TIME INCONSISTENCY AND WELFARE PROGRAM PARTICIPATION: EVIDENCE FROM THE NLSY

By

Hanming Fang and Dan Silverman

July 2004

COWLES FOUNDATION DISCUSSION PAPER NO. 1465

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

YALE UNIVERSITY

Box 208281

New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

Time-inconsistency and Welfare Program Participation: Evidence from the NLSY*

Hanming Fang[†] Dan Silverman[‡]

First Version: April 2001
This Version: June 2004

Abstract

We empirically implement a dynamic structural model of labor supply and welfare program participation for agents with *potentially* time-inconsistent preferences. Using panel data on the choices of single women with children from the NLSY 1979, we provide estimates of the degree of time-inconsistency, and of its influence on the welfare take-up decision. With these estimates, we conduct counterfactual experiments to quantify the utility loss stemming from the inability to commit to future decisions, and the potential utility gains from commitment mechanisms such as welfare time limits and work requirements.

Keywords: Time Inconsistent Preferences, Welfare Reform, Labor Supply

JEL Classification Number: J22, I38.

*We are deeply indebted to Ken Wolpin for his advice and encouragement on this project. We also thank Steve Berry, Stefano Della Vigna, Zvi Eckstein, Michael Keane, David Laibson, Donghoon Lee, Ulrike Malmandier, Robert Miller, Daniele Paserman, Andrew Postlewaite, Matthew Rabin, Mark Rosenzweig, John Rust, Kent Smetters and participants at many seminars and conferences for helpful suggestions and discussions. We are responsible for the remaining errors and shortcomings.

†Department of Economics, Yale University, P.O. Box 208264, New Haven, CT 06520-8264. Email: hanming.fang@yale.edu

‡Department of Economics, University of Michigan, 611 Tappan Street, Ann Arbor, MI 48109-1220. Email: dansilv@umich.edu

1 Introduction

Economists studying choice over time typically assume decision makers are impatient and, traditionally, this impatience is modelled in a very particular way: agents discount future streams of utility or profits *exponentially* over time. Strotz (1956, p. 172) showed that exponential discounting is not just an analytically convenient assumption; in fact, without this assumption, intertemporal marginal rates of substitution will change as time passes, and preferences will be time-inconsistent.

A recent growing literature has built on the work of Strotz (1956) and others to explore the consequences of relaxing the standard assumption of time-consistent discounting. Drawing both on experimental research in psychology and on common intuition, economists have built models of quasi-hyperbolic discounting to capture the tendency of decision makers to seize short term rewards at the expense of long-term preferences.[1] With quasi-hyperbolic discounting, the relative value of utility received in period t versus period $t+1$ increases as period t draws nearer to the present. Thus agents have a bias towards near-term utility and immediate gratification. As a result, intrapersonal conflicts arise when earlier "selves" prefer a future sequence of trade-offs that their later selves will not find optimal and therefore will not make. Moreover, to the extent that they recognize this intrapersonal conflict, earlier selves have an incentive to commit their future selves to a preferred sequence of actions. Therefore time-inconsistent preferences may generate an incentive to exercise self-control that is absent from the standard framework.[2] The literature includes studies of the implications of self-control incentives for a variety of economic problems.[3] In some of these applications, researchers have used time-inconsistent preferences to parsimoniously explain a set of apparently common but irrational behaviors, and explore when and why agents might benefit from "commitment devices."

This paper is an empirical investigation of the influence of time-inconsistent preferences on work

[1]A body of experimental research, reviewed in Ainslie (1992) and several papers in Loewenstein and Elster (1992), indicates that hyperbolic time discounting may parsimoniously explain some basic features of the intertemporal decision making of both animal and human subjects that are inconsistent with exponential discounting. In the experiments using human subjects, standard decision models with exponential discounting are difficult to reconcile with commonly observed preference reversals: subjects choose the larger and later of two prizes when both are distant in time, but prefer the smaller but earlier one as both prizes draw nearer to the present (see Rubinstein 2003, however, for an alternative explanation of preference reversals).

[2]For thorough presentations of economic problems of self control, see Thaler (1991), Ainslie (1992) and O'Donoghue and Rabin (1999b), among others.

[3]For example, models of time-inconsistent preferences have been applied by Laibson (1997) and O'Donoghue and Rabin (1999a,b) to consumption and savings; by Barro (1999) to growth; by Gruber and Koszegi (2001) to smoking decisions; by Krusell, Kuruşçu, and Smith (2002) to optimal tax policy; by Carrillo and Mariotti (2000) to belief formation; by Della Vigna and Paserman (2001) to job search.

and welfare program participation decisions. Using panel data on the choices of single women with dependent children, we estimate a dynamic structural model of labor supply and welfare program participation that allows, but *does not assume*, time-inconsistent preferences. Our estimates, which also allow for a correlation between observable initial conditions and unobserved heterogeneity in skills and tastes, provide evidence of time-inconsistency in preferences. With reasonable precision, we estimate a present-bias factor considerably less than one, and reject a standard exponential discounting model at common confidence levels.

While the literature has used stylized models to demonstrate the potentially large behavioral effects of time-inconsistency in preferences, we emphasize that whether an estimate of time-inconsistency in preferences necessarily implies economically important behavioral consequences is inherently an empirical question. To illustrate this point, consider the extreme case in which agents' choices among discrete options are made very far from the margin;[4] then their choices are unlikely to change even when they are able to commit their future selves' behavior. Indeed, simulations of our estimated model indicate that, among women with very low human capital living in States with relatively high welfare benefits, self-control problems due to present-biased time preferences have relatively little impact on choices over work versus welfare (see Table 8). Among this group, commitment ability is predicted to generate more rather than less welfare participation, as these women no longer delay their entry onto the welfare rolls out of fear of welfare stigma. Thus, the model suggests that low-human capital women in high-benefit States participate in welfare simply because the return to welfare is quite high relative to that from low-wage work, and not because they cannot overcome self-control problems. In contrast, for women with more education residing in States with relatively low welfare benefits, simulations of the model indicate that self-control problems lead to substantial under-investment in human capital with respect to long-term preferences; that is, the model predicts that if they could commit themselves to future decisions, they would choose to work considerably more often.

We also emphasize that, even when the ability to commit may dramatically alter agents' *behavior* it does not necessarily imply large *utility gains*. To illustrate this point, consider another extreme case in which agents' choices are made very close to the margin; then their choices are likely to change when they can commit their future selves' behavior. But such changes in choices will have little utility consequence since they were initially close to the margin. Indeed, in our simulations we find that, even in low benefit States where agents' participation in work is increased substantially by the ability to commit, the utility gains are relatively modest. Our simulations also indicate that,

[4]That is, if an agent chooses alternative A over B, her utility from A is much larger than that from B; and vice versa.

while the changes in behavior resulting from imperfect commitment policies such as welfare time limits and workfare can be large, the expected utility of the welfare eligible is invariably reduced by these restrictions, though for some only slightly. Only when workfare generates meaningful human capital, and compensates lost home production through, for example, child care subsidies, does the model predict both meaningful gains in employment and expected utility (see Tables 9 and 10).

This paper makes two contributions to the literature on time-inconsistent preferences. First, by applying a model that allows quasi-hyperbolic preferences to the problem of labor supply and welfare program participation, we provide an economically significant setting for an evaluation of the importance of time-inconsistency. As a source of information about time-inconsistency, this context has the advantage that labor supply decisions are among the most consequential economic choices that individuals make; they largely define time use for working-age adults. One might expect, therefore, that choices about labor supply would be less subject to prolonged lapses in strict intertemporal rationality. Nevertheless, recent welfare reforms and anecdotal evidence indicate a commonly held view: the trade-off between the short-term costs of entering the labor force at a low wage relative to the welfare benefit, and the long term reward of higher wages from the accumulation of work experience, may generate problems of self-control. Our previous research has shown that this trade-off can, in theory, produce important observable differences in the behavior of time-consistent and time-inconsistent agents (Fang and Silverman, 2004a). Second, economists have so far *calibrated* models of time-inconsistent preferences to match important moments of aggregate data sets (for example, Laibson *et al.* 1998). In this paper, we take a different method to estimate the structural parameters of the model, including the present-bias parameter, from a single panel data set.[5] Two recent papers also used field data to structurally estimate discount factors. Paserman (2003) estimates a structural job search model using data on unemployment spells and accepted wages from the NLSY. Laibson, Repetto and Tobacman (2004) estimate a structural model of consumption and saving. They calibrate some of their model parameters and estimate the discount factors using the method of simulated moments.

This paper is also related to a small literature on labor supply and welfare participation that structurally implements models of dynamic decision-making. Miller and Sanders (1997) estimates a dynamic discrete choice model in which women decide monthly whether to work or receive welfare. In an effort to explain both the low welfare take-up rate among eligible families and the persistence of welfare choices among the families who do enroll, Miller and Sanders incorporate wage growth

[5]Some researchers have tested the reduced form implications of hyperbolic discounting. For example, Della Vigna and Paserman (2001) consider the influence of self control problems on job search; and Della Vigna and Malmandier (2002) find evidence supporting time-inconsistent preferences in data on health club contracts and usage.

through work experience, and preferences that adapt both to labor supply and to welfare experience. As in our paper, fertility and marriage are exogenous. Swann (1996) adds marriage to the choice set, and looks at women's decisions annually. In a working paper, Keane and Wolpin (2000) endogenize education, employment, fertility and marriage decisions. Each of these prior papers assumes exponential discounting preferences. Our paper contributes to this literature and to the welfare reform debate with, to our knowledge, the first empirical examination of the importance of time-inconsistency for the welfare take-up decision. We are also able to use the structural estimates of the model to calculate the value of commitment to potential welfare recipients and predict the behavioral response and utility consequences of welfare reform measures such as time limits and work requirement, thus quantifying the influence of time-inconsistency.

The remainder of the paper is structured as follows: Section 2 presents our model and describes the intrapersonal game played by the decision maker and describes the numerical method for obtaining the game's solution. Section 3 presents the estimation strategy and discusses identification issues. Section 4 describes the data and variable definitions. Section 5 presents the estimation results and associated simulations. Section 6 provides estimates of both the behavioral consequences and utility value of an ability to commit, and of various policy changes such as time limits and workfare. Section 7 concludes.

2 The Model

We consider a discrete time model of work-welfare decisions by a single parent (interchangeably, an agent). Each agent has a finite decision horizon starting from her age at the birth of her first child, a_0, and ending at age A.[6] At each age $a \in \{a_0, ..., A\}$, the agent must choose from a choice set D that includes three mutually exclusive and exhaustive alternatives: receive welfare, work in the labor market, or stay at home without work or welfare. The alternatives of welfare, work and home are respectively referred to as alternative $0, 1$ and 2, thus $D = \{0, 1, 2\}$.[7] The agent's decision in period a is denoted by $d_a \in D$.

The return from choosing alternative d for an agent at age a represents all of the current-period benefits and costs associated with the choice and it is denoted by $R_a(d; s_a, \epsilon_{da})$ where s_a is a set of payoff-relevant state variables at age a to be detailed below and ϵ_{da} is a random shock to the value

[6] Agents will obtain a continuation value in the terminal period A as a function of her payoff-relevant state variables.

[7] In reality, it is possible that an agent chooses more than one actions in any period, and there are also distinctions between part- and full-time work. For example, Edin and Lein (1997) report that, in their study of 379 low-income single mothers, many welfare recipients both work in the (unofficial) labor market and rely on family and neighborhood resources.

of alternative d at age a. We parameterize $R_a(d; s_a, \epsilon_{da})$ as follows.

Welfare. An agent's age-a return to welfare $R_a(0; s_a, \epsilon_{0a})$ depends on her State of residence j; the number of her children in period a, denoted by n_a; her age-$(a-1)$ choice d_{a-1}. In the absence of a time limit, $R_a(0; s_a, \epsilon_{0a})$ is specified as follows:[8]

$$R_a(0; s_a, \epsilon_{0a}) = e(n_a) + G_j(n_a) - \phi(d_{a-1}) + \epsilon_{0a}, \tag{1}$$

where $e(n_a)$ is the monetary value of her home production skills or leisure as a function of the number of her children; $G_j(n_a)$ is the monetary value of the cash and food welfare benefits in State j as a function of the number of her children; $\phi(d_{a-1})$ is the net stigma associated with welfare participation denominated in dollars; and ϵ_{0a} is an idiosyncratic, choice-specific shock.

