person has retired, creating a mechanism for reversals from states of greater to lesser work.

In the model, for each year i , the individual chooses consumption C_i and leisure L_i to maximize expected utility:

$$E_t U = \frac{1}{1-\alpha} C_t^{1-\alpha} + h_t L_t^\gamma + e^{-\rho_1} E_t \left[\sum_{i=t+1}^T \left\{ e^{-\rho_2[i-(t+1)]} \sum_{m=1}^3 s_{m,i} \left(\frac{1}{1-\alpha} C_{m,i}^{1-\alpha} + h_i L_{m,i}^\gamma \right) \right\} \right]$$

In this utility function, the individual has a time preference rate of ρ_1 between year t and year $t+1$ and a time preference rate of ρ_2 between all subsequent periods. The expectations of future values of C and L are formed on the assumption that the individual knows that in future periods he will behave in the same way. That is, in period $t+1$ the choices of C_{t+1} and L_{t+1} will be made on the basis of a time preference rate of ρ_1 between year $t+1$ and $t+2$ and a time preference rate of ρ_2 for all subsequent periods. This is the sophisticated hyperbolic assumption, which assumes that the individual realizes he will have the same self-control problem in future years, as opposed to the naïve hyperbolic assumption, which would assume that the individual thinks that he can adhere to this year's preferences in future years. We focus on the sophisticated assumption because a naïve hyperbolic individual would never be able to save any substantial sum for retirement, always figuring that he could start saving next year.

L takes on a value of 1 if the individual is retired, 0 if the individual is working, and $\frac{1}{2}$ if the individual is partially retired. h_t indicates the strength of the individual's preference for retirement, which may vary from one year to the next according to the specification below. The model is estimated for married men, and the income of the spouse is assumed to be exogenous and non-stochastic. The index m takes on three values indicating whether both members of the couple survive until the relevant year, only the husband survives, or only the spouse survives. $s_{m,t}$ is the conditional probability that the household will have the composition described by m in year t . T corresponds to the maximum age beyond which the household's survival probabilities are too small to matter, taken to be 100 in the current work.

The asset constraint is given by:

$$A_t = (1+r_t) A_{t-1} + W_t(1 - L_{m,t}) + E_{m,t} + B_{m,t} - C_{m,t}$$

A_t is the level of assets in year t , and r_t is the stochastic return on those assets in year t . Assets are assumed to start out at 0 at the beginning of the working life and are not permitted to be negative. W_t is the wage rate at time t , which will depend on whether the individual has stayed on his career job, has moved on to a partial retirement job, or has previously retired and gone back to full-time work. The career job is considered to be the job the individual holds until he fully or partially retires for the first time. The term $E_{m,t}$ is the income accruing to the spouse, including earnings and pensions. The spouse is assumed to have a retirement date unaffected by the husband's choices, and the term is taken to be zero in states where the spouse is no longer alive.

$B_{m,t}$ is the amount of the husband's pension and the household's Social Security benefits, both of which will be affected by the husband's retirement decisions. For defined benefit pensions, the benefit amount is determined by the retirement date and continues until death. For defined contribution plans, the contributions are put into an account and allowed to accrue subject to the same stochastic return as is applied to assets. The account is assumed to be made available to the husband when he retires from his career job. Household Social Security benefits are calculated according to the Social Security rules, depending on previous retirement decisions and the composition of the surviving household. Since most individuals claim benefits as soon as eligible (Coile et al., 2002; Gustman and Steinmeier, 2002), we do not try to model the acceptance decision here and instead assume that individuals will claim the benefits as soon as they can, consistent with the earnings test.⁷

Retirement preferences are reflected in the coefficient to the leisure term in the utility function and are characterized by:

$$h_t = e^{\beta X_t + \varepsilon_t}$$

The linear form βX_t has three terms: a constant, age, and health status. The coefficient of age is taken to be positive, so that retirement gradually becomes more desirable as the individual ages and finds work to be more difficult. The ε_t term in h_t reflects the relative preference for leisure. An individual starts out with a value of ε and keeps this value until he retires from the career job. Upon retirement, the individual may find that retirement is more or less fulfilling than anticipated, or perhaps the individual may find that he values consumption relatively more than he had thought. In any case, experience allows the value of ε to change after retirement, and the model reflects this by allowing the value of ε_t to vary after retirement, with values in successive years correlated with a correlation parameter ρ_ε . If the individual finds that retirement is substantially less fulfilling than anticipated, a return to work, albeit at a reduced wage as compared to the career job, may be the optimal decision.

L^γ is proportional to the utility value of leisure. It has a value of zero if the individual is working full time ($L = 0$) and a value of one if the individual is fully retired ($L = 1$). For an individual who is partially retired ($L = 1/2$), $(1/2)^\gamma$ should take on a value between $1/2$ and 1 if diminishing marginal utility of leisure is to be satisfied. Call this value V_p . If V_p is close to unity, full-time work is particularly onerous compared to partial retirement work, and most

⁷For a model that includes benefit claiming as an endogenous decision, see Gustman and Steinmeier (2008a).

people should go through a period of partial retirement. If, on the other hand, V_p is close to $\frac{1}{2}$, the marginal disutility of work is rising very slowly with additional work, and the likelihood of partial retirement diminishes.

The individual carries several state variables from one period to the next; these are variables which are consequences of past decisions and random events that have a bearing on the current decision. The state variables in this model are the level of assets A_t , whether or not the individual is still in the career job, the size of the defined contribution pension balance (if applicable), the value of defined benefit pension amounts (also if applicable), and the value of Social Security benefits. The pension and Social Security amounts come from matched records from the firm and the Social Security Administration and hence properly reflect the heterogeneous retirement incentives inherent in pensions and Social Security.

IV. Stochastic Specification and Estimation

The model is completed by specifying the stochastic and heterogeneous elements. There are three elements of heterogeneity in the model. The first is the value of the two time preference parameters ρ_1 and ρ_2 . Time preference is considered to be a fixed effect, meaning that its value is separately estimated for each observation in the sample. The estimated values of time preference are taken as those values for which the assets simulated by the model match the actual value of assets at some moment in time. In this estimation, the assets are matched as of 1992, which is the beginning of the sample period. The second element of heterogeneity of the model is the initial value of ϵ_t , which is the preference for leisure (and therefore retirement). We take this initial value to be a random effect whose value is drawn from a normal distribution with a mean of zero and a standard deviation of σ_ϵ .

The third element of heterogeneity is utility value of partial retirement V_p . This value is also taken to be a random effect whose initial value is drawn from a truncated exponential distribution $f(V_p) = ke^{\delta V_p}$, defined on the interval $\frac{1}{2}$ to 1. For a given δ , k is the value needed to make the distribution function integrate to unity over the interval. If δ is positive, values of V_p toward unity will be more common, and partial retirement will be higher, while if δ is negative, values of V_p toward $\frac{1}{2}$ will be more common. Since partial retirement becomes a greater percentage of total employment at later ages, we allow δ to be a function of age: $\delta = \delta_0 + \delta_1 \text{ Age}$. This allows the distribution to shift as the individual ages. However, the individual maintains his relative position within the distribution; it is this position within the distribution which is the random effect.

Mortality of the two spouses is considered to be stochastic and is based on life tables provided by the Social Security Administration. The real interest rate on assets is also considered to be stochastic and is based on a portfolio which is around half stocks and half money market funds. We take the real interest rate to be uncorrelated over time since the actual historical interest rates on these assets is approximately uncorrelated over time periods of a year, which is the time period used in the model estimation and simulation. This means that the model does not have to introduce an additional state variable reflecting the value of interest rates, which would increase the computational complexity.

There are eight estimated parameters of the model: The exponent of consumption α ; three elements of β corresponding to a constant, the coefficient of age, and the coefficient of health; two elements δ_0 and δ_1 describing the distribution of partial retirement utilities; the standard deviation of the leisure preference term σ_ϵ ; and the correlation of leisure preference after retirement ρ_ϵ . The method of estimation is the generalized method of simulated moments (MSM). In a nutshell, this method chooses the parameter values which allow the model to best match a specified set of moments. There are 43 moments used in the estimation, as follows: 13 moments measuring the fraction of individuals in full-time work

between ages 54 and 66; 5 moments to measure the fraction of individuals not in full-time retirement at ages 55, 58, 60, 62, and 65; 5 moments to measure the fraction of individuals in the lower third of lifetime earnings in full-time work at those ages; 5 moments for a similar measure of individuals in the upper third of lifetime earnings; 5 moments to measure the fraction of individuals in poor health in full-time work at those ages and an additional 5 moments for similar individuals not in full retirement; and a final 5 moments to measure the fraction of individuals who went from a state of greater retirement to more work between the pairs of survey dates.