The value of home production skills (leisure) is allowed to depend on the number of children to capture the additional demands or rewards of having more children. We assume a quadratic function for $e(n_a)$:

$$e(n_a) = e_0 + e_1 n_a + e_2 n_a^2. \tag{2}$$

The welfare benefits schedule $G_j(n_a)$ is assumed to be an affine function of the number of the agent's children:[9,10]

$$G_j(n_a) = \theta_{j0} + \theta_{j1} n_a. \tag{3}$$

The welfare benefit schedule $G_j(n_a)$ is estimated separately for each State. Finally, the net welfare stigma $\phi(d_{a-1})$ is specified as

$$\phi(d_{a-1}) = \begin{cases} 0, & \text{if } d_{a-1} = 0 \\ \phi, & \text{otherwise.} \end{cases} \tag{4}$$

In words, we assume stigma lasts for only one period after switching into welfare from some other choice.[11] The specification of welfare stigma in (4) is most natural if we interpret the stigma as the psychic and administrative costs associated with welfare take-up. If we take a more general interpretation of stigma, then (4) imposes a particular form of stigma decay with continued participation.

[8] To decrease the dimension of the state variables in our empirical estimation, we select samples of women who have children below age 18, thus are eligible for welfare, during all periods that we analyze.

[9] The actual welfare benefits schedule deviates from a linear function approximation by a few dollars at most. However, we abstract from asset and income restriction on welfare eligiblity.

[10] We assume residential location at age a_0 is given exogenously and remains unchanged through age A. Indeed, 85% of the sample described below continued to reside in the State in which they lived at the birth of their first child.

[11] The net stigma parameter $\phi(d_{a-1})$ has, since Moffitt (1983), become standard in empirical studies of welfare participation. Its primary function is to help explain the number of welfare-eligible adults who remain at home without work or welfare.

In the presence of welfare time limits, it will be necessary to keep track of the cumulative number of periods an agent has received welfare prior to age a. This is denoted by κ_a. Given a lifetime limit of L periods, the return to welfare is then

$$R_a\left(0; s_a, \epsilon_{0a}\right) = \begin{cases} e\left(n_a\right) + G_j\left(n_a\right) - \phi\left(d_{a-1}\right) + \epsilon_{0a}, & \text{if } \kappa_a < L; \\ -\infty, & \text{otherwise.} \end{cases}$$

Work. An agent's age-a return from work $R_a(1; s_a, \epsilon_{1a})$ is her wage. Following the standard theory of human capital, we model this wage as the product of a (constant) rental price of human capital, r, and the quantity of skill units held by the individual $h_a\left(s_a, \epsilon_{1a}\right)$:

$$R_a(1; s_a, \epsilon_{1a}) = r h_a\left(s_a, \epsilon_{1a}\right).$$

An agent's age-a quantity of skill units when the state variable is $s_a, h_a\left(s_a, \epsilon_{1a}\right)$, is given by

$$h_a\left(s_a, \epsilon_{1a}\right) = \exp\left[h_0 + \alpha_1 g_0 + \alpha_2 x_a + \alpha_3 x_a^2 + \alpha_4 \mathtt{I}\left(x_a > 0\right) + \alpha_5 \mathtt{I}\left(d_{a-1} \neq 1\right) + \epsilon_{1a}\right], \tag{5}$$

where h_0 is the agent's (unobserved) skill endowment at the birth of her first child; g_0 is her completed years of schooling at the birth of her first child;[12] x_a is her total work experience prior to age-a; and d_{a-1} is her choice in the previous period; and ϵ_{1a} is the age-a skill shock. In specification (5), $\mathtt{I}\left(\cdot\right)$ is an indicator function equal to one if the expression in parentheses is true. Thus $\alpha_4 \mathtt{I}\left(x_a > 0\right)$ takes value α_4 if the agent acquired any experience before age a, captures a persistent first-year experience effect. The term $\alpha_5 \mathtt{I}\left(d_{a-1} \neq 1\right)$ takes value α_5 (which is presumably negative) whenever the agent was not working in the previous period. Thus the parameter α_5 represents the one-time depreciation of human capital that occurs whenever the agent leaves work to choose welfare or home. Note that the functional form (5) implies that only $\left(\ln r + h_0\right)$, but not $\ln r$ and h_0 separately, can be identified.

Home. An agent's current-period return from staying home without work or welfare $R_a\left(2; s_a, \epsilon_{2a}\right)$ is specified as follows:

$$R_a\left(2; s_a, \epsilon_{2a}\right) = e\left(n_a\right) + \eta \mathtt{I}\left(d_{a-1} = 2\right) + \epsilon_{2a}$$

where $e\left(n_a\right)$ is the same monetary value of home production as in (2); η captures the possible decay or appreciation of the value of home production when a woman stays home without welfare; and ϵ_{2a} is a choice-specific shock.

[12] We assume the agent's level of schooling remains unchanged after the birth of her first child, thus do not model the schooling choice. In fact, 34% of our sample goes on to acquire additional schooling after the birth of their first child. Of this fraction, approximately half acquires less than one additional year of schooling.

We assume that the choice-specific shocks $\boldsymbol{\epsilon}_a = (\epsilon_{0a}, \epsilon_{1a}, \epsilon_{2a})$ are distributed according to a joint normal distribution $\mathrm{N}(0, \boldsymbol{\Omega})$, and they are serially uncorrelated. Now we describe the payoff-relevant state variables s_a in greater detail.

We analyze a single woman's work/welfare/home decisions from the age when she was first surveyed in the NLSY or when she gave birth to her first child, whichever occurred later. When an agent first enters our analysis at age a_0, we observe a set of initial conditions, including her State of residence j, her years of completed schooling g_0, her prior work experience x_0, and her decision in the period prior to the birth of her first child d_{a_0-1}. As we mentioned earlier, we assume that an agent's State of residence remains unchanged during the course of the data, and she does not further attend school. Thus (j, g_0) are the agent's payoff-relevant state variables that are constant over time. We do not model the process that generated these differences in initial conditions among agents; instead we assume that the differences in initial conditions are captured by some persistent, unobservable heterogeneity. We describe in Section 3 below how our methodology allows for such unobservable heterogeneity that may be correlated with observable initial conditions.[13]

An agent's period-a specific state variables include her prior work experience x_a, the number of her children n_a, the number of prior periods she had participated in welfare κ_a, and her last period decision $d_{a-1} \in D$. Thus $(x_a, n_a, \kappa_a, d_{a-1})$ represents the potentially time-varying period-a state variables. To summarize, an agent's period-a state variables $s_a = (j, g_0, x_a, n_a, \kappa_a, d_{a-1})$, and we write the space for the state variables at period a as $\mathcal{S}_a$. The evolution of the state variables are straightforward except for n_a, the number of children. We treat the arrival of additional children as exogenously determined and model births as a stochastic process given by

$$n_{a+1} = \begin{cases} n_a + 1, & \text{with probability } \rho(a, n_a, d_a) \\ n_a, & \text{with probability } 1 - \rho(a, n_a, d_a), \end{cases}$$

where $\rho(a, n_a, d_a)$ is a logistic function:

$$\rho(a, n_a, d_a) = \frac{\exp[\gamma_0 + \gamma_1 a + \gamma_2 n_a + \gamma_3 \mathtt{I}(d_a = 0) + \gamma_4 \mathtt{I}(d_a = 1)]}{1 + \exp[\gamma_0 + \gamma_1 a + \gamma_2 n_a + \gamma_3 \mathtt{I}(d_a = 0) + \gamma_4 \mathtt{I}(d_a = 1)]}. \tag{6}$$

Preferences. We now describe an agent's intertemporal preferences. We assume that an agent consumes all of the returns associated with her choice d in each period, and obtains an instantaneous utility $u_a = R_a(d; s_a, \epsilon_{da})$. An agent in period a is concerned about both her present and future instantaneous utilities. Let $U^a(u_a, u_{a+1}, ..., u_A)$ represent an agent's intertemporal preferences from the perspective of period a. We adopt a simple and now commonly used formulation of agents'

[13]We allow for unobservable heterogeneity in labor market skill endowment h_0; home production skill endowment e_0; non-welfare home production decay η; and welfare stigma ϕ.

potentially time-inconsistent preferences: (β, δ)-preferences (Phelps and Pollak 1968, Laibson 1997 and O'Donoghue and Rabin 1999a):

Definition 1. (β, δ)-preferences are intertemporal preferences represented by

$$U^a(u_a, ..., u_A) \equiv \delta^a u_a + \beta \sum_{t=a+1}^{A} \delta^t u_t$$

where $\beta \in (0,1], \delta \in (0,1]$, and $a \in \{a_0, a_0+1, ..., A\}$.

Following the terminology of O'Donoghue and Rabin (1999a), the parameter δ is called the *standard discount factor* and it represents the long-run, time-consistent discounting; and the parameter β is called the *present-bias factor* and it represents short-term impatience. The standard exponential discounting model is nested as a special case of (β, δ)-preference when $\beta = 1$. When $\beta \in (0,1)$, (β, δ)-preferences are "quasi-hyperbolic" discounting preferences in the terminology of Laibson (1997). We say that an agent's preference is time-consistent if $\beta = 1$, and is present-biased if $\beta \in (0,1)$.

Following previous studies of time-inconsistent preferences, we will analyze the behavior of an agent by thinking of the single individual as consisting of many autonomous *selves,* one for each period. Each period-a self chooses her current behavior to maximize her current preference $U^a(u_a, ..., u_A)$, while her future selves control her subsequent decisions. The literature on time-inconsistent preferences distinguishes between *naive* and *sophisticated* agents (Strotz 1956, Pollak 1968, O'Donoghue and Rabin 1999a, b). An agent is *partially naive* if the self in every period a underestimates the present-bias of her future selves, believing that her future selves' present bias is $\tilde{\beta} \in (\beta, 1)$; in the extreme, if the present self believes that her future selves are time-consistent, i.e. $\tilde{\beta} = 1$, she is said to be *completely naive.* On the other hand, an agent is *sophisticated* if the self in every period a correctly knows her future selves' present-bias β and anticipates their behavior when making her period-a decision.

2.1 Strategies, Payoffs and Equilibrium

We restrict our attention to Markov strategies and define a *feasible strategy* for a period-a self as a mapping $\sigma_a : \mathcal{S}_a \times \mathbb{R}^3 \to D$, where $\sigma_a(s_a, \boldsymbol{\epsilon}_a) \in \{0,1,2\}$ is simply the choice of the agent's period-a self over welfare, work or home when her state variables are s_a and the period-a shock vector is $\boldsymbol{\epsilon}_a = (\epsilon_{0a}, \epsilon_{1a}, \epsilon_{2a})$. With slight abuse of notation, we write $R_a(\sigma_a(s_a, \boldsymbol{\epsilon}_a); s_a, \boldsymbol{\epsilon}_a)$ as the instantaneous period-a utility the agent obtains from strategy σ_a when the state is s_a and shocks are $\boldsymbol{\epsilon}_a$.

A *strategy profile* for all selves is $\boldsymbol{\sigma} \equiv \{\sigma_t\}_{t=a_0}^{A}$. It specifies for each self her action in all possible states and under all possible realizations of shock vectors. For any strategy profile $\boldsymbol{\sigma}$, write $\boldsymbol{\sigma}_a^+ \equiv \{\sigma_t\}_{t=a}^{A}$ as the *continuation strategy profile* from period a to A. To define and characterize the equilibrium of the intrapersonal game of an agent with potentially time-inconsistent preferences, we first introduce a useful concept: write $V_a\left(s_a, \boldsymbol{\epsilon}_a; \boldsymbol{\sigma}_a^+\right)$ as the agent's period-a expected continuation utility when the state variable is s_a and the shock vector is $\boldsymbol{\epsilon}_a$ under her *long-run* time preference for a given a continuation strategy profile $\boldsymbol{\sigma}_a^+$. We can think of $V_a\left(s_a, \boldsymbol{\epsilon}_a; \boldsymbol{\sigma}_a^+\right)$ as representing (hypothetically) her intertemporal preferences from some prior perspective when her own present-bias is irrelevant. Specifically, $V_a\left(s_a, \boldsymbol{\epsilon}_a; \boldsymbol{\sigma}_a^+\right)$ can be calculated recursively as follows. First, let

$$V_A\left(s_A, \boldsymbol{\epsilon}_A; \boldsymbol{\sigma}_A^+\right) = R_A\left(\sigma_A\left(s_A, \boldsymbol{\epsilon}_A\right); s_A, \boldsymbol{\epsilon}_A\right) + \delta \mathrm{E}\left[W\left(s_{A+1}\right) | s_A, \sigma_A\left(s_A, \boldsymbol{\epsilon}_A\right)\right], \tag{7}$$

where $W\left(s_{A+1}\right)$ is the continuation value at the terminal age A as a function of the period-$(A+1)$ state variables; and the expectation is taken over the fertility shock conditional on s_A, as modelled by (6), and decision $\sigma_A\left(s_A, \boldsymbol{\epsilon}_A\right)$.[14] Recursively, for $a = A-1, ..., a_0$,

$$V_a\left(s_a, \boldsymbol{\epsilon}_a; \boldsymbol{\sigma}_a^+\right) = R_a\left(\sigma_a\left(s_a, \boldsymbol{\epsilon}_a\right); s_a, \boldsymbol{\epsilon}_a\right) + \delta \mathrm{E}\left[V_{a+1}\left(s_{a+1}, \boldsymbol{\epsilon}_{a+1}; \boldsymbol{\sigma}_{a+1}^+\right) | s_a, \sigma_a\left(s_a, \boldsymbol{\epsilon}_a\right)\right] \tag{8}$$

where the expectation is taken over both the conditional fertility shock and $\boldsymbol{\epsilon}_{a+1}$.

We will define the equilibrium for partially naive agents whose period-a self believes that, beginning next period, her future selves will behave optimally with a present-bias factor of $\tilde{\beta} \in [\beta, 1]$. Following O'Donoghue and Rabin's (1999b), we first define the concept of an agent's *perceived continuation strategy profile* by her future selves.

Definition 2. The *perceived continuation strategy profile* by a partially naive agent is a strategy profile $\tilde{\boldsymbol{\sigma}} \equiv \{\tilde{\sigma}_a\}_{a=a_0}^{A}$ such that for all $a \in \{a_0, ..., A\}$, all $s_a \in \mathcal{S}_a$, and all $\boldsymbol{\epsilon}_a \in \mathbb{R}^3$,

$$\tilde{\sigma}_a\left(s_a, \boldsymbol{\epsilon}_a\right) = \arg\max_{d \in D}\left\{R_a\left(d; s_a, \epsilon_{da}\right) + \tilde{\beta}\delta \mathrm{E}\left[V_{a+1}\left(s_{a+1}, \boldsymbol{\epsilon}_{a+1}; \tilde{\boldsymbol{\sigma}}_{a+1}^+\right) | s_a, d\right]\right\}.$$

That is, if an agent is partially naive with perceived present-bias by future selves at $\tilde{\beta}$, then her period-a self will anticipate that her future selves will follow strategies $\tilde{\boldsymbol{\sigma}}_{a+1}^+ \equiv \{\tilde{\sigma}_t\}_{t=a+1}^{A}$. Given this perception, the period-a self's best response is called *perception-perfect strategy profile.*

[14] In the empirical implementation, we approximate the continuation value by the following function of state variables:

$$W\left(s_{A+1}\right) = \omega_1 n_{A+1} + \omega_2 n_{A+1}^2 + \omega_3 x_{A+1} + \omega_4 x_{A+1}^2 + \omega_5 \mathtt{I}\left(d_A = 1\right) + \omega_6 \mathtt{I}\left(d_A = 2\right).$$

Definition 3. A *perception-perfect strategy profile* for a partially naive agent is a strategy profile $\boldsymbol{\sigma}^* \equiv \{\sigma_a^*\}_{a=a_0}^A$ such that, for all $a \in \{a_0, ..., A\}$, all $s_a \in \mathcal{S}_a$, and all $\boldsymbol{\epsilon}_a \in \mathbb{R}^3$,

$$\sigma_a^*(s_a, \boldsymbol{\epsilon}_a) = \arg\max_{d \in D} \left\{ R_a(d; s_a, \epsilon_{da}) + \beta\delta \mathrm{E}\left[V_{a+1}\left(s_{a+1}, \boldsymbol{\epsilon}_{a+1}; \tilde{\boldsymbol{\sigma}}_{a+1}^+\right) | s_a, d \right] \right\}.$$

When the agent is sophisticated, i.e., when $\tilde{\beta} = \beta$, we immediately have $\tilde{\boldsymbol{\sigma}} = \boldsymbol{\sigma}^*$. Thus the perception-perfect strategy profile is simply the familiar subgame perfect equilibrium of the intrapersonal conflict game. In our empirical implementation, we will report results for both completely naive $\left(\tilde{\beta} = 1\right)$ and sophisticated agents $\left(\tilde{\beta} = \beta\right)$.

2.2 Numerical Solution of $\boldsymbol{\sigma}^*$

In our empirical implementation, the terminal age A is finite. This allows us to numerically solve the perception-perfect strategy profile $\boldsymbol{\sigma}^*$ recursively. The solutions for sophisticated, and completely naive agents are merely special cases of the partially naive solution, so we describe how $\boldsymbol{\sigma}^*$ can be numerically solved for a partially naive agent.

First consider the terminal period A. For any $s_A \in \mathcal{S}_A$, and $\boldsymbol{\epsilon}_A \in \mathbb{R}^3$, the period-$A$ self's optimal strategy is simple:

$$\sigma_A^*(s_A, \boldsymbol{\epsilon}_a) = \arg\max_{d \in D} \left\{ R_A(d; s_A, \epsilon_{da}) + \beta\delta \mathrm{E}\left[W(s_{A+1}) | s_A, d \right] \right\}.$$

A partially naive agent at period-$(A-1)$, however, would perceive her period-A self would follow

$$\tilde{\sigma}_A(s_{A,}\boldsymbol{\epsilon}_A) = \arg\max_{d \in D} \left\{ R_A(d; s_A, \epsilon_{da}) + \tilde{\beta}\delta \mathrm{E}\left[W(s_{A+1}) | s_A, d \right] \right\}.$$

Now for every $a = A-1, ..., a_0$, every $s_a \in \mathcal{S}_a$, and every $\boldsymbol{\epsilon}_a \in \mathbb{R}^3$, we will have, recursively,

$$\begin{aligned} \tilde{\sigma}_a(s_a, \boldsymbol{\epsilon}_a) &= \arg\max_{d \in D} \left\{ R_a(d; s_a, \epsilon_{da}) + \tilde{\beta}\delta \mathrm{E}\left[V_{a+1}\left(s_{a+1}, \boldsymbol{\epsilon}_{a+1}; \tilde{\boldsymbol{\sigma}}_{a+1}^+\right) | s_a, d \right] \right\} \\ \sigma_a^*(s_a, \boldsymbol{\epsilon}_a) &= \arg\max_{d \in D} \left\{ R_a(d; s_a, \epsilon_{da}) + \beta\delta \mathrm{E}\left[V_{a+1}\left(s_{a+1}, \boldsymbol{\epsilon}_{a+1}; \tilde{\boldsymbol{\sigma}}_{a+1}^+\right) | s_a, d \right] \right\}, \end{aligned}$$

where $V_{a+1}(\cdot, \cdot; \cdot)$ is recursively defined by (7) and (8). This completes the recursion.

Informally, in equilibrium the individual's decision making proceeds as follows. Beginning at age a_0, the period-a_0 self observes her state variable s_{a_0} and then draws three choice-specific shocks $\boldsymbol{\epsilon}_{a_0} \in \mathrm{N}(0, \Omega)$. Given the anticipated behavior of her future selves, represented in the above recursive procedure by $\tilde{\boldsymbol{\sigma}}_{a_0+1}^+$, she calculates the realized current rewards and the expected future rewards from each of her three alternatives, using her own discount factors (β, δ). This calculation yields $\sigma_{a_0}^*(s_{a_0}, \boldsymbol{\epsilon}_{a_0})$, representing the alternative that offers the highest discounted present value. Then the state variable is updated for period-(a_0+1) according to the alternative chosen and the process is repeated. The perception-perfect strategy at each age a, for each $s_a \in \mathcal{S}_a$, is identified by the region in the three dimensional space of $\boldsymbol{\epsilon}_a$ over which each of the alternatives is optimal, for the given state s_a. Since there is no closed-form representation of this solution, we will, in the estimation and simulations below, solve the game numerically by backward recursion using crude Monte Carlo integration to approximate the expected continuation values $\mathrm{E}\left[V_{a+1}\left(s_{a+1}, \boldsymbol{\epsilon}_{a+1}; \tilde{\boldsymbol{\sigma}}^{+}_{a+1}\right)|s_a, d\right]$.[15]

3 Estimation Strategy

The solution to the intrapersonal game described above provides the inputs for estimating the parameters of the model by the following method. We first describe the structure of our data (see also Section 4). We have data on choices, state variables and related outcomes (such as welfare benefit levels and accepted wages) from a sample of agents each of whom solves the intrapersonal conflict game. In what follows, we use superscript $i \in \{1, ..., N\}$ to index the agents. Our data set consists of three sets of information: (1) agent i's sequence of state variables represented by $\mathbf{s}^i \equiv \left\{s_a^i\right\}_{a=a_0^i}^{A^i}$ where a_0^i denotes the age at which individual i gave birth to her first child (which is the time she becomes part of our analysis) and A^i is the age at which we last observe the agent; (2) agent i's sequences of choices $\mathbf{d}^i \equiv \left\{d_a^i\right\}_{a=a_0^i}^{A^i}$; (3) if agent i chooses to work, we observe her accepted wages, which we write as $\mathbf{w}^i \equiv \left\{w_a^i\right\}_{a=a_0^i}^{A^i}$ with the understanding that $w_a^i = \emptyset$ if $d_a^i \neq 1$. We also have a separate data set that provides the welfare benefit levels for families of different sizes for all the States, denoted by G_j where j indexes the State. We denote our data set by $\mathcal{D}$.

The decision at any age a is deterministic for the agent for a given vector $(s_a, \boldsymbol{\epsilon}_a) \in \mathcal{S}_a \times \mathbb{R}^3$, but it is probabilistic from our perspective since we do not observe the shock vector $\boldsymbol{\epsilon}_a$. As we described in the last paragraph of Subsection 2.2, for a given parameters of the model $\boldsymbol{\Theta}$, we can numerically solve for the perception-perfect strategy profile as the solution to the game, and it then provides the probability of choosing alterative d_a^i at state s_a^i; and if $d_a^i = 1$, receiving wage w_a^i, denoted by

$$\Pr\left[d_a^i, w_a^i|s_a^i; \boldsymbol{\Theta}\right].$$

We can therefore consistently estimate $\boldsymbol{\Theta}$ by maximizing with respect to $\boldsymbol{\Theta}$ the sample likelihood

$$\prod_{i=1}^{N} \prod_{a=a_0^i}^{A^i} \Pr\left[d_a^i, w_a^i|s_a^i; \boldsymbol{\Theta}\right].$$

[15] The numerical solution method we employ follows closely Keane and Wolpin (1994). However, because the state space of our model is, conditional on type, relatively small (roughly 150,000 elements at age $A = 34$), we do not use Keane and Wolpin's method for approximating the expected continuation values using only a subset of the state space. Instead we approximate the expected continuation value for *every* element of the state space by Monte Carlo integration. Based on sensitivity analysis we chose to rely on 150 draws from the $\boldsymbol{\epsilon}$ distribution to perform this integration.

To ease the computational burden, we take two shortcuts before maximizing the sample likelihood. First, the parameters $\boldsymbol{\theta}_j \equiv (\theta_{0j}, \theta_{1j})$ in the welfare benefits function $G_j(\cdot)$ [see Eq. (3)] are taken as the mean of estimated real benefit function in the agent's State of residence j over the period observed.[16] Table A2 of the Appendix presents these estimated parameters, and summary statistics for the 20 U.S. States represented in the sample.

Second, we estimate the parameters $\boldsymbol{\gamma} \equiv (\gamma_0, \gamma_1, \gamma_2, \gamma_3, \gamma_4)$ of the fertility function $\rho(a, n_a, d_a)$ [See Eq. (6)] separately by estimating a logit.

Given this set of estimated parameters, the remaining parameters of the model, including those in the utility function, the returns functions, and the variance-covariance matrix of the shocks Ω, denoted by $\tilde{\boldsymbol{\Theta}}$, are estimated by maximizing $\tilde{\boldsymbol{\Theta}}$ over a restricted likelihood function:[17]

$$\mathcal{L}\left(\tilde{\boldsymbol{\Theta}};\mathcal{D}\right) = \prod_{i=1}^{N} \prod_{a=a_0^i}^{A^i} \Pr\left[d_a^i, w_a^i | s_a^i; \left(\tilde{\boldsymbol{\Theta}}, \hat{\boldsymbol{\theta}}_j, \hat{\boldsymbol{\gamma}}\right)\right]. \tag{9}$$

For each observation i, $\Pr\left[d_a^i, w_a^i | s_a^i; \left(\tilde{\boldsymbol{\Theta}}, \hat{\boldsymbol{\theta}}_j, \hat{\boldsymbol{\gamma}}\right)\right]$ is a three-dimensional integral which we approximate using 300 Monte Carlo draws to form kernel-smoothed simulators of the probabilities.[18]

3.1 Unobserved Heterogeneity

The likelihood function in (9) applies to a sample that is homogenous except for the following observable initial conditions at the birth of the first child: age a_0^i, education g_0^i, work experience $x_{a_0^i}^i$, previous period choice $d_{a_0^i-1}^i$ and the State of residence j^i. The skills and preferences of individuals are likely to vary, however, in unobservable ways that are both persistent, and correlated with observable initial conditions. For example, those with greater endowments of unobservable human capital may be more likely to prolong schooling and postpone both childbirth and entry into the workforce. Ignoring this persistent heterogeneity would generate inconsistent estimates of the model's parameters.