The model is estimated for members of the original HRS cohort. It uses matched Social Security records and the matched pension records when available. Individuals working at least 30 hours per week and 1560 hours per year are counted as full-time. Individuals working at least 100 hours per year but no more than 25 hours per week or 1250 hours per year are counted as partially retired, and individuals not doing any work at all are counted as fully retired. Individuals who fall between full-time work and partial retirement or between partial retirement and full-retirement are classified on the basis of self reports. For instance, an individual who works 35 hours per week for 39 weeks (1365 hours per year) is counted as full-time if he self-reports that he is “not retired at all” and as partially retired otherwise. Wealth is measured in the 1992 survey, and an individual is counted as being in poor health after the first time he self-reports his health as being “fair” or “poor” for two successive surveys. Earnings are taken from the matched Social Security records if possible and otherwise imputed on the basis of whatever earnings information is available.

The final sample consists of 2,231 respondents for whom we can construct, at least approximately, the details of their earnings and income opportunities, and for whom the model seems appropriate. This is slightly less than half of the number of married men available in the HRS sample.

We estimate two versions of the model. The first version constrains ρ_1 to be equal to ρ_2 for each individual, effectively making the model an exponential model. In this model, the common value of ρ is taken to be that value for which the assets implied by the model equal observed assets of the individual in 1992. The second version allows the two values of ρ to be different. As indicated in an earlier section, the ability to distinguish these two values for a particular individual depends on small differences in the growth rate of assets or in the level of consumption, differences that are likely to be overwhelmed by measurement error in household surveys such as the HRS. Thus, rather small differences in wealth in surveys at either or both ends of the observation period are likely to result in large differences in the implied values of ρ_1 and ρ_2 . Rather than have the estimate values of ρ_1 and ρ_2 be so imprecisely estimated, we try an alternative approach, namely by assigning a particular value to one of the two values. Since much of the theoretical discussions use a first period discount rate of around $\frac{2}{3}$, we assume a value of ρ_1 ($\rho_1 = 0.5$) which generates this discount rate and allow ρ_2 to vary so as to make the assets calculated by the model equal to observed assets in 1992. In some cases a value of ρ_2 equal to 0 is insufficient to produce calculated assets equal to observed assets in 1992; in those cases we set ρ_2 equal to 0 and allow ρ_1 to adjust until the two asset levels are equal. In other cases the implied value of ρ_2 exceeds ρ_1 ; here we assume that the two values of ρ are equal at a value which equates the two asset levels.

Table 2 gives the results of the two estimations. The first column corresponds to exponential preferences, and the second column to hyperbolic preferences. All the coefficients, with the exception of the δ_{age} coefficient, are significant in the both models, and the two sets of coefficients are quite similar to one another. One of the central coefficients in the model, β_{age} , is 0.058 in one model and 0.056 in the other. This means that the value of leisure is

increasing by around 5.5% to 6 % per year. This coefficient determines the sensitivity of the retirement age to changes in economic incentives. The faster the value of leisure is increasing with age, the less room there is for economic incentives to make much of an impact on the retirement age. With the value of leisure increasing by a relatively small amount per year, changes in economic incentives to retire can play a meaningful role in determining retirement. The coefficient of the health variable indicates that poor health has an effect equivalent to being around 7 years older, and the standard deviation of ϵ suggests that a one standard deviation in leisure preferences has an effect equivalent to being around 5 years older. This indicates that differences in leisure preferences play a substantial role in the dispersion of retirement ages.

With MSM, the overall fit of the model is governed by the q-statistic, which is derived from the differences between the calculated moments of the model and the observed moments in the data. The null hypothesis is the model is correct. Under this null hypothesis, it can be shown that the q-statistic is distributed in a χ^2 distribution with degrees of freedom equal to the differences between the number of moments used in the estimation and the number of parameters estimated. Large values indicate a poor fit that is unlikely if the model is indeed specified correctly. Since there are 43 moments used in the estimation and there are 8 parameters, the appropriate χ^2 distribution has 35 degrees of freedom and has a 5 percent significance value of 49.80 and a 4% significance value of 50.93. Thus the q-statistic for the model with exponential preferences is approximately included in the interval for 5% significance, and the statistic for the model with hyperbolic preferences, though slightly higher, is approximately included in the interval for 4% significance.

The closeness of the q-statistics for the two models is an indication that the present-bias parameter, which is the difference between the first period discount rate and the rate for subsequent periods, is not well identified. Fang and Wang (2010) argue that in hyperbolic models, exclusion variables which affect the transition of the payoff-relevant states but do not enter the instantaneous payoff function are important to distinguish the present-bias parameter from the exponential discounting factor. The lack of such identifying variables may explain why the two models we have estimated fit the data approximately equally well. However, in this paper we are not primarily interested in whether the present-bias parameter is identified or what its precise value is, but whether its value makes any substantial difference in the labor market implications of the model with respect to potential policy changes.

V. Simulations

The Base Simulation

We begin with a simulation of the model under the actual set of circumstances that applied to the original HRS cohort. This includes a normal retirement age of 65, a delayed retirement credit that increasing with successive cohorts toward 8%, and a phase-out of the earnings test for individuals over the normal retirement age in 2001.

Figure 3 gives the percentage of the cohort that retired from full-time work between the ages of 55 and 67. Since individuals can go back to full-time work in the model, the percentages in Figure 3 are the difference in the percentage retired from full-time work at a particular age and the corresponding percentage at the previous age. These percentages are better thought of as pseudo-percentages of individuals who retired from full-time work within the year, and they are composed of a net flow of individuals who leave full-time work that year, minus the flow of individuals who had been fully or partially retired and within the year have returned to full-time work.

The percentages corresponding to the dashed lines come from the model with exponential preferences, and those corresponding to the solid lines come from the model with hyperbolic preferences. The two lines are relatively close, indicating that both models are able to reproduce the peak in retirement at age 62 fairly well. The peak at age 62 is not quite two percentage points higher when the exponential model is used, while the peak at age 65 is not quite two percentage points higher when the hyperbolic model is used.

Figure 4 gives the percentage of each cohort that is retired from full-time work. This contrasts to Figure 3, which gives the pseudo-percentage that retired in the past year. As might be expected from the estimated values of σ_e , roughly two thirds of the HRS cohorts retire between the ages of 55 and 67. The percentages of individuals retired from full-time work include only those individuals simulated to be retired from full-time work at that age and do not include individuals who had previously partially or fully retired and then returned to full-time work. At the earlier ages in the figure, the group of full-time workers is mostly comprised of individuals who have never previously retired, but by the late 60's the percentage of full-time workers who have been retired at some time in the past rises to almost a third for the model with exponential preferences and a fourth for the model with hyperbolic preferences. Overall, though, the dominant message of Figure 4 is that both models fairly closely track each other.

Figure 5 shows the percentage of individuals who are partially retired, by age. This percentage starts off at around 5 percent at age 55 and rises gradually to around 9 percent at age 61. At age 62, it jumps to around 16 percent and stay at that level for three of four years before a gradual decline. The jump at age 62 is almost certainly due to the imposition of the earnings test on Social Security benefits, which for individuals with high time preference rates appears to be similar to a tax that reduces effective compensation. Since the earnings test has a disregard amount before it applies, it hits full-time earnings relatively more than part-time earnings and thus provides an incentive to move from full-time work to part-time work. In the mid 50's age range, over three-quarters of the individuals entering partial retirement come from the ranks of the full-time workers. By the late 60's, the situation reverses, with between 50 and 60 percent of individuals entering partial retirement coming from those who had previously been fully retired. For the comparison of models with exponential preferences with those with hyperbolic preferences, however, the main point of this figure is that the behavior of the models with the two types of preferences is in fact very similar.

Over the course of their working careers, some 30% of individuals are simulated to return to full-time work at some time after a period of full or partial retirement. 15% are simulated to engage in part-time work after a period of full retirement. Since there is some overlap between the two groups, the percentage of individuals who undergo some type of movement from a state of more retirement to a state of more work, what we call "reversals," is around 40%. Thus, a substantial minority of workers at some point reverse the usual flow from work to retirement and instead increase their work effort for at least a while.