---

[16] We thank Ken Wolpin for providing us with these estimates. The State of residence is defined as the State in which the respondent resided at the birth of her first child.

[17] To ease identification and the computational burden we assume $\mathrm{Cov}(\epsilon_{0a}, \epsilon_{1a}) = \mathrm{Cov}(\epsilon_{1a}, \epsilon_2 a) = 0$. The remaining elements of the variance-covariance matrix are estimated.

[18] We chose 300 draws after tests for sensitivity of the simulated probabilities and data fit to changes in the number of repetitions. The kernel of the simulated integral is given by:

$$\exp\left[\frac{Q_d^a - \max_{d\in D}\left(Q_d^a\right)}{\tau}\right] / \sum_{d=0}^{2} \exp\left[\frac{Q_d^a - \max_{d\in D}\left(Q_d^a\right)}{\tau}\right]$$

where $Q_d^a = R_a(d; s_a, \boldsymbol{\epsilon}_a) + \beta\delta \mathrm{E}\left[V_{a+1}\left(s_{a+1}, \boldsymbol{\epsilon}_{a+1}; \tilde{\boldsymbol{\sigma}}_{a+1}^{+}\right) | s_a, d\right]$ is the present value of choosing alternative d at period a; and τ is the smoothing parameter. In the estimation results that follow, τ is set to 150, again based on sensitivity analysis. For related application of kernel smooother, see Eckstein and Wolpin (1999).

To allow for the possibility of persistent heterogeneity correlated with initial conditions, we posit that agents can be of K possible types, indexed by $k \in \{1, ..., K\}$; and allow different types of agents to differ in their home production skill endowment e_0, unobservable labor market skill endowment h_0, welfare stigma ϕ, and non-welfare home production decay parameter η. In our estimation, these parameters will be type specific and denoted by $e_0^{(k)}, \phi^{(k)}, h_0^{(k)}$ and $\eta^{(k)}$ for each $k \in \{1, ..., K\}$ respectively.[19]

The *ex ante* probability that an agent i is of type k is denoted by P_k^i. To capture correlation between an agent's unobservable type and her initial conditions, we allow P_k^i to depend on all of her observable initial conditions except State of residence in the form of a multinomial logit.[20] That is, for $k = 2, ..., K$,

$$\begin{aligned}
P_k^i &= P_k\left(s_{a_0^i}; \boldsymbol{\pi}\right) \\
&= \frac{\exp\left[\pi_0^{(k)} + \pi_1^{(k)} a_0^i + \pi_2^{(k)} g_0^i + \pi_3^{(k)} x_{a_0^i}^i + \pi_4^{(k)} \mathrm{I}\left(d_{a_0^i-1}^i = 0\right) + \pi_5^{(k)} \mathrm{I}\left(d_{a_0^i-1}^i = 1\right)\right]}{1 + \sum_{l=2}^K \exp\left[\pi_0^{(l)} + \pi_1^{(l)} a_0^i + \pi_2^{(l)} g_0^i + \pi_3^{(l)} x_{a_0^i}^i + \pi_4^{(l)} \mathrm{I}\left(d_{a_0^i-1}^i = 0\right) + \pi_5^{(l)} \mathrm{I}\left(d_{a_0^i-1}^i = 1\right)\right]}
\end{aligned}$$

and normalize $P_1^i\left(s_{a_0^i}\right)$ as

$$\begin{aligned}
P_1^i &= P_1\left(s_{a_0^i}; \boldsymbol{\pi}\right) \\
&= \frac{1}{1 + \sum_{l=2}^K \exp\left[\pi_0^{(l)} + \pi_1^{(l)} a_0^i + \pi_2^{(l)} g_0^i + \pi_3^{(l)} x_{a_0^i}^i + \pi_4^{(l)} \mathrm{I}\left(d_{a_0^i-1}^i = 0\right) + \pi_5^{(l)} \mathrm{I}\left(d_{a_0^i-1}^i = 1\right)\right]}.
\end{aligned}$$

where $\boldsymbol{\pi} \equiv \left\{\pi_0^{(l)}, ..., \pi_5^{(l)}\right\}_{l=2}^K$. Now write $\tilde{\boldsymbol{\Theta}}^k$ as the set of model parameters for type-k agent to be estimated by simulated maximum likelihood, the sample likelihood, integrating over all types, can be written as:

$$\tilde{\mathcal{L}}\left(\tilde{\boldsymbol{\Theta}}^1, ..., \tilde{\boldsymbol{\Theta}}^K, \boldsymbol{\pi}; \mathcal{D}\right) = \prod_{i=1}^N \sum_{k=1}^K P_k\left(s_{a_0^i}; \boldsymbol{\pi}\right) \prod_{a=a_0^i}^{A^i} \Pr\left[d_a^i, w_a^i | s_a^i; \left(\tilde{\boldsymbol{\Theta}}^k, \hat{\boldsymbol{\theta}}_j, \hat{\boldsymbol{\gamma}}\right)\right]. \tag{10}$$

3.2 Identification of β and δ

We now address the separate identification of the present-bias factor β and the standard discount factor δ. In some models, it has been shown that the decisions of sophisticated present-biased agents are observationally equivalent to those of time-consistent exponential discounters, and identification

[19] In our estimation, we choose $K = 3$ after sensitivity analysis.

[20] We omit State of residence because the variation in welfare benefits in the data provides an important source of identification for the model's parameters. Allowing type to depend on State of residence would sharply weaken our ability to identify, in particular, unobserved home production and welfare stigma parameters from variation in decisions correlated with variation in initial conditions, welfare benefits and wages.

of the two parameters is thus *a priori* precluded. For example, Barro (1999) demonstrates the observational equivalence in a growth model with sophisticated agents, perfect credit markets, and log utility. The equivalence does not hold more generally. Harris and Laibson (2001) show, for example, that observational equivalence is not obtained when the assumption of perfect credit markets is relaxed. This illustrates that, as is true in any structural empirical paper, the ability to separately identify β and δ results from both the structures imposed by the model and the variations in the data. In what follows, we hope to provide identification arguments from three different angles.

Qualitative Differences in the Behavioral Effects of Changes in β and δ. We first argue that changes in β (while holding δ constant) will have a qualitatively different effect on welfare versus work decisions than similar changes in δ (while holding β constant). Specifically, in a companion paper (Fang and Silverman 2004a), we show in a deterministic setting that, if stigma ϕ lasts for only one period after switching into welfare (as it does here), then: greater short-term patience (a higher β) will lead to *more* welfare participation, while greater long-term patience (a higher δ) will lead to *less* welfare participation. Consider the following deterministic three-period numerical example. Suppose that net wages are a deterministic increasing function of work experience x given by $w(0) = -1.5, w(1) = -1.0$ and $w(2) = 4.5$; and the welfare benefits are constant at 2, but the one shot stigma is $\phi = 4$. Thus, the net welfare benefits as a function of welfare experience κ is given by $b(0) = -2, b(1) = b(2) = 2$. Table 1 shows the discounted payoffs to work versus welfare from the perspective of the first period self for different values of β and δ. When $\beta = 0.80$ and $\delta = 0.95$, the discounted payoff from work for the first period self is $w(0) + \beta \sum_{x=1}^{2} \delta^x w(x) = 0.99$, and the discounted payoff from welfare is $b(0) + \beta \sum_{\kappa=1}^{2} \delta^\kappa b(\kappa) = 0.96$. Given the increasing profiles, it follows that in equilibrium the agent chooses work in all three periods. An increase in the long term patience δ from 0.95 to 1.0 (while keeping the short term patience β fixed at 0.8) merely serves to widen the gap between the returns to work versus welfare. If we think of these payoffs as expectations, the increase in long-term patience serves to increase the probability the agent works. On the other hand, if we increase the short term patience β to 0.90 (while keeping the long term patience fixed at $\delta = 0.95$), the discounted payoff from welfare now *exceeds* that from work. That is, an increase in the short term patience leads the agent to choose welfare in all three periods.

[Table 1 About Here]

Intuition for these countervailing effects of increased short-term versus long-term patience comes from consideration of a single mother who was living at home in the previous period. For such a decision-maker, the costs of welfare stigma, and the loss of special home production loom large but the gains from enduring the stigma today will be realized in large measure beginning next period. For present-biased agents, this one-time stigma and their taste for immediate gratification may induce a delay in welfare take-up – a delay that the same agent would not choose if she were time-consistent. The greater the degree of short term-impatience, the more likely she is to delay the welfare take-up decision. Holding the present-bias factor constant, however, increases in long-term patience will disproportionately affect the incentives to work that derive only from the returns to experience obtained in the relatively distant future. These same increases in long-term patience may, however, leave the relatively immediate trade-offs between home and welfare little changed. As a result, greater long-term patience will decrease the likelihood of choosing welfare in the expected way. Thus, we would expect that changes in β could affect the conditional probability of choosing welfare in the opposite way from similar changes in δ, and thus reveal through the structure of the model, rates of time discount in the near versus the long term.

[Table 2 About Here]

In Table 2, we summarize the results of a series of simulation experiments suggesting that this intuition from the deterministic case carries through and that the countervailing effects of β and δ hold locally. Panel 1 of Table 2 considers the simulated decisions of individuals at various levels of the discount factor δ, holding the present-bias factor (and all other parameters fixed).[21] As δ increases, the fraction of individuals choosing welfare declines monotonically. When $\delta = 0.84$, agents on average choose welfare 47.7 percent of the time. Increasing the standard discount factor one standard deviation to $\delta = 0.86$ leads to a small decline in average welfare participation to 46.9 percent. As we continue to increase δ, average welfare participation continues to decline ending at 40.3 percent participation when agents have a standard discount factor of 0.92.

Panel 2 performs the parallel exercise for variations in β, holding δ fixed. Here the results differ qualitatively. Following the intuition provided above, greater short-term patience *increases* the probability of choosing welfare. When $\beta = 0.20$, agents choose welfare on average 43.8 percent of the time. When $\beta = 0.41$ this fraction increases to 46.5 percent, and when β rises by another standard deviation, average participation increases to 47.2 percent. Similarly, we find that the simulated effects of changes in β and δ may differ qualitatively for the same individual depending on her age.

[21]Except for β and δ, all other parameters are set to their estimated levels for sophisticated present-biased model agents (see Tables 5 and 6).

Formal Arguments for Distinguishing Exponential and Hyperbolic Discounting in a Simpler Model. In another related paper (Fang and Silverman 2004b), we studied the non-parametric distinction of exponential and hyperbolic discounting models of welfare program participation. The model there is more complicated in allowing for bias in predicting future cost of working, but is simpler in only allowing for stochastic component in the utility from work, but not from home or welfare. In that model, we show that if we have three or more periods of observations, then a present bias model with $\beta \in (0,1)$, $\delta \in (0,1)$ can be distinguished from exponential discounting model with $\beta = 1$ and $\delta \in (0,1)$ using standard data without making parametric assumptions on the distribution of stochastic shock in the payoff from working (see Proposition 2). While the arguments used in the proof do not generalize to the case when we allow for stochastic components in all three choices, this result informs us that, (β, δ)-discounting *per se* does not create problems in identification in this context. Finally, note that, while it is true that our model in this paper allows for stochastic shocks in all three choices, we have, as is standard, imposed a parametric functional form on their joint distribution. In this sense, we have modelled a setting where identification is easier than that considered in Fang and Silverman (2004b).

Likelihood Surface. In practice, whether the two discount parameters are separately identified with reasonable precision depends on the curvature of the likelihood surface as we vary β and δ. As an illustration, Figure 1 presents the three-dimensional surface of the log-likelihood as a function of β and δ when other parameters are set at their respective estimates. It is particularly worth noting that when β is 1, the log likelihood is substantially lower than its global maximum.

[**Figure 1 About Here**]

The above three arguments do not, of course, provide a definitive statement about the identification of the highly non-linear model we estimate in this paper. Taken together with the reasonably small standard errors presented below, however, they do suggest that the theoretical bases for identification suggested in our companion papers may be valid in this more realistic model. We believe that the considerable amount of variations in wages and welfare benefits across agents, States and time in the data, together with our assumption of a one-period welfare stigma in the model [see Eq. (4)], indeed provide a field data analogue of the lab experiments in which subjects reveal their rates of time discount in both the near and the longer term.

4 Data

4.1 Sample Definition

The data are taken from the 1979 youth cohort of the National Longitudinal Surveys of Labor Market Experience (NLSY). The NLSY began in 1979 with 6,283 women ages 14-22, and has interviewed this cohort annually up to 1994, and biannually since 1994. We restrict attention to the 675 women who, as of their interview in 1992, had both remained unmarried and given birth to at least one child during the years they were surveyed. We then consider only the decisions each individual made after the birth of her first child and during the calendar years 1978-1991, assuming she continued to reside in the State in which she lived at the birth of her first child.

Our purpose in selecting this subsample of individuals and years is threefold. First, to be consistent with our model, we want to restrict attention to those who, if they do not work, are almost certainly eligible for welfare by virtue of having a child and being unmarried. Second, to better justify our assumption that marriage decisions are not germane, we restrict attention to women who never marry. Third, we want to limit our analysis to decisions made before the changes in welfare eligibility rules beginning in 1993, and easily anticipated by 1992. Finally, again to ease the computational burden, we further limit our sample to residents of the 20 States best represented in the data. This final restriction leaves us with 483 individuals taken from the NLSY's core random sample and its oversamples of blacks and Hispanics. The women in our subsample were observed with at least one child for an average of 9.3 of the 14 years from 1978-1991, providing us with 4,487 state-choice observations for the estimation. Note, however, our sample selection criterion also suggests caution in generalizing the estimates in this paper to the overall population.

4.2 Period and Variable Definition

At each interview, the NLSY collects welfare participation data as a monthly event history recorded back to the preceding interview. The survey's employment data are collected as a weekly event history. We assume the decision period of the model corresponds to a calendar year, and identify an agent as age a in a year if she was a years old for at least half that year. The decisions at each age a are defined as follows: An individual chose welfare at age a if she received Aid to Families with Dependent Children (AFDC) for at least six months of the year during which she was a years old. An individual chose work at age a if she was employed for at least 1,500 hours of the year during which she was a years old. An agent chose to stay home if she chose neither of the above.[22,23]

4.3 Descriptive Statistics

Descriptive statistics of the subsample are presented, by age, in Table A1 of Appendix A. Since none of the women in the subsample marries during the period she is observed, the group we study is not typical of the general US population. To better understand the ways in which members of the subsample differ from the average population, Table A1 also compares their statistics with those of the entire sample of women in the NLSY from 1978-1991. Broadly, this comparison suggests that while the subsample represents the targets of the U.S. welfare policy, it is atypical of the population as a whole.

[Table 3 About Here]

The distribution of choices among welfare, work, and home is presented by age in Table 3. We concentrate on the decisions made at ages 16-32, that represent 98 percent of the data. The fraction of the subsample choosing welfare increases considerably between ages 16 and 22. Of the 16 year-olds with at least one child, 32 percent chose welfare while 54 percent of 22 year-olds with children chose welfare. The proportion choosing work exhibits a comparable increase over the same period: rising from zero percent of 16 year-olds with children to 17 percent of 22 year-olds. Given these changes in welfare and work participation we, by definition, observe a more dramatic decline in the fraction of women with children choosing to remain at home; with 68 percent choosing to stay home at age 16 and just 29 percent choosing to stay at home at age 22.

While these basic trends continue for the fractions choosing work and home beyond age 22, the fraction choosing welfare stops increasing and instead exhibits a slow decline after age 22. By age 25, 47 percent of the sample is now choosing welfare, despite having on average more children. By age 29, the fraction is 43 percent.

Not all of the movements in these age-decision profiles reflect the changing choices of the same individuals. The observed transitions are partly due to the fact that the composition of the sample is changing as the women of the NLSY age and, by virtue of having a child, join the subsample. To investigate the degree to which the choices of same individuals change over time, Table 4 presents the one-period transition rates between decisions by the same agent. Here we see evidence of

[22] On 19 occasions a respondent reported that she both received AFDC for at least 6 months of the previous calendar year *and* worked more than 1,500 hours that year. In these cases the agent was defined as having chosen welfare.

[23] In about 9% of our observations, the respondent was attending school. This part of the sample is concentrated among agents younger than 18. These observations are also concentrated in the sample we classify as "choosing home," among which 15.7% was attending school.

considerable persistence in individuals' choices. The rows of Table 4 represent the choices made in period $t-1$; the columns describe the choices made in period t. The top figure (Row %) in each cell represents the fraction of the subsample that made the row choice in period $t-1$ who went on to make the column choice in period t. The bottom figure (Column %) in each cell shows the fraction of the subsample that made the column choice in period t who made the row choice in the previous period. We find that 84.3% of those who chose welfare in period $t-1$ went on to choose it again in period t. Conversely, of those who chose welfare in period t, 76.7% had chosen welfare in the previous period. Of those who chose work in period $t-1$, 79.3% went on to choose it again in period t. Decisions to remain at home are considerably less persistent. Of those who chose to stay home in period $t-1$, 59.7% chose it again in period t.

[**Table 4 About Here**]

5 Results

5.1 Estimates of Welfare Benefit Function G_j and Fertility Function ρ

Table A2 in Appendix A presents the parameters of the benefit rule for the twenty selected States used in our estimation. As has been often noted, there is considerable variation in benefits levels across States. In our sample, the estimated average annual benefit for a mother with two children ranges from $4,856 (1987 dollars) to $9,490. Patterns of welfare participation vary with the level of benefits in ways consistent with optimizing behavior. In our sample, 56 percent of the residents in the 5 States with the highest benefits received welfare, while 37 percent of these in the 5 States with the lowest benefits were on welfare.

Table A3 in Appendix A presents the parameter estimate of the fertility function (6). These parameter estimates suggest that the probability of an additional birth is decreasing with age and with the number of children. The estimate also indicates that, relative to those who stay home, the probability of an additional birth is lower for workers and higher for those on welfare. We note, however, that our simple exogenous model of subsequent fertility beyond the first child explains very little of the variation in the timing of births in this subsample. The pseudo-R^2 is less than two percent.

5.2 Parameter Estimates

In our estimation, we assume that agents are of three possible types, i.e., $K = 3$. Tables 5 and 6 present the parameter estimates under three different restrictions of the model. Column (1) presents estimates when we restrict agents to be time-consistent, that is, restricting $\beta = 1$;

Column (2) presents estimates when agents are assumed to be sophisticated and present-biased (i.e. $\tilde{\beta} = \beta$); and Column (3) presents the estimates when agents are assumed to be completely naive present-biased $\left(\tilde{\beta} = 1\right)$. We present both the point estimates and their asymptotic standard errors.[24]

[Tables 5-6 About Here]

In the sophisticated present-bias model, the estimated present-bias factor β equals 0.33802 with a reasonably small standard error of 0.06943. A Wald test would also reject the hypotheses of time-consistency (t-statistic 9.53 against the null of $\beta = 1$), indicating that β is significantly different from 1. Allowing for present-bias improves the data fit in a statistically significant way, and a likelihood ratio test easily rejects the time-consistent model (the χ^2 statistic for the likelihood ratio test is over 32). This provides evidence that the behavior of the single mothers are affected by present bias. However, the likelihood ratio test does not yield overwhelming evidence in favor of the completely naive or sophisticated model.[25] Note that both imposed a particular restrictions on $\tilde{\beta}$. This also indicates that a completely unrestricted model (i.e., $\tilde{\beta}$ is also a parameter to be estimated) is probably not identified. In what follows, we will focus on the results from the sophisticated present-biased agent model.[26]

Combined with the estimated standard discount factor $\delta = 0.87507$, our estimate of the present-bias factor implies a one-year ahead discount rate of 238%. Our estimate of the present-bias factor is low relative to most of those estimated in experimental studies, though more similar to Paserman's (2002) structural estimate for low wage workers. Inferential studies such as Hausman (1979), and Warner and Pleeter (2001) estimate discount rates ranging from 0 to 89% depending on the characteristics of the individual and intertemporal trade-offs at stake. Paserman finds, for low-wage workers a discount rate of about 149%. Laibson, Repetto and Tobacman's (2004) point estimates of β and δ are respectively 0.7031 (with standard error of 0.1093) and 0.9580 (with standard error of 0.0068), which imply a one-year ahead discount rate of 48.5%. There are two possible explanations for the difference between our finding and others. First, the samples cover different subpopulations. Our sample includes mostly poor, never-married women with children (see Table A1). Thus it is

[24]Asymptotic standard errors are estimated using the BHHH, or outer product of gradients, method. See Berndt, et al (1974).

[25]Technically we can not use likelihood ratio test to distinguish completely naive and sophisticated present biased models because they are not nested: the completely naive model restricts $\tilde{\beta} = 1$, and the sophisticated model restricts $\tilde{\beta} = \beta$.

[26]The results for naive present-biased agents are qualitatively and quantitatively similar. The key simulation results are included for completeness in Appendix B.

possible that this subpopulation is more susceptible to present-biases. Second, different papers focus on different spheres on decision making. It is very possible that the magnitudes of present bias differ by specific decisions.

Besides the discount factors, Tables 5 and 6 also present estimates, by (unobservable) type, the net welfare stigma, home production functions, wage functions, continuation value functions, and variance-covariance matrix of the shocks, etc. Of particular interest is the substantial estimated return to experience in the wage offer function, and the considerable variation in the estimated skills and tastes across types. There is an important average gain in wages for every year of additional work experience. The unobservable skill levels that determine those wages vary importantly, however, by type.

5.3 Within-sample fit

Age-Choice Profiles. Summarizing the interaction of the potentially complex and countervailing effects of time preferences and basic incentives, Figures 2-4 compare the estimated model's predicted distributions over the three alternatives welfare, work and home, to the actual distributions in the data, by age. The model's predictions represent the simulated decisions of 1,000 agents in each of 16 cells defined to reflect the sample variation in initial conditions j, a_0, g_0, x_{a_0}, and d_{a_0-1}. There are four different j categories defined as high, medium-high, medium-low, and low benefits municipality. Similarly there are four g_0 categories defined as 10 years of schooling or less, 11 years of schooling, 12 years of schooling, and some college at the birth of the first child. Within each of these 16 cells, the initial conditions are given by the sample average (benefits, age, schooling, experience) level in the cell. These sample averages imply probabilities of the agent being of the three different unobservable types. The distribution of the 1,000 simulated decisions in each of these cells is then weighted by the probability of each type and the proportion of the data falling into that initial condition cell to generate the predicted distributions appearing in Figures 2-4.

[**Figures 2-4 About Here**]

The simulated age profiles match the data quite well. Each of the profiles implied by the estimated model assumes approximately the correct shape, and mostly matches the levels of the data quite closely. More formally, Table A4 of the appendix presents the within-sample χ^2 goodness of fit statistics for the model with respect to the choice distribution, by age. These statistics confirm the impression given by Figures 2-4.

Transition Probabilities. Table 7 presents the simulated one-period transition probabilities for the sophisticated present-biased agent model. This table is to be compared with the transition probability matrix in the data (see Table 4). The model matches the persistence and relative rates of transition quite well. To illustrate, the estimated model predicts that 84.4% of those who chose welfare in period $t-1$ will go on to choose it again the following period while 11.4% will choose to stay at home. These figures should be compared with 84.3% and 12.3% observed in the data. Similarly the model predicts that 57.0% of those choosing home in period $t-1$ will remain at home next period, while 25.9% will switch to welfare, comparable to 59.7% and 28.3% respectively in the data.

[**Table 7 About Here**]

Wage Profiles. Figures 5 and 6 compare, respectively, the model's mean wage-age and wage-experience profiles, with the parallel moments in the data. Save the outlying wages of age-18 workers, the model somewhat underestimates of average wages for those who choose to work (see Figure 5). Overall, however, the average accepted wages, by age, of the model and data are quite similar. Save the accepted wages of those with no experience, the model slightly underestimates wage levels while replicating the observed shape of the wage-experience profile (see Figure 6).

[**Figures 5-6 About Here**]

5.4 Out-of-Sample Fit

As we mentioned in Subsection 4.1, we have used only residents of the 20 States best represented in the NLSY in our empirical estimation. The sample of single mothers from the remaining States allow us to examine the out-of-sample fit of our model and the estimates. Figure 7 compares the proportions of single mothers choosing welfare, work and home by age predicted by our present-biased sophisticated model using the parameter estimates in Section 5.2 with their data counterparts. The model is able to capture the relative shape of changes in the participation rates as the women get older. For example, the model's prediction of the increase in the proportion of working single mothers mirrors that in the data. However, our model consistently overestimates the proportion of single mothers on welfare and underestimates the proportion at home.

[**Figure 7 About Here**]

6 Numerical Simulations

The estimates and simulations presented in Subsection 5.2 indicate that the work-welfare-home decisions of never-married women with children can be well described by a model of time-inconsistent preferences. With a reasonable degree of precision, the estimated model indicates a present-bias factor (β) substantially less than unity; and the model matches many aspects of observed decision and accepted wage profiles. But as we have said earlier, we could not clearly distinguish sophisticated and naive present-biased models - both models are able to fit the data much better than a time-consistent model. In this section, we present simulation results for sophisticated present-biased agents. Analogous results for completely naive present-biased agents are included in Appendix B.