A major question, of course, is whether using a model with exponential preferences gives a different answer to the effect of policy changes than does a model with hyperbolic preferences. Three simulations with different policy changes are generated to answer this question. The results are in the following paragraphs. It should be emphasized that these results examine the effects of using exponential vs. hyperbolic models in the estimation, and not the effects of introducing hyperbolic preferences *per se*. Authors such as Fang and Silverman (2009) have examined the effects of introducing a substantial first period discount rate and found that, holding the later discount rate constant, the effects of policy on labor supply may be substantial. However, in the current paper the transition from an exponential

model to a hyperbolic one is associated with a lower estimated discount rate in the later years whose effects may offset the effect of the higher discount rate in the first year.

Raising the Early Entitlement Age to 64

The first set of simulations pertains to raising the early entitlement age to 64, leaving the other parameters of the Social Security system the same. In particular, the reduction rate for benefits for early retirement remains the same, so that an individual with a normal retirement age of 65 would receive a reduction of $6\frac{2}{3}\%$ if he were to retire one year earlier, at age 64. That is, a person retiring at 64 would receive exactly the same benefits; it is just that he would not have the opportunity to retire even earlier with less benefits.

Figure 6 shows the effects of this policy change on the percentage individuals retiring from full-time work by age. Note that this simulation does not show the effects of a sudden change in the early entitlement age; rather, it assumes that the individuals have been aware of this change from the beginning of their working lives. Hence, this figure should show the long-run effects of this change. The figure clearly indicates that most of the peak in retirement from full-time work that formerly occurred at age 62 will shift to age 64 if the early entitlement age is shifted to this age. The peak of retirement is about 2 percentage points higher for the model with exponential preferences than for the model with hyperbolic preferences, but this primarily reflects the fact that the same thing is true in the base simulation illustrated in Figure 3. The simulations for both the exponential model and the hyperbolic model suggest that the peak at age 62 has not moved in its entirety; there is still a minor peak there, probably reflecting the fact that some defined benefit pension plans at that time still used 62 as an age for either early or normal retirement. But the bulk of the peak does move with the shift in the early entitlement age, and whether the model is framed with exponential preferences or hyperbolic preferences does not seem to affect this result materially.

Figure 7 shows another aspect of the shift in the early entitlement age, namely the percentage of the cohort that is in partial retirement at any particular age. In the base results in Figure 5, the percentage in partial retirement grows slowly up until age 62, jumps considerably at that age, and stays at the higher level for three or four years before gradually tapering off. Figure 7 suggests that if the early entitlement age increases, the jump in partial retirement will also shift to the new early entitlement age. Before age 62, and at age 64 and beyond, the amount of partial retirement is about the same as the original base simulations. It is partial retirement ages 62 and 63 that are substantially affected if the early entitlement age is raised to 64. Again, this result does not appear to depend very much on whether the model being simulated is based on exponential preferences or hyperbolic preferences.

Raising the Delayed Retirement Credit to 8%

The next set of simulations analyzes the situation wherein the delayed retirement credit is raised uniformly to 8% for the entire sample. Over the course of a couple of decades the delayed retirement credit, which is the percent that benefits are raised for each additional year after full retirement age the individual goes before claiming them, has been raised from 3 percent to 8 percent in a series of $\frac{1}{2}$ percent increases. For the households in the HRS, which contain at least one member born between 1931 and 1941, this generally means that the delayed retirement credit varied between 5 percent and 7.5 percent. Thus, relative to the base simulation, this set of simulations represents an increase in the delayed retirement credit of between 0.5 percentage points and 3 percentage points.

Figure 8 indicates the effect of raising the delayed retirement credit to 8 percent. The vertical axis measures the difference between the percentage of individuals employed full-

time in the simulation with the 8 percent delayed retirement credit and the corresponding percentage in the base simulation. Since the delayed retirement credit only applies for delays in retirement after the full retirement age, which was age 65 for most individuals in the HRS sample, the effect on full-time work was minimal for ages 64 and earlier. At age 65, the simulation indicates that full-time work would increase by a little over 2 percentage points, with the effect declining at later ages. By age 65, however, only about 28 percent of individuals in the sample are working full-time in the base simulation, so an increase of 2 percentage points in the number of full-time workers would translate to an increase of around 7 percent in the number of full-time workers in this age group. Moreover, this increase would be caused by an average increase in the delayed retirement credit for this sample of a little under 2 percentage points, from a little over 6 percent to 8 percent. The effect of the full increase in the delayed retirement credit from the level of 3 percent per year to 8 percent per year would probably be around three times as great.

For present purposes, the salient point of Figure 8 is that the increases in full-time work effort are very close whether we use the model with exponential preferences or the model with hyperbolic preferences. The similarities in results between the two models extend to other measures not included in the figure. For instance, at age 65 the percentage of individuals who are partially retired declines by 0.4 percentage points in both models. At the same age, the percent of individuals who are fully retired declines by 1.8 percentage points if the model with exponential preferences is used and 1.9 percentage points if the model with hyperbolic preferences is used. In this particular set of simulations, there do not appear to be any significant differences in the results of the model using exponential preferences and the model using hyperbolic preferences.

Eliminating the Earnings Test

The final set of simulations explores the effects of eliminating the earnings test, which reduces Social Security benefits if recipients earn more than a few thousand dollars annually. Although future benefits are increased to make up for current benefits lost to the earnings test, for individuals with high time preference the future benefit increases are not sufficient to compensate for the current loss. On an actuarial basis, the increases are roughly actuarially fair for those under age 65, and for the HRS cohorts they are slightly less than fair for those over age 65. In the base simulation, the earnings test is eliminated for those over the normal retirement age in 2001, reflecting changes in the actual Social Security law.

Figure 9 compares the results from two simulations. The first is a simulation where the earnings test is uniformly applied to all individuals in all years, including years after 2001. The second eliminates the earnings test for all individuals in all years and effectively allows everyone to begin collecting benefits at the early entitlement age regardless of current earnings. The vertical axis on this figure measures the difference in the percentage of individuals working full-time in the two simulations. One set of simulations pertains to the model with exponential preferences, and the other to the model with hyperbolic preferences.

Unlike the previous two sets of simulations, the choice of model does seem to make a difference when analyzing the effects of eliminating the earnings test. Using the model with exponential preferences, eliminating the earnings test would increase full-time employment by over 4 percentage points in the 62 to 64 age range and by an additional point or two in the 65 to 69 age range. The extra effect for those over 65 is probably due to the fact that in that age range, the increases in future benefits to make up for lost current benefits is a bit less than actuarially fair. On the other hand, using the model with hyperbolic preferences there is only a percentage point or so gain in full-time employment for those in the 62 to 64 age range and a somewhat greater increase in the 65 to 69 age range.

These results are more or less what one might expect. The net perceived wage for an individual is $w^p = w - et + cv$, where w is the actual wage, et is the amount lost to the earnings test, and cv is the perceived compensating variation of the resulting future Social Security increases. Moving from a situation where the earnings test is in effect to one in which there is no earnings test ($et = 0$ and $cv = 0$) results in a perceived wage change of $et - cv$. Since et is the same whether we are considering exponential or hyperbolic preferences, the perceived wage change depends on the values of cv . For low time preference individuals who have generally high levels of wealth, the differences in cv between exponential and hyperbolic preferences is minimal, as indicated in Table 1. For higher time preference workers with generally lower levels of wealth, cv is lower for exponential individuals than for hyperbolic individuals. This means that the perceived wages increase more (or decrease less) for exponential than for hyperbolic individuals, and they increase their work effort more in response to the higher wages. After the normal retirement age, which is 65 for most of this sample, the earnings test results in smaller future increases in benefits and hence proportionately smaller values of cv . This shifts upward the perceived wage change from an elimination of the earnings test for both exponential and hyperbolic individuals and is the reason for the greater increase in work effort at these ages for both these groups.

VI. Conclusions

Individuals with hyperbolic time preferences have inconsistent time preferences, in the sense that their relative valuation of consumption in two periods depends on when they are doing the valuation. In particular, they may value current consumption highly relative to future consumption, but have much less strong preferences between consumption at two future dates. Contrary to what one might expect, such individuals may nevertheless be able to accumulate substantial amounts of wealth for retirement, at least as long as they realize that in the future, they will behave the same way as today, and that today's preferences will not match future preferences.