On its own, an estimated β less than one does not imply that time-inconsistency importantly influences the work-welfare decisions of never-married women with children. It may be that the ability to commit to future decisions influences behavior in statistically identifiable, but economically insubstantial, ways. This possibility is particularly relevant for the model estimated here. In our model, initial conditions such as welfare benefits in State of residence, and years of schooling, and unobservable skills differ across individuals; and these differences would be expected to importantly influence decision making. While time-inconsistency in preferences may affect marginal decisions, it may be that the influence of initial conditions typically places individuals far from these margins, and the ability to commit would have little effect on decisions. By extension, if most individuals are little influenced by their inability to commit to future decisions, the behavioral and utility consequences of policies such as time limits or workfare that may serve as commitment devices will not much depend on the time-inconsistency of their targets.[27] In the next sections we use the estimated parameters of the sophisticated present-biased model to quantify both the behavioral and utility consequences of the ability to commit, and consider how different policy reforms affect both behavior and utility in the presence of time-inconsistency.

6.1 Consequences of an Ability to Commit

To evaluate the consequences of an ability to commit to future decisions, we use the estimated parameters of the sophisticated present-biased model to simulate the decisions of agents with various initial conditions both with and without commitment ability. In this experiment, an individual has commitment ability if, starting from the period in which her first child is born, her future selves behave as though they were time-consistent (i.e. $\beta = 1$), and believed all of their future selves also to be time-consistent. Equilibrium behavior represents the optimal plan of an individual considering the sequence of decisions to begin at the birth of her first child.

Evaluating utility effects in a setting with time-inconsistency is often thought to be especially problematic because sequences of utility flows may be valued differently by the different selves of

[27] See Fang and Silverman (2004a) for a discussion of how time-limits could serve as commitment mechanisms.

the same individual.[28] In the literature on time-inconsistency, two criteria have been proposed to serve as a basis for comparing an agent's well being: the *Pareto* criterion (Laibson 1997) and the *long-run utility* criterion (O'Donoghue and Rabin 1999a). The Pareto criterion asks if all the selves are made better off; while the long-run utility criterion takes the perspective of an effectively time-consistent agent just prior to the decision-making sequence, and asks if she is made better off. From the perspective of policy evaluation, it is not obvious which of these criteria is the more appropriate. Based on its similarity to prior utility evaluations in structural estimation (see, e.g., Keane and Wolpin 1997), we adopt the latter criterion for estimating changes in well-being. Specifically, we calculate the discounted stream of expected lifetime utility for period-a_0 self, i.e., the self when her first child was born if, counterfactually, $\beta = 1$. Explicitly, we first numerically solve for the perception-perfect strategy profile, denoted by $\boldsymbol{\sigma}^{\mathrm{c}*} \equiv \{\sigma_a^{\mathrm{c}*}\}_{a_0}^{A}$ for an agent with $\beta = 1$ and $\delta = \hat{\delta} = 0.875$, the point estimate presented in Table 5. Of course, $\boldsymbol{\sigma}^{\mathrm{c}*}$ depends on agents' initial conditions at period a_0. Conditional on an agent's initial conditions, her utility with commitment ability is given by

$$U^{\mathrm{c}} = \mathrm{E} \sum_{a=a_0}^{A} \delta^t R_a \left(\boldsymbol{\sigma}^{\mathrm{c}*}\left(s_a, \boldsymbol{\epsilon}_a\right); s_a, \boldsymbol{\epsilon}_a\right).$$

As a benchmark for comparison, when agents do not have ability to commit, we also numerically solve for the perception-perfect strategy profile, denoted by $\boldsymbol{\sigma}^{\mathrm{n}*} \equiv \{\sigma_a^{\mathrm{n}*}\}_{a_0}^{A}$ for an agent with $\beta = \hat{\beta} = 0.338$ and $\delta = \hat{\delta} = 0.875$. Conditional on an agent's initial conditions, the utility without commitment ability that we use as benchmark comparison is given by

$$U^{\mathrm{n}} = \mathrm{E} \sum_{a=a_0}^{A} \delta^t R_a \left(\sigma_a^{\mathrm{n}*}\left(s_a, \boldsymbol{\epsilon}_a\right); s_a, \boldsymbol{\epsilon}_a\right).$$

Note that U^{n} is not how period-a_0 self would have evaluated the lifetime utility with her (β, δ) preference. We reported below $\left(U^c - U^{\mathrm{n}}\right)/U^{\mathrm{n}}$ as the percentage change in lifetime utility as a result of the ability to commit. Representative results of the simulations are presented, by initial conditions cell, in Table 8. The cells (1-8) vary according to the level of the benefits in the State of residence, age and years of schooling at first birth, and thus probability of being types 1 and 2. The levels of these initial conditions are presented in second panel in Table 8. The same initial conditions (by cell) are used in subsequent tables.

[Table 8 About Here]

[28] As Caplin and Leahy (2000) points out, models of time-consistent discounting also exhibit this feature. Suppose, at the birth of her first child, a time-consistent decision maker rationally chooses a career of welfare rather than work because the short-term costs of work exceed the discounted long-term gains. Ten years later she would regard the streams of utility coming from work versus welfare very differently; she would strictly prefer that she had worked for the previous 10 years.

Panel 1 of Table 8 indicates that while the behavioral effects of an inability to commit may be large, they differ both in size and sign depending on initial conditions. For example, among individuals in cell 2, who were relatively young and little educated at the birth of their first child, and who live in a high benefits State, work is relatively unattractive and commitment ability leads them to work somewhat *less* (2.74%) of the time between ages 18 and 34. For this group, the inability to commit generated costly delay, not in work, but in the takeup of welfare. With commitment ability, they are quicker to endure welfare stigma in exchange for the future benefit of welfare receipt. Compare this effect of commitment to that of similarly young and but better educated individuals in a low benefits State (cell 3). In this second group, for whom working is relatively more attractive, the ability to commit leads them to work an additional 23.32% of the time, representing a 66 percent increase in their probability of working. Comparing across other cells, we observe similar disparities in the behavioral reaction depending on the relative attractiveness of welfare and work. Among the more educated, and those living in lower welfare benefits States, the ability to commit leads to significantly more work; among those with less education and living in high benefits States commitment generates either little, or negative changes in work behavior.

Importantly, the results of Table 8 also indicate that while the behavioral changes produced by an ability to commit may be large, the utility effects are invariably modest. The change in lifetime utility as a result of commitment ranges from \$1737 (a 5.03 percent increase) for those in cell 7 with the highest levels of education and medium welfare benefits, to \$1525 (a 5.33 percent increase) for those in cell 5 with medium levels of education and low welfare benefits, and to \$1092 (a 2.37 percent increase) for among those in cell 1 with very low levels of education and welfare benefits.

It may seem puzzling that the behavioral effects of commitment could be large while the utility gains among the same group are relatively small. This result derives from two mechanisms. First, for those delaying welfare takeup in favor of home, the delay in the absence of commitment is fairly short – typically less than two years. Thus the cumulative gains are relatively modest. Second, for those delaying entry into the labor force, the delay is typically longer, but the gains are realized only in the relatively distant future. So while it may be optimal from the perspective of the period a_0 self to commit herself to a career of work, the gains from that decision (relative to the decisions made in the absence of commitment) will be realized only after substantial work experience has accumulated, and will thus discounted by time. The costs required in order to acquire that work experience are, on the other hand, realized in the relatively near term, and thus discounted less by time. As a result, from the perspective of the period a_0 self, the net gains from commitment may be relatively small even when the behavioral consequences are substantial. If, however, we evaluate the change in utility from the perspective of the agent in her late 20s, the utility gains from commitment can be as high as 11 percent of continuation utility.

6.2 Consequences of Time Limits

The experiment of the previous section sets an upper bound on the utility gains from commitment. While we know that imperfect commitment devices such as time limits and workfare can at most deliver some fraction of these gains, it is not clear how they would influence work decisions. Table 9 presents the results of simulation exercises when we impose welfare eligibility time limits of varying lengths. Again we consider the behavioral and utility consequences for individuals with different initial conditions.

[**Table 9 About Here**]

While each of the time limits increases the frequency of work, in doing so they almost always reduce the lifetime utility of individuals in the model. Regardless of the limit's length, the predicted increases in work and decreases in utility are most dramatic for those with little education living in high benefits States (see, e.g., cells 2 and 4). The model implies that time limits are too crude a commitment device. They fail to induce more work while increasing expected lifetime utility. Though, for those living in low benefits States, and for those with higher levels of skills and education, the utility losses from the actual five-year time limit are quite modest (see cells 1, 3, 5 and 7). Thus, in these lower benefits States, the model suggests that if the policy goal is to promote work while limiting the utility consequences to the welfare eligible, a five-year time limit is a reasonable tool. In higher benefits States and among those with low human capital, however, the estimated utility consequences are relatively severe.

From the perspective of the period a_0 self, the preferred length of the time limit depends somewhat on education level and type. Among those with less education (cell 3), longer limits induce less work but generate more utility. Among those with more education (cells 5, 7, and 8) the longest limit is most preferred, but among the shorter limits, the shorter the better. Indeed, among these groups, eliminating welfare is preferred to a year-long time limit; and in particular, cell 5 and 7 groups may strictly prefer the elimination of the welfare system.

6.3 Consequences of Workfare

Table 10 presents the results of a parallel analysis with workfare policies. In these experiments, two dimensions of the policy are varied: (1) the degree to which workfare contributes to human capital, and (2) the extent to which workfare compensates for lost home production.

[**Table 10 About Here**]

Policy version 1 assumes that workfare is merely "make work" – participation in the program adds nothing to human capital. In this version of the policy, home production is compensated by 50% while on workfare through, for example, a child care subsidy. (With each policy the stigma of welfare participation is assumed to apply.) **Policy version 2** assumes workfare approximates market work – participation in workfare contributes to work experience just as labor market work would.[29] Again home production is compensated by 50%. Last, **policy version 3** replicates the human capital structure of policy version 2, but increases home production compensation to 75%.

For the first version of the workfare policy, in which the work requirement adds nothing to human capital while reducing home production by half, the model predicts substantial increases in market work. Among those with less schooling, the increases in time spent in the labor market are somewhat larger for those in low benefit States (see, cells 1 and 3). Among those with more schooling, the opposite holds: make-work policies lead to the largest increases in market work for those living in higher benefit States (see cells 6 and 8). Regardless of education or welfare benefits level, however, this first workfare policy reduces expected lifetime utility, though for those with greater human capital living in low benefits States, the declines are quite modest.

The predicted effects of workfare can be qualitatively different, however, when the work required adds to human capital (policy versions 2 and 3). When workfare provides the opportunity to accumulate human capital, there are two countervailing effects on decision making. On one hand, the access to human capital at a guaranteed "wage" makes welfare a relatively attractive choice. On the other hand, the accumulation of human capital while receiving welfare will make a transition into market work more appealing. The simulations indicate that the dominating effect vary with the initial conditions of the agents. Relative to the "make-work" policy, the second version of the policy leads to greater increases in market work among welfare-eligibles with relatively low human capital in high benefit States (cells 2, 4 and 6). For those with higher human capital, and/or living in low benefits States (cells 1, 3, 5, 7 and 8) are the employment gains smaller with this second policy.

When home production is compensated by half (policy version 2), the utility effects of the policy experiment are somewhat mixed. Among those with more human capital living in low benefits States, the commitment effect of the policy combined with the ability to accumulate human capital while on welfare leads to modest increases in expected lifetime utility. But those with relatively low human capital living in high welfare benefit States (cells 2 4 and 6), lifetime expected utility declines, though the declines are quite modest. When home production is compensated by 75% (policy version 3), the model predicts, more uniformly, lifetime utility gains. These gains are arguably

[29]The decay of human capital still occurs when an individual leaves market work for workfare.

modest, but mostly derive from increases in employment of a size comparable to those derived from make-work workfare. Thus, these simulations indicate that sizeable increases in employment among the welfare eligible can be achieved at relatively low utility cost (or indeed with utility gains) from workfare that both generates marketable human capital and substantially compensates for lost home production.[30]

7 Conclusions

Estimates of the structural parameters of a dynamic model of labor supply indicate that the work-welfare-home decisions of never-married women with children are well described by a model of time-inconsistent preferences. With reasonable precision, we estimate a present-bias factor (β) less than unity; and the model matches many aspects of observed decision and wage profiles. We reject a model of standard discounting at standard levels of confidence.

Simulation exercises using the model's estimated parameters indicate that while the behavioral consequences of an inability to commit to future decisions may be substantial, by one measure the utility consequences of the self-control problem are modest. The model suggests that the ability to commit to future decisions would often lead to considerably more work and less welfare participation. However, for those with low levels of human capital, and living in high welfare benefits States, procrastination leads to costly delays in welfare takeup. For this group, commitment ability leads to slightly more, not less, welfare participation. Moreover, among those entering the labor force earlier, this entry involves costs in terms of welfare benefits and home production forgone; and the benefits in terms of higher wages are accrued only in the relatively distant future. As a result, the discounted lifetime utility gains from commitment may be small even when the behavioral consequences are large.

Further simulations of the model indicate that behavioral and utility consequences of welfare reform policies that serve as imperfect commitment devices vary according to both the characteristics of the intended targets and the design of the policy. Simulations suggest that time limits are too crude to enhance expected utility. While limits serve to substantially increase employment, they do so at a sometimes substantial utility cost for the welfare-eligible. For those living in low benefits States, and for those with higher levels of skills and education, however, the utility losses from a five-year time limit are quite modest. The estimated model indicates that workfare policies also better serve those with more education living in States with lower welfare benefits. However, when workfare leads to the accumulation of valuable human capital, and includes compensation

[30] The gains would be more subtantial if, as is plausible, the policy also reduced the stigma of welfare participation.

for lost home production through, for example, child care subsidies, the estimated model suggests that most potential recipients will increase both their employment and their lifetime utility.

We interpret these results as qualified support for the extension of standard models of dynamic labor supply to allow for time-inconsistency. Our findings with this model indicate that allowing for time-inconsistency may be both feasible and fruitful, adding to our understanding of the potential consequences of policy. We also view our findings as a caution against simple arguments for accounting for the role of psychological biases in public policy. As our simulations indicate, even when individuals display substantial present-bias in preferences, simple policies that resemble commitment devices may not function effectively as such.

References

[1] Ainslie, George (1992). *Picoeconomics: The Strategic Interaction of Successive Motivational States Within the Person*. Cambridge [England]; New York: Cambridge University Press.

[2] Barro, Robert J. (1999). "Ramsey Meets Laibson in the Neoclassical Growth Model." *Quarterly Journal of Economics* **114(4):** 1125-1152.

[3] Berndt, E., B. Hall, R. Hall, and J. Hausman (1974). "Estimation and Inference in Nonlinear Structural Models." *Annals of Economic and Social Measurement.* **3/4**:653-665.

[4] Caplin, Andrew and John Leahy (2000). "The Social Discount Rate." *NBER Working Paper* 7983.

[5] Carrillo, J. and T. Mariotti (2000). "Strategic Ignorance as a Self-Disciplining Device." *Review of Economic Studies* **67(3)**:529-544.

[6] Della Vigna, Stefano, and Ulrike Malmendier (2002). "Overestimating Self-Control: Evidence from the Health Club Industry." Mimeo University of California.

[7] Della Vigna, Stefano, and M. Daniele Paserman (2001). "Job Search and Impatience." Mimeo University of California.

[8] Edin, Kathryn and Laura Lein (1997). *Making Ends Meet: How Single Mothers Survive Welfare and Low-wage Work*. New York: Russell Sage.

[9] Eckstein, Zvi and Kenneth I. Wolpin (1999). "Why Youth Drop out of High School: The Impact of Preferences, Opportunities and Abilities." *Econometrica,* **67(6)**, 1295-1339.

[10] Fang, Hanming, and Dan Silverman (2004a). "On the Compassion of Time-limited Welfare Programs." *Journal of Public Economics*, **88(2):** 1445-1470.

[11] ——— and ——— (2004b). "Distinguishing Between Cognitive Biases: Belief vs. Time Discounting in Welfare Program Participation. Mimeo, Yale Univeristy and University of Michigan.

[12] Gruber, Jonathan, and Botond Koszegi (2001). "Is Addiction 'Rational'? Theory and Evidence." *Quarterly Journal of Economics* **116(4)**:1261-1303.

[13] Harris, Christopher, and David Laibson (2001). "Dynamic Choices of Hyperbolic consumers." *Econometrica* **69(4)**: 935-957.

[14] Hausman, Jerry A. (1979) "Individual Discount Rates and the Purchase and Utilization of Energy-using Durables." *Bell Journal of Economics* **10(1):**33-54.

[15] Krusell, Per, Burhanettin Kuruşçu, and Anthony Smith, Jr. (2002). "Equilibrium Welfare and Government Policy with Quasi-Geometric Discounting." *Journal of Economic Theory* **105(1)**:42-72.

[16] Keane, Michael P., and Kenneth I. Wolpin, (1994). "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence." *The Review of Economics and Statistics* **76(4):**648-672.

[17] ——— and ——— (1997). "The Career Decisions of Young Men." *Journal of Political Economy* **105(3):**473-522.

[18] ——— and ——— (2000). "Estimating the Effect of Welfare on the Education, Employment, Fertility and Marriage Decisions of Women." Mimeo, New York University and University of Pennsylvania.

[19] Laibson, David (1997). "Golden Eggs and Hyperbolic Discounting." *Quarterly Journal of Economics* **112:** 443-477.

[20] Laibson, David, Andrea Repetto and Jeremy Tobacman, (1998). "Self-Control and Saving for Retirement." *Brookings Papers on Economic Activity* **1:**91-196.

[21] ———, ———, and ——— (2004). Estimating Discount Functions from Lifecycle Consumption Choices. Mimeo, Department of Economics, Harvard University.

[22] Loewenstein, George, and Jon Elster (1992). *Choice Over Time.* Russell Sage: New York.

[23] Miller, Robert A. and Seth G. Sanders (1997). "Human Capital Development and Welfare Participation." *Carnegie-Rochester Conference Series on Public Policy*, **46**: 1-43.

[24] Moffitt, Robert (1983). "An Economic Model of Welfare Stigma." *American Economic Review* **73(5)**:1023-35.

[25] Mulligan, Casey B. (1997). "A Logical Economist's Argument Against Hyperbolic Discounting." Mimeo University of Chicago.

[26] O'Donoghue, Ted and Matthew Rabin (1999a). "Doing It Now or Later." *American Economic Review* **89(1)**: 103-124.

[27] ——— and ——— (1999b). "Addiction and Self-Control." In Jon Elster, editor, *Addiction: Entries and Exits*. New York: Russell Sage.

[28] Paserman, M. Daniele (2002). "Job Search and Hyperbolic Discounting: Structural Estimation and Policy Evaluation." Mimeo Hebrew University.

[29] Phelps, Edmund S., and Robert A. Pollak (1968). "On Second-best National Saving and Game-equilibrium Growth." *Review of Economic Studies* **35(2)**: 185-199.

[30] Pollak, Robert A. (1968). "Consistent Planning." *Review of Economic Studies* **35(2)**: 201-208.

[31] Rubinstein, Ariel (2003). "Economics and Psychology"?: The Case of Hyperbolic Discounting." Forthcoming, *International Economic Review*.

[32] Strotz, Robert H. (1956). "Myopia and Inconsistency in Dynamic Utility Maximization." *Review of Economic Studies* **23(3):** 165-180.

[33] Swann, Christopher (1996). "Welfare Reforms when Agents are Forward-looking." Mimeo, University of Virginia.

[34] Thaler, Richard H. (1991). *Quasi Rational Economics.* New York: Russell Sage.

[35] Warner, John T., and Saul Pleeter (2001). "The Personal Discount Rate: Evidence from Military Downsizing Programs." *American Economic Review* **91(1):**33-53.

Discount Factors		Payoffs from Work	Payoffs from Welfare
β	δ	$w(0) + \beta \sum_{x=1}^{2} \delta^{x} w(x)$	$b(0) + \beta \sum_{\kappa=1}^{2} \delta^{\kappa} b(\kappa)$
0.80	0.95	0.99	0.96
0.80	1.00	1.30	1.20
0.90	0.95	1.30	1.33

Table 1: Qualitative Differences in Behavioral Effects of Changes in β and δ

Panel 1: Average Percent of Time Observed Choosing Welfare with $\beta = 0.34$				
$\delta = 0.84$	$\delta = 0.86$	$\delta = 0.88$	$\delta = 0.90$	$\delta = 0.92$
47.73	46.92	45.62	43.61	40.27
Panel 2: Average Percent of Time Observed Choosing Welfare with $\delta = 0.88$				
$\beta = 0.20$	$\beta = 0.27$	$\beta = 0.34$	$\beta = 0.41$	$\beta = 0.48$
43.38	44.74	45.62	46.46	47.23

Table 2: Simulated Decisions for Varying Levels of the Discount Parameters, Ages 18-32.

	Welfare		Work		Home		Total	
Age	Percent	Number	Percent	Number	Percent	Number	Percent	Number
16	31.9	15	0.0	0	68.1	32	100.0	47
17	38.2	34	0.0	0	61.8	55	100.0	89
18	38.5	60	1.9	3	59.6	93	100.0	156
19	46.6	109	8.6	20	44.9	105	100.0	234
20	50.2	143	11.9	34	37.9	108	100.0	285
21	50.5	165	14.1	46	35.5	116	100.0	327
22	53.7	188	16.9	59	29.4	103	100.0	350
23	51.2	191	20.6	77	28.2	105	100.0	373
24	48.5	182	25.6	96	25.9	97	100.0	375
25	47.3	187	27.1	107	25.6	101	100.0	395
26	48.6	196	30.5	123	20.8	84	100.0	403
27	44.3	167	32.1	121	23.6	89	100.0	377
28	45.1	142	33.0	104	21.9	69	100.0	315
29	42.8	109	37.7	96	19.6	50	100.0	255
30	47.9	91	35.3	67	16.8	32	100.0	190
31	43.1	62	39.6	57	17.4	25	100.0	144
32	35.6	32	42.2	38	22.2	20	100.0	90
Total	47.1	2073	23.8	1048	29.1	1284	100.0	4405

Table 3: Choice Distribution, Ages 16-32, Never-married Women with at Least One Child.

	Choice at t		
Choice at $t-1$	Welfare	Work	Home
Welfare			
Row %	84.3	3.5	12.3
Column %	76.7	6.3	17.9
Work			
Row %	5.3	79.3	15.3
Column %	2.6	76.4	12.1
Home			
Row %	28.3	12.0	59.7
Column %	20.7	17.3	70.0

Table 4: Transition Matrix, Never-married Women with at Least One Child.

Parameters		(1) Time Consistent Estimate	(1) S.E.	(2) Present-Biased (sophisticated) Estimate	(2) S.E.	(3) Present-Biased (Naive) Estimate	(3) S.E.
Preference Parameters							
Discount Factors	β	1	n.a.	0.33802	0.06943	0.355	0.0983
	δ	0.41488	0.07693	0.87507	0.01603	0.868	0.02471
Net Stigma	$\phi^{(1)}$	7537.04	774.81	8126.19	834.011	8277.46	950.77
(by type)	$\phi^{(2)}$	10100.9	1064.83	10242.01	955.878	10350.20	1185.27
	$\phi^{(3)}$	13333.2	1640.18	12697.25	1426.40	12533.69	1685.92
Home Production	$e_0^{(1)}$	2684.97	427.85	2209.48	405.26	2224.98	456.85
(by type)	$e_0^{(2)}$	3324.79	516.96	3502.66	509.07	3492.15	617.64
	$e_0^{(3)}$	1729.53	1418.21	2126.86	879.54	2182.17	1227.66
	e_1	84.83	441.45	124.92	48.95	121.58	130.57
	e_2	-36.21	105.61	-603.29	215.67	-608.39	560.31
	$\eta^{(1)}$	2484.69	494.09	4565.06	399.07	4588.88	756.19
	$\eta^{(2)}$	4432.11	573.40	6547.94	503.62	6557.07	933.40
	$\eta^{(3)}$	9858.23	1290.18	12149.5	869.089	12054.63	1670.74
Wage and Skill Parameters							
Constant	$h_0^{(1)}$	0.12881	0.09963	0.16329	0.0676	0.1672	0.1362
(by type)	$h_0^{(2)}$	0.59176	0.10073	0.6121	0.06828	0.61628	0.13625
	$h_0^{(3)}$	1.11547	0.12045	1.10907	0.08089	1.12299	0.14646
years of schooling	α_1	0.01995	0.0082	0.02153	0.00501	0.02166	0.00976
experience	α_2	0.13513	0.01056	0.12252	0.00853	0.12142	0.01203
experience^2	α_3	-0.00736	0.0009	-0.00623	0.00068	-0.00605	0.00099
1^{st} year experience	α_4	0.09352	0.04291	0.06681	0.02949	0.06742	0.04535
experience decay	α_5	-0.22702	0.03601	-0.23105	0.03096	-0.23694	0.03731

Table 5: Parameter Estimates for Time Consistent, Sophisticated and Naive Present-biased Agents.