Simulations with two models, one using exponential preferences and one using hyperbolic preferences, suggest that it is frequently difficult to distinguish empirically between hyperbolic and exponential discounters, at least on the basis of asset accumulation paths or consumption paths around the period of retirement. The simulations also suggest that, despite the much higher initial time preference rate, individuals with hyperbolic preferences may actually value a real annuity more than individuals with exponential preferences who have accumulated roughly equal amounts of assets. This appears to be especially true for individuals with relatively high time preference rates or who have low assets for whatever reason. This affects the tradeoff between current benefits and future benefits on which many of the retirement incentives of the Social Security system rest.

The policy simulations have presented a mixed bag. The simulations involving increasing the early entitlement age and increasing the delayed retirement credit do not show a great deal of difference whether exponential or hyperbolic preferences are used, but the simulations for eliminating the earnings test show a non-trivially greater effect when exponential preferences are used. This latter result probably reflects the differential tradeoffs between current and future benefits for the two kinds of preferences.

Whether to use exponential preferences or hyperbolic preferences is an unsettled point in the literature. Proponents of hyperbolic preferences point to modes of behavior, such as simultaneously holding credit card debt and IRA's, which are difficult to explain otherwise. On the other hand, while there continues to be disagreement, some empirical studies conclude that in the absence of some kind of shock, older individuals do not spend down their wealth, which can be even more of a problem for hyperbolic behavior than it is for

exponential behavior (Poterba, Venti and Wise, 2010). We have certainly not resolved the issue of whether to use exponential preferences or hyperbolic preferences in this paper. Since the predictions are the same for most outcomes and policies, the current literature, which is based on the exponential model, would seem to be adequate for most purposes. However, the paper does suggest that in the absence of a consensus, it would probably be wise to investigate the effects of potential policy changes in both scenarios and be aware of instances in which the effects are sensitive to which set of preferences is assumed.

Research Highlights

- This paper constructs a structural retirement model with hyperbolic preferences.
- Estimated effects of policies are compared under hyperbolic and standard exponential preferences.
- We find it is frequently difficult to distinguish empirically between models with the two types of preferences on the basis of asset accumulation paths or consumption paths around the period of retirement.
- The simulations also suggest that, despite the much higher initial time preference rate, individuals with hyperbolic preferences may actually value a real annuity more than individuals with exponential preferences who have accumulated roughly equal amounts of assets.
- Simulations involving increasing the early entitlement age and increasing the delayed retirement credit do not show a great deal of difference whether exponential or hyperbolic preferences are used, but simulations for eliminating the earnings test show a non-trivially greater effect when exponential preferences are used.

Acknowledgments

This paper was supported by a grant from the National Institute on Aging, 5R01AG24337. All opinions are those of the authors, and not of the National Bureau of Economic Research nor of the National Institute on Aging.

References

- Coile, Courtney C.; Diamond, Peter; Gruber, Jonathan; Jouten, Alain. Delays in Claiming Social Security Benefits. *Journal of Public Economics*. 2002; 84(3):357–385.
- Diamond, Peter; Koszegi, Botond. Quasi-hyperbolic Discounting and Retirement. *Journal of Public Economics*. 2003; 87:1839–1872.
- Fang, Hanming; Silverman, Dan. Time-Inconsistency and Welfare Program Participation: Evidence from the NLSY. *International Economic Review*. 2009; 50(4):1043–1077.
- Fang, Hanming; Wang, Yang. Estimating Dynamic Discrete Choice Models with Hyperbolic Discounting, with an Application to Mammography Decisions. NBER Working Paper 16438. 2010 Oct.
- French, Eric. A The Effects of Health, Wealth, and Wages on Labor Supply and Retirement Behavior@. *Review of Economic Studies*. 2005; 72:395–427.
- French, Eric; Jones, John Bailey. A The Effects of Health Insurance and Self-Insurance on Retirement Behavior@. Chicago: Federal Reserve Bank of Chicago, Working Paper 2004–12; 2004.
- Goss, Stephen C. The Future Financial Status of the Social Security Program. *Social Security Bulletin*. 2010; 70(3):111–125. [PubMed: 20737861]
- Gustman, Alan L.; Steinmeier, Thomas L. A Structural Retirement Model. *Econometrica*. 1986; 54(3): 555–584.

- Gustman, Alan L.; Steinmeier, Thomas L. Retirement and the Stock Market Bubble. NBER Working Paper 9404. 2002 Dec.
- Gustman, Alan L.; Steinmeier, Thomas L. Projecting Behavioral Responses to the Next Generation of Retirement Policies. *Research in Labor Economics*. 2008a; 28:141–196.
- Gustman, Alan L.; Steinmeier, Thomas L. How Changes in Social Security Affect Recent Retirement Trends. *Research in Aging*. 2008b; 31(2):261–290.
- Gustman, Alan L.; Steinmeier, Thomas L. Policy Effects in Hyperbolic vs. Exponential Models of Consumption and Retirement. NBER Working Paper 16503. 2010 Oct.
- Harris, Christopher J.; Laibson, David. Hyperbolic Discounting and Consumption. In: Dewatripont, Mathias; Hansen, Lars Peter; Turnovsky, Stephen, editors. *Advances in Economics and Econometrics: Theory and Applications; Eighth World Congress; 2002*. p. 258-298.
- Laibson, David. Golden Eggs and Hyperbolic Discounting. *Quarterly Journal of Economics*. 1997; 112:443–477.
- Poterba, James M.; Venti, Steven F.; Wise, David A. Family Status Transitions, Latent Health and the Post-Retirement Evolution of Assets. NBER Working Paper 15789. 2010
- Rust, John; Phelan, Christopher. A How Social Security and Medicare Affect Retirement@. *Econometrica*. 1997; 65(4):781–831.
- Scholz, Seshadri; Khitatrakun. Are Americans Saving Optimally for Retirement. *Journal of Political Economy*. 2006; 114(4):607–643.
- U.S. Senate, Special Committee on Aging. Washington D.C: U.S. Government Printing Office; 2010. Social Security Modernization Options to Address Solvency and Benefit Adequacy.
- Van der Klaauw, Wilbert; Wolpin, KennethI. A Social Security and the Retirement and Savings Behavior of Low Income Households@. *Journal of Econometrics*. 2006; 145(1–2):21–42. [PubMed: 21566719]

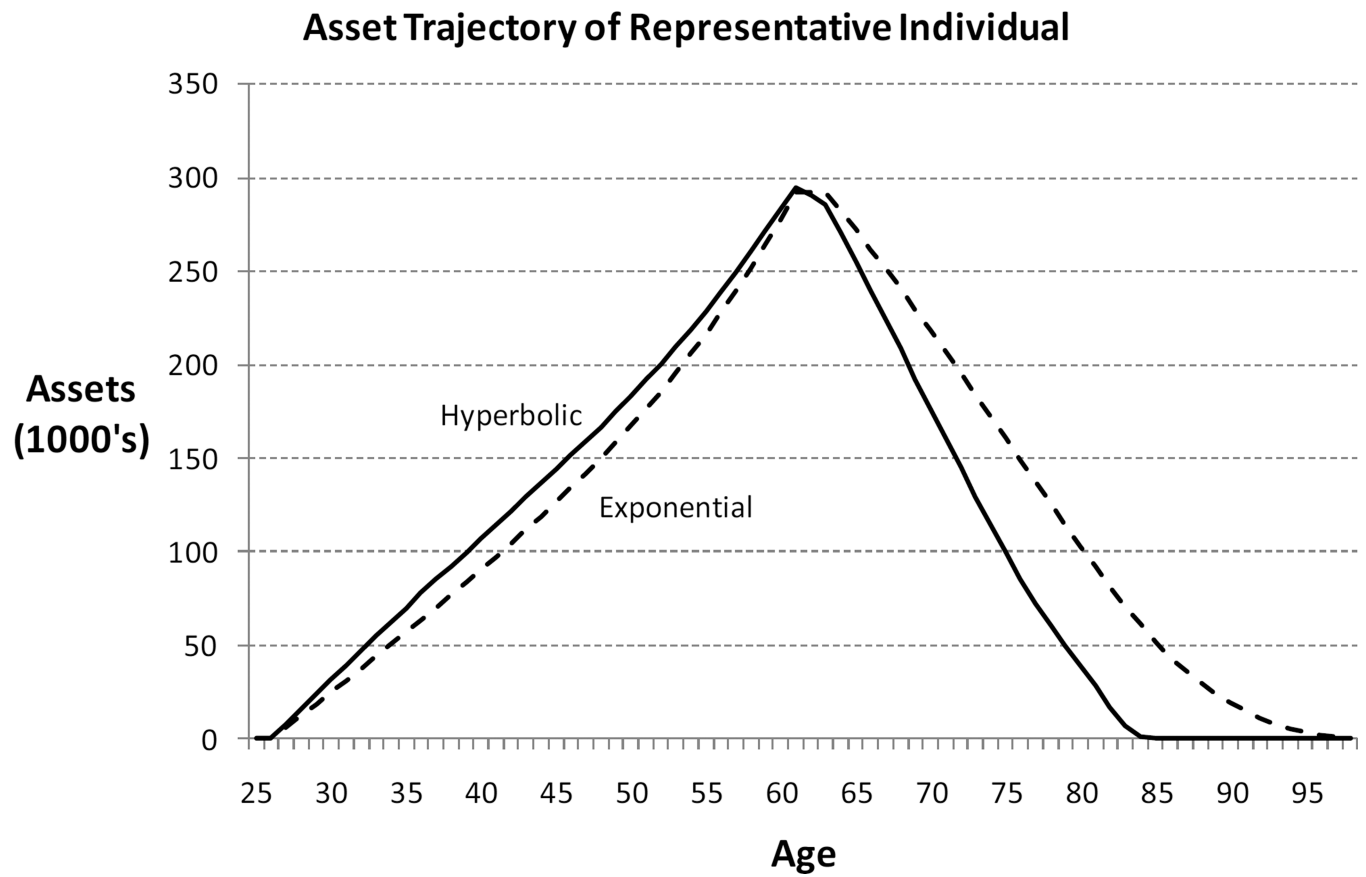


Figure 1.

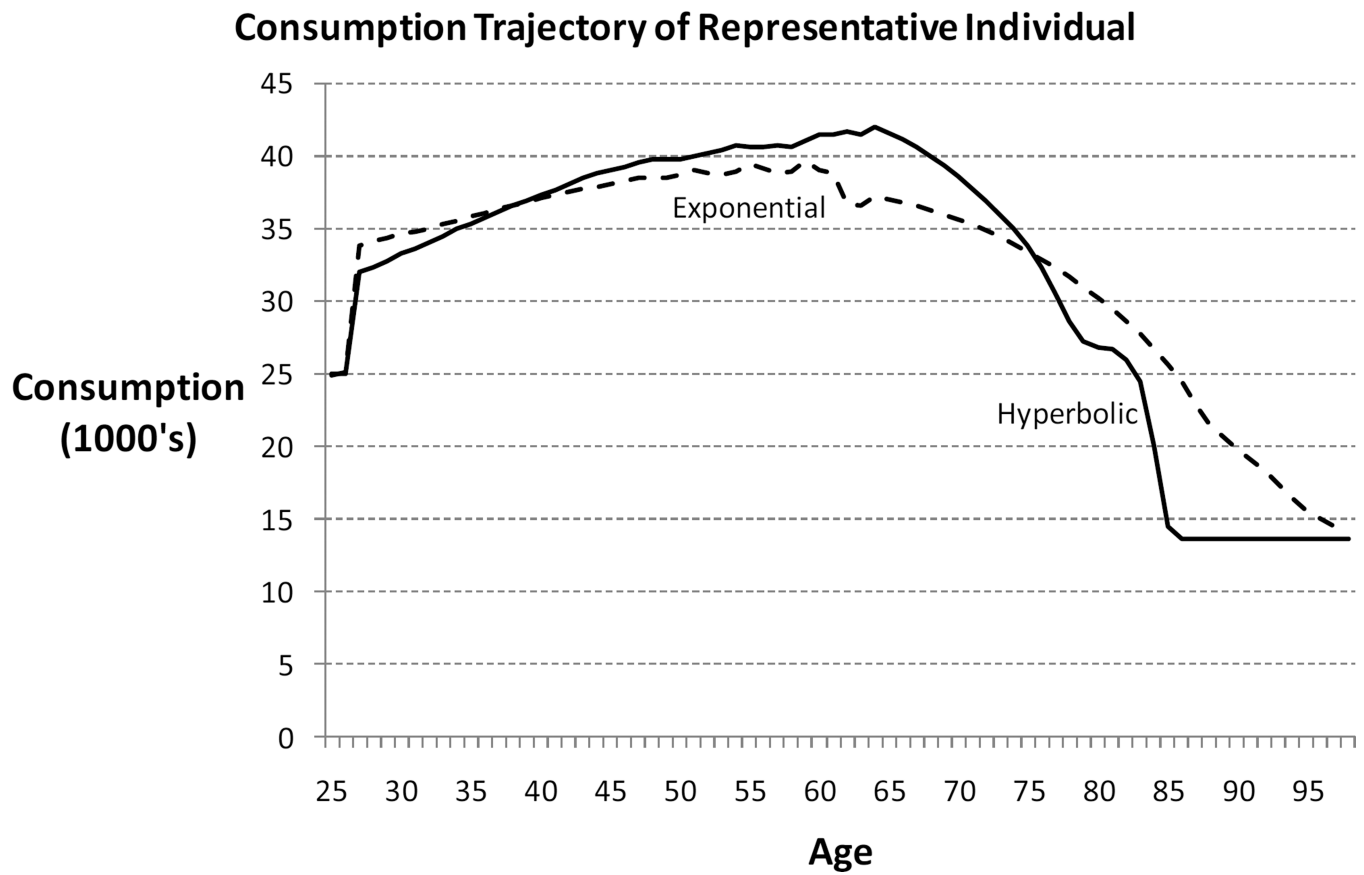


Figure 2.

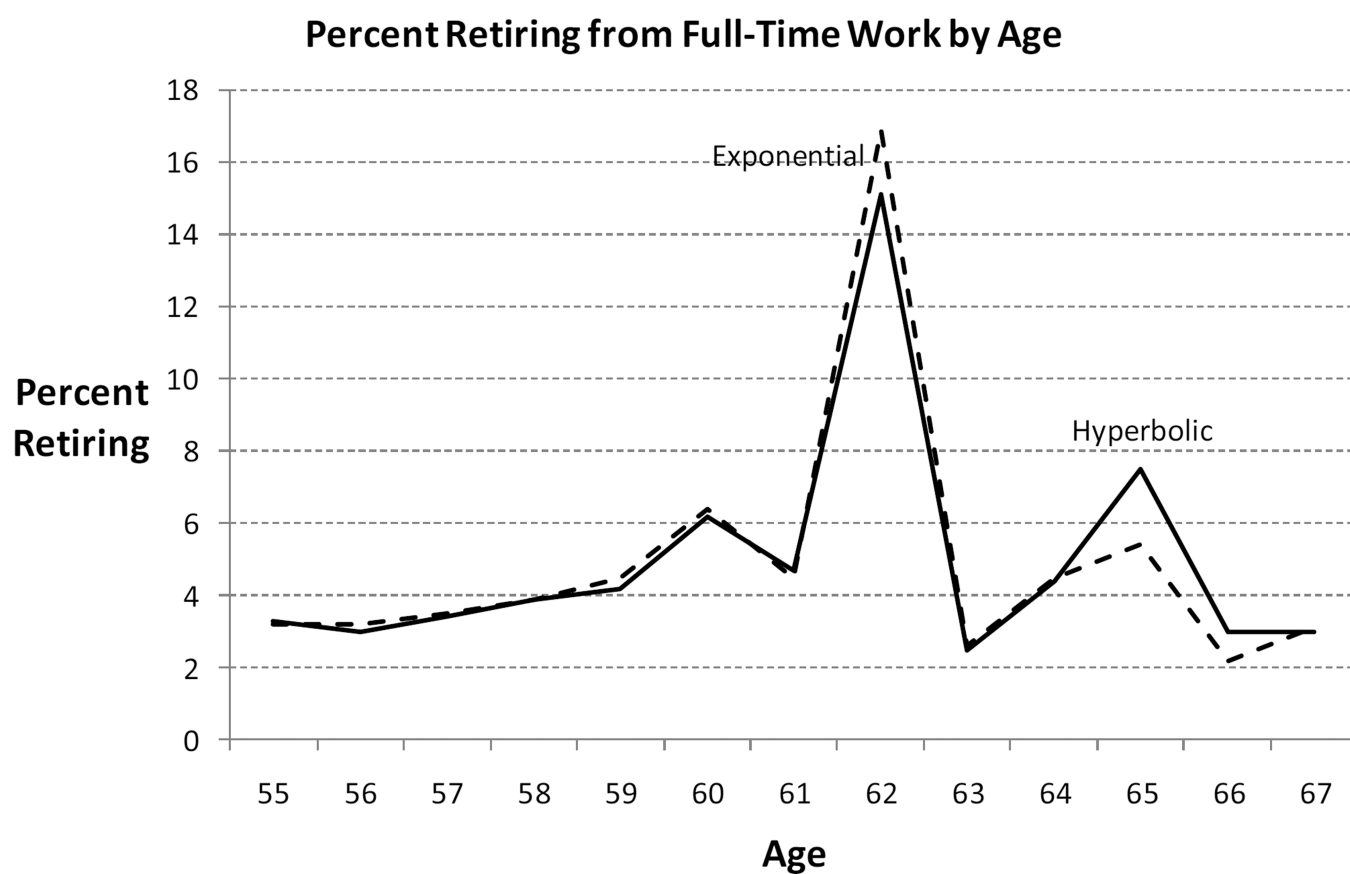


Figure 3.

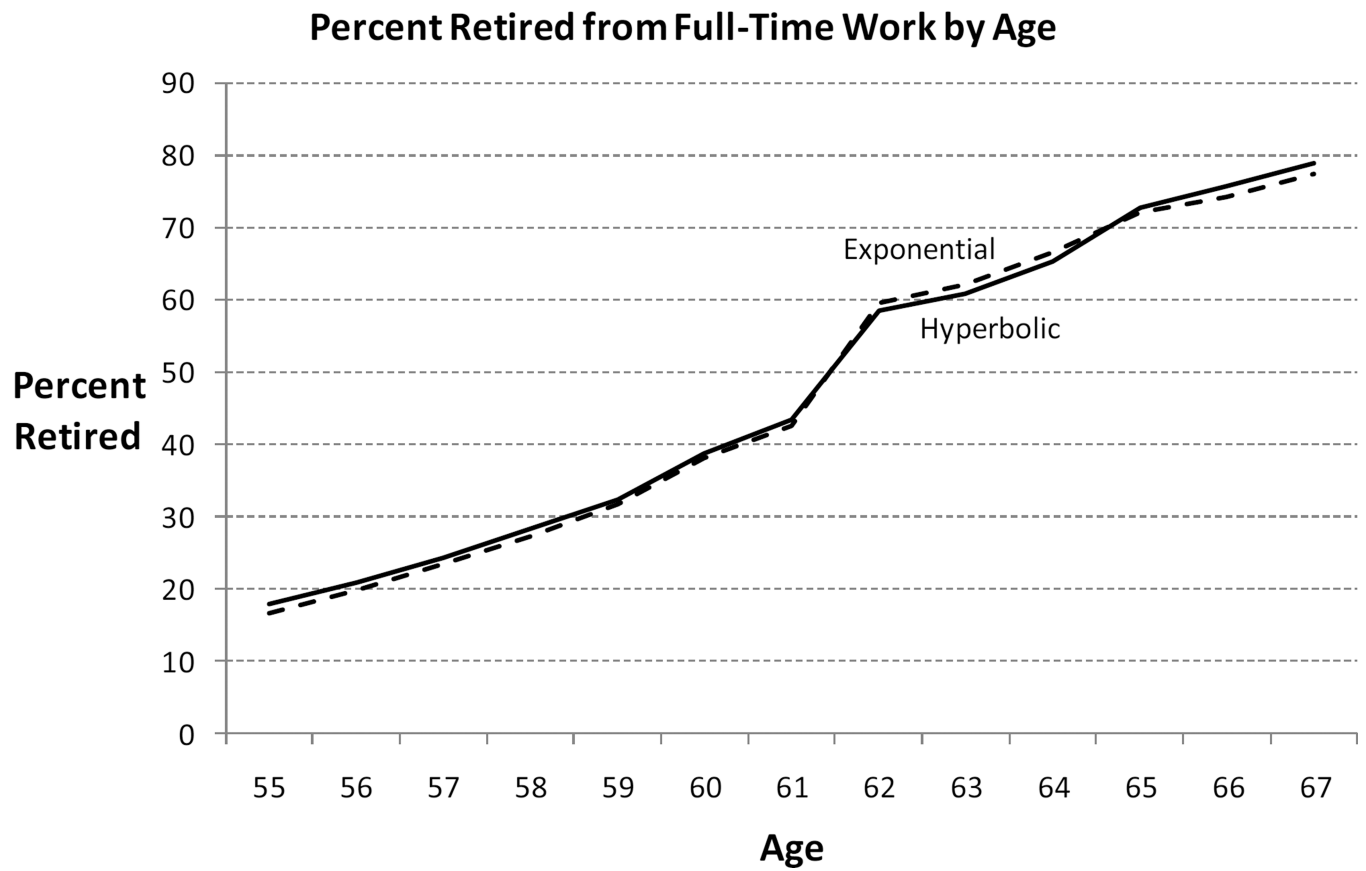


Figure 4.

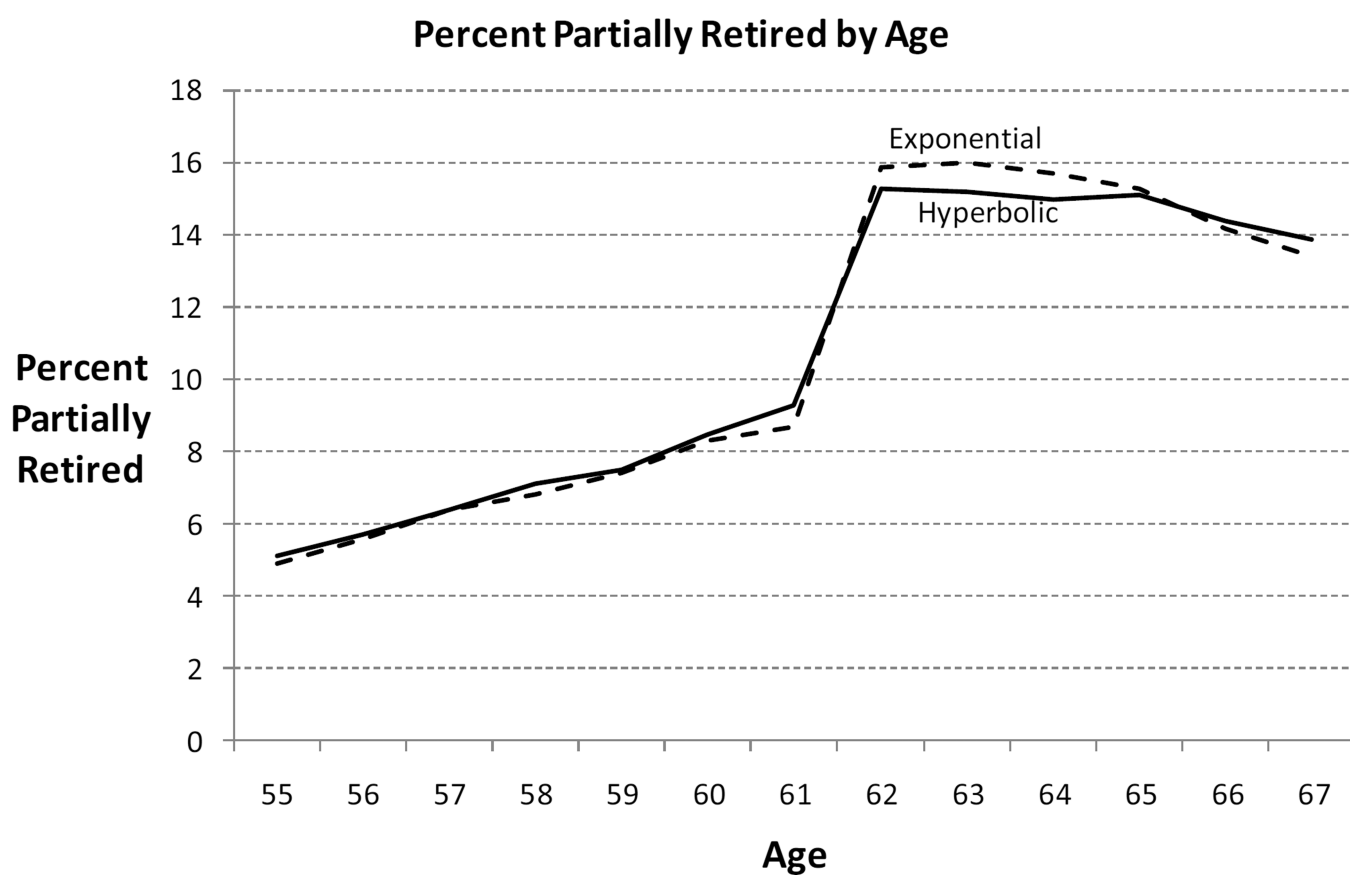


Figure 5.

Percent Retiring from Full-Time Work by Age Early Entitlement Age Raised to 64

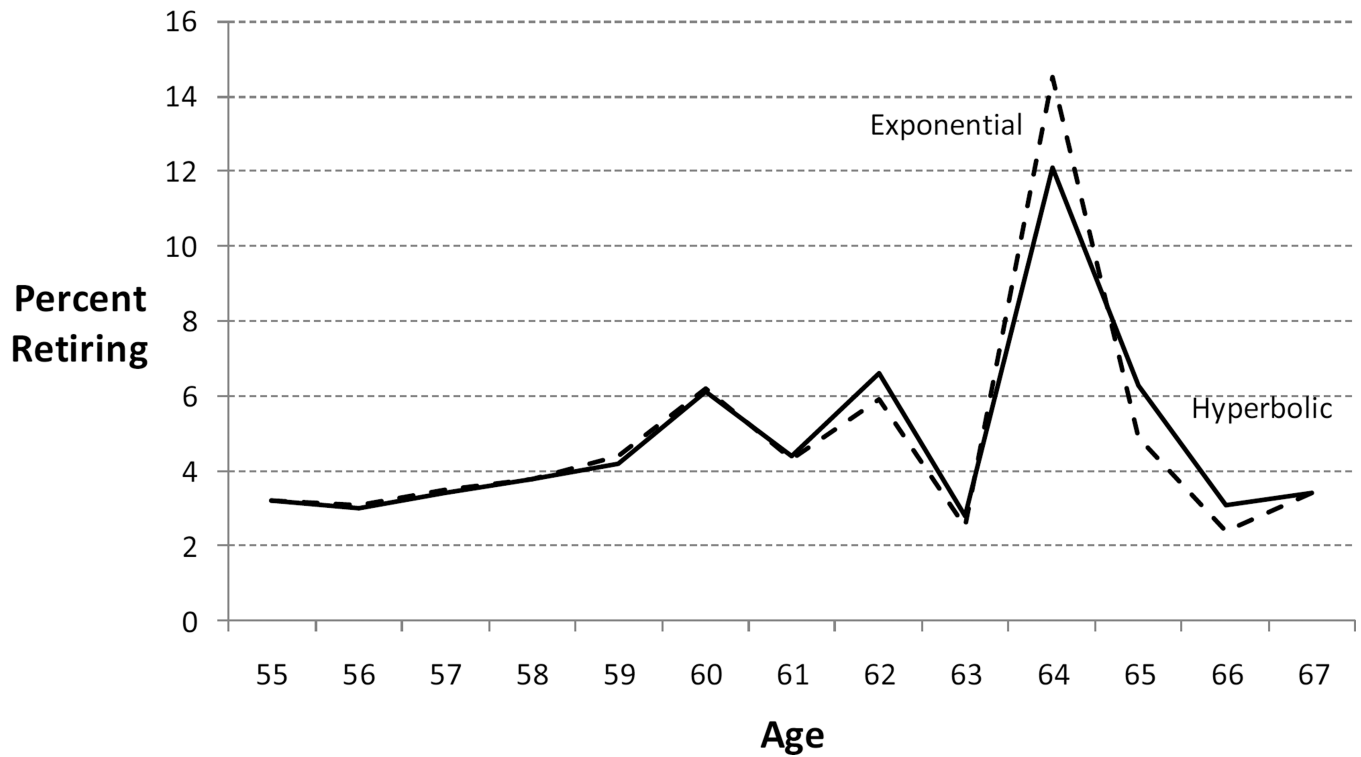


Figure 6.

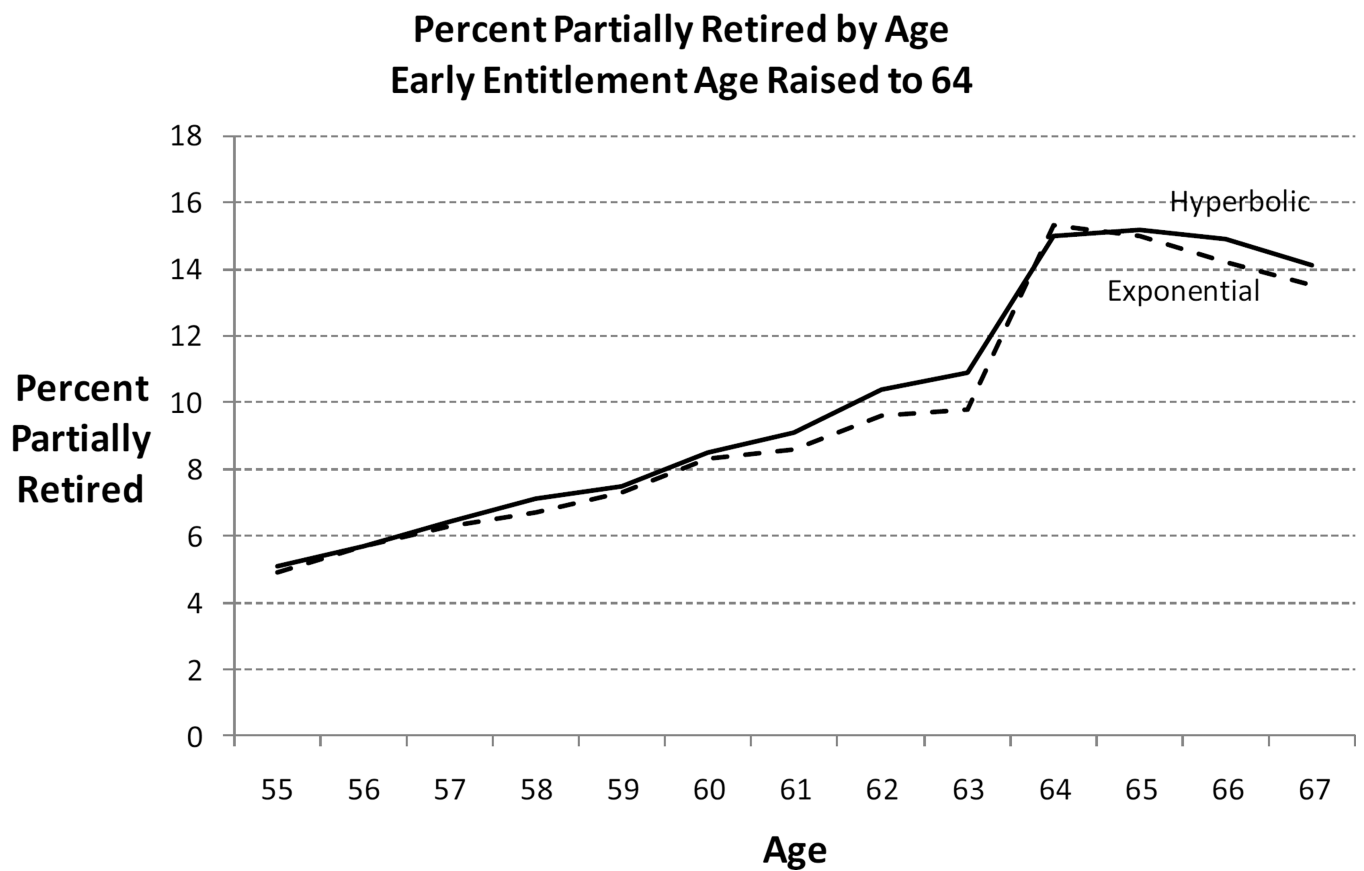


Figure 7.

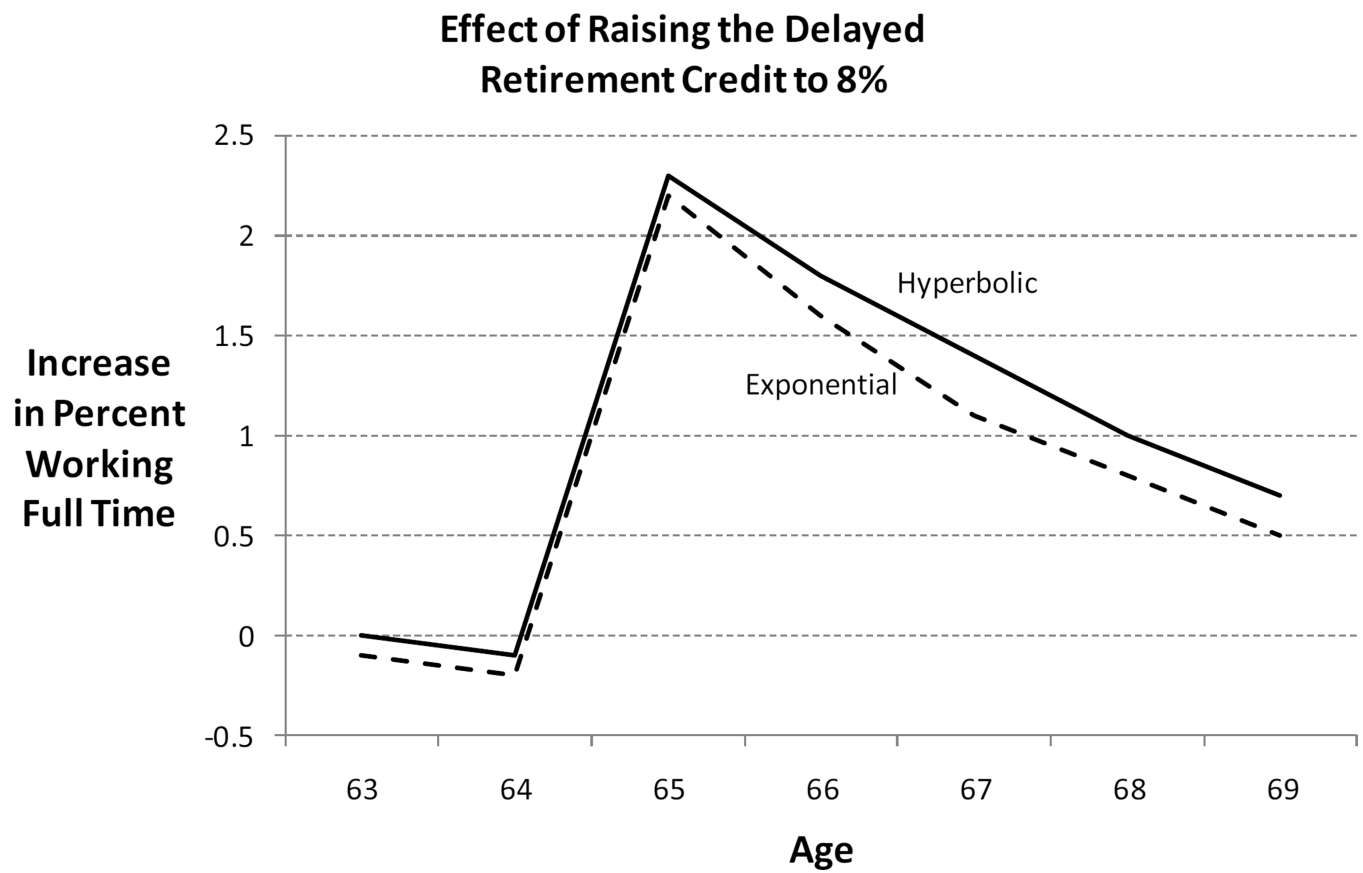


Figure 8.

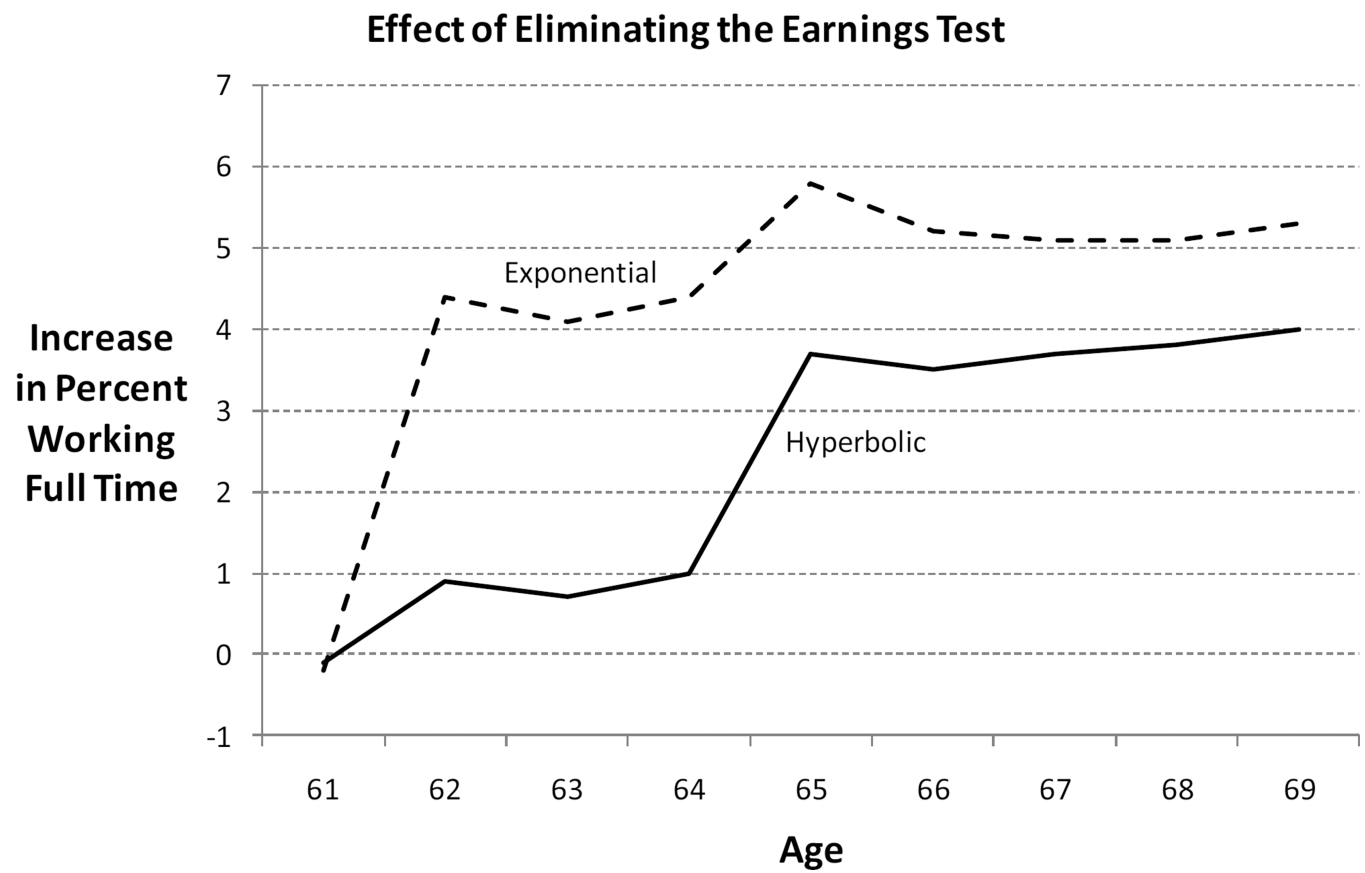


Figure 9.

Table 1

Value of a \$100 Real Annuity at age 62 Compensating Variation at Age 62 for a \$100 Real Immediate Annuity

Asset Level	Exponential	Hyperbolic
Low Time Preference		
	$\rho = 0.041$	$\rho_1 = 0.50$
		$\rho_2 = 0.02$
0	1,217	1,833
10,000	1,240	1,829
25,000	1,263	1,361
50,000	1,320	1,344
100,000	1,392	1,430
250,000	1,530	1,529
500,000	1,645	1,671
Higher Time Preference		
	$\rho = 0.18$	$\rho_1 = 0.50$
		$\rho_2 = 0.10$
0	791	1,005
10,000	846	1,137
25,000	909	1,199
50,000	981	1,240
100,000	1,078	1,318
250,000	1,225	1,448
500,000	1,352	1,657

Table 2

Estimates of the Exponential and Hyperbolic Models (absolute standard errors in parentheses)

Variable	Exponential Model	Hyperbolic Model
α	1.04 (0.02)	1.06 (0.02)
β_o	- 9.910 (0.020)	- 9.728 (0.018)
β_{age}	0.058 (0.008)	0.056 (0.007)
β_{health}	7.17 (0.95)	7.14 (0.75)
δ_o	- 4.46 (0.55)	- 4.24 (0.52)
δ_{age}	- 0.01 (0.10)	0.08 (0.11)
σ_e	5.38 (0.72)	5.50 (0.38)
ρ_e	0.72 (0.12)	0.71 (0.10)
q-statistic	50.02	51.16
Number of observations	2231	

Note: The coefficients β_{health} and σ_e are expressed as a multiple of β_{age} in order to facilitate their interpretation.