Parameters		(1) Time Consistent Estimate	(1) Time Consistent S.E.	(2) Present-Biased (sophisticated) Estimate	(2) Present-Biased (sophisticated) S.E.	(3) Present-Biased (Naive) Estimate	(3) Present-Biased (Naive) S.E.
Continuation Value Function at Age 35							
num. of children	ω_1	794.52	743350.2	2618.55	3511.39	2496.75	197163.19
num. of children^2	ω_2	-8938.74	82101.40	-8918.7	5258.05	-8638.95	27929.24
experience	ω_3	62.74	20429.20	235.24	268.94	231.11	4500.37
experience^2	ω_4	-54.47	516.04	378.36	115.00	374.21	185.64
welfare lag	ω_5	2617.59	7515.73	8707.61	6322.23	8725.00	10638.20
work lag	ω_6	1544.06	13820.09	6151.05	4142.20	6260.41	14140.67
Log Odds as Function of Initial Conditions for Types 2 and 3							
Type 2: constant	$\pi_0^{(2)}$	-1.842	1.544	-1.070	1.550	-1.179	1.593
age	$\pi_1^{(2)}$	0.0067	0.087	-0.0406	0.086	-0.0385	0.0867
yrs. of schooling	$\pi_2^{(2)}$	0.129	0.124	0.133	0.122	0.139	0.127
experience	$\pi_3^{(2)}$	0.217	0.194	0.227	0.190	0.221	0.187
welfare lag	$\pi_4^{(2)}$	0.0865	0.662	0.398	0.618	0.406	0.633
work lag	$\pi_5^{(2)}$	0.0131	0.578	0.062	0.587	0.0576	0.578
Type 3: constant	$\pi_0^{(3)}$	-3.948	2.423	-5.627	2.328	-5.562	2.273
age	$\pi_1^{(3)}$	-0.687	0.126	-0.360	0.168	-0.356	0.167
yrs. of schooling	$\pi_2^{(3)}$	1.303	0.156	0.9322	0.268	0.918	0.263
experience	$\pi_3^{(3)}$	0.1055	0.2811	0.314	0.278	0.318	0.277
welfare lag	$\pi_4^{(3)}$	-0.526	1.252	-0.640	1.508	-0.463	1.305
work lag	$\pi_5^{(3)}$	0.575	0.881	-0.13387	0.874	-0.144	0.846
Variance and Covariance of Shocks							
std. dev. of ϵ_0	σ_{ϵ_0}	5262.40	548.55	5656.61	446.56	5708.50	579.19
std. dev. of ϵ_1	σ_{ϵ_1}	0.3751	0.0122	0.3726	0.0071	0.3707	0.0076
std. dev. of ϵ_2	σ_{ϵ_2}	4168.06	334.76	4116.96	331.49	4074.99	459.82
$cov\,(\epsilon_0, \epsilon_2)$	$\sigma^2_{\epsilon_0\epsilon_2}$	-3046.77	168.32	-2849.19	202.06	-2861.02	247.60
Log-likelihood		-3505.96		-3489.80		-3486.44	
χ^2-Statistics		32.32		n.a.		6.72	

Table 6: Parameter Estimates for Time Consistent, Sophisticated and Naive Present Agents (continued from Table 5). NOTE: χ^2 statistics are calculated under the null hypothesis of the present-biased sophisticated model.

Choice at $t-1$	Choice at t		
	Welfare	Work	Home
Welfare			
Row %	84.4	4.2	11.4
Column %	78.4	7.4	19.8
Work			
Row %	10.9	74.2	14.9
Column %	5.6	72.7	14.5
Home			
Row %	25.9	17.1	57.0
Column %	15.9	19.9	65.7

Table 7: Simulated Transition Probability Matrix for Sohpisticated Present-Biased Agents.

	Initial Conditions Cell							
	1	2	3	4	5	6	7	8
PANEL 1: SIMULATED EFFECTS								
changes in % working	14.07	-2.74	23.32	-3.38	24.44	-1.71	24.69	13.07
% change in lifetime utility	3.41	2.42	4.78	2.37	5.33	2.47	5.03	3.26
PANEL 2: INITIAL CONDITIONS FOR DIFFERENT CELLS								
wel. benefits (1 child)	4126.53	7103.51	4103.53	7278.74	4073.25	7116.39	4278.39	7023.98
wel. benefits (2 children)	5383.13	8781.58	5340.66	8969.49	5315.70	8809.24	5529.23	8746.56
age at first birth	17	18	19	19	21	20	22	22
years of schooling	9	9	11	11	12	12	14	14
work yr. befor first birth	No	No	No	No	Yes	No	No	Yes
yrs of work exper. at first birth	0	0	0	0	1	0	1	1
Prob (type = 1)	0.622	0.636	0.556	0.556	0.469	0.513	0.335	0.339
prob (type = 2)	0.356	0.349	0.383	0.383	0.454	0.387	0.383	0.412

Table 8: Simulated Effects of the Ability to Commit for Sophisticated Present-Biased Agents, by Initial Conditions.

Time Limits		Initial Conditions Cell							
		1	2	3	4	5	6	7	8
7 years	% change in lifetime util.	−1.86	−9.70	−1.10	−8.10	−0.15	−6.63	−0.14	−1.70
	changes in % working	11.23	22.28	7.44	20.22	2.41	17.80	1.77	9.69
5 years	% change in lifetime util.	−2.52	−13.05	−1.36	−10.83	−0.10	−9.07	−0.17	−2.36
	changes in % working	17.90	33.29	13.01	31.54	4.93	28.48	3.66	16.72
3 years	% change in lifetime util.	−3.53	−17.29	−1.99	−14.23	−0.20	−11.97	−0.23	−3.14
	changes in % working	26.02	43.66	19.74	42.38	8.71	40.73	6.63	24.84
1 year	% change in lifetime util.	−5.54	−22.93	−3.11	−18.96	−0.43	−15.76	−0.32	−4.20
	changes in % working	34.97	55.06	29.18	55.51	14.91	53.74	12.03	35.55
0 year	% change in lifetime util.	−5.55	−23.92	−2.92	−19.66	0.21	−16.12	0.16	−3.51
	changes in % working	40.22	61.03	34.50	61.42	19.40	60.33	15.50	41.90

Table 9: Simulated Effects of Time Limits of Varying Lengths for Sophisticated Present-Biased Agents, by Initial Conditions.

Policy		Initial Conditions Cell							
		1	2	3	4	5	6	7	8
Workfare version 1	% change in lifetime util.	−4.88	−14.74	−2.89	−12.90	−0.38	−11.26	−0.16	−3.15
	changes in % working	22.51	13.55	20.49	19.33	12.04	22.22	10.58	23.58
Workfare version 2	% change in lifetime util.	1.78	−2.33	2.68	−1.75	2.80	−0.77	1.96	2.58
	changes in % working	21.61	20.32	18.93	23.85	8.15	26.35	7.29	18.93
Workfare version 3	% change in lifetime util.	6.58	6.15	6.36	6.09	5.00	6.36	3.47	6.33
	changes in % working	13.85	13.00	12.88	16.11	2.39	18.84	2.62	9.46

Table 10: Simulated Effects of Workfare Policies for Sophisticated Present-Biased agents, by Initial Conditions.

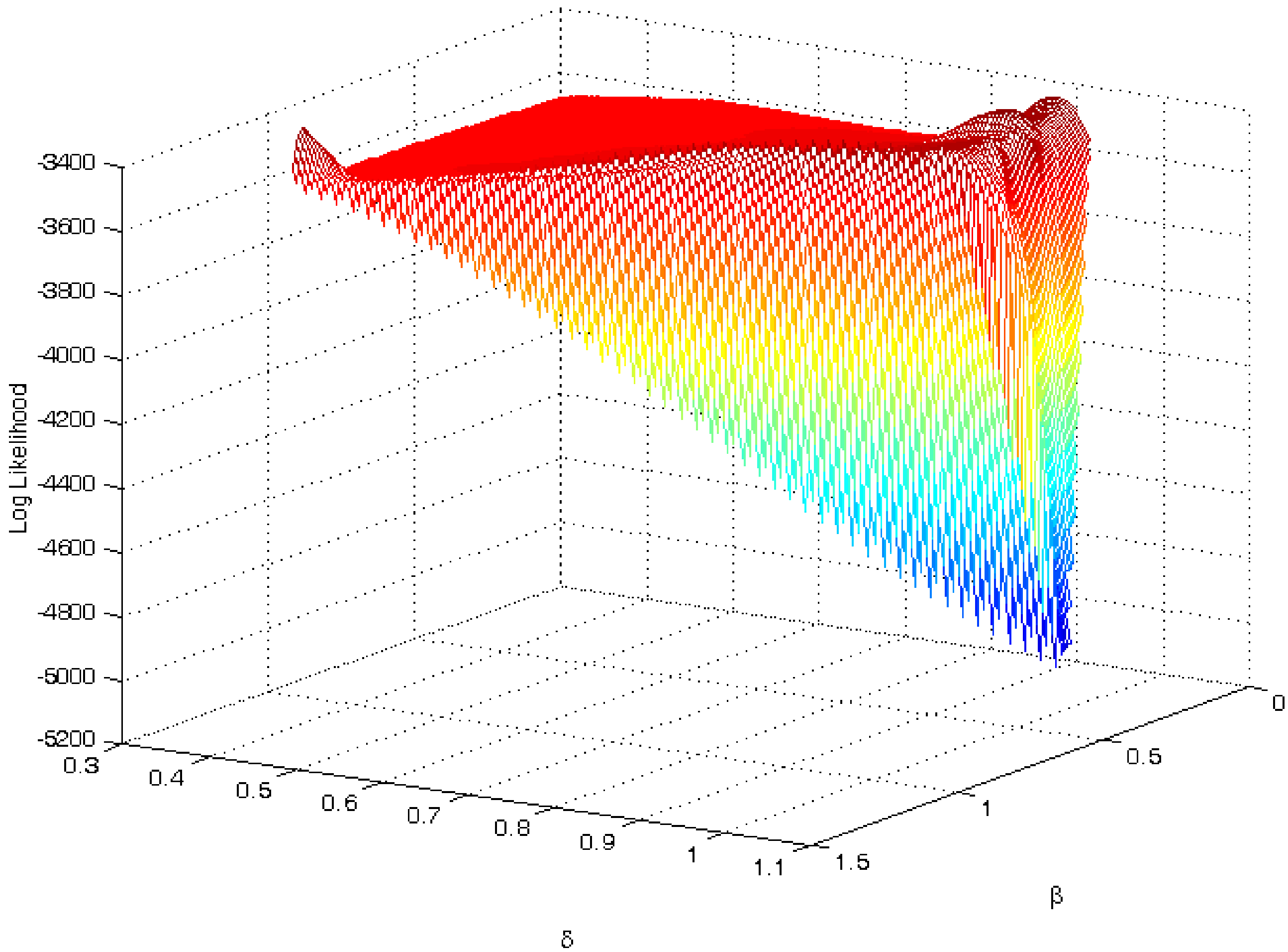

Figure 1: Surface of the Log-Likelihood as a Function of β and δ: An Illustration.

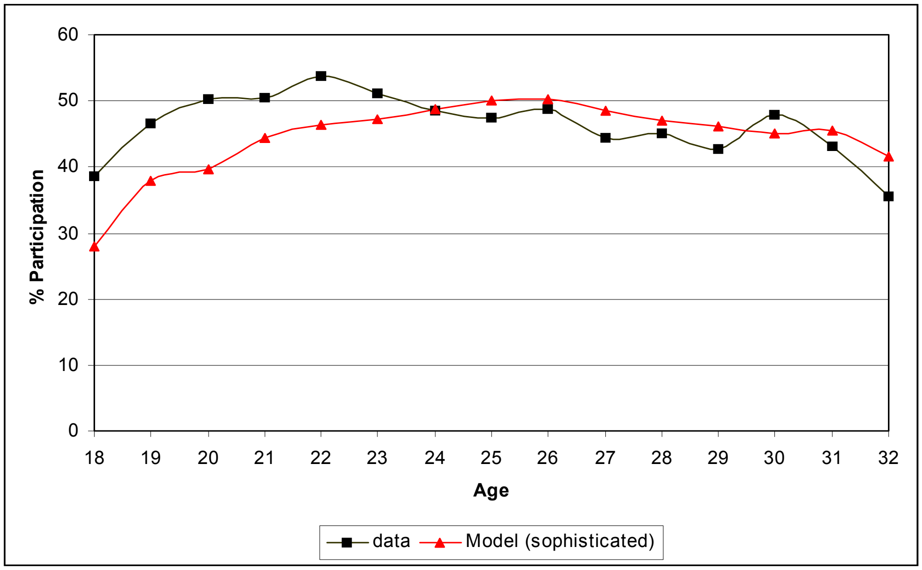

Figure 2: Age-Welfare Participation Profiles: Data vs. and Model Simulation for Sophisticated Agents.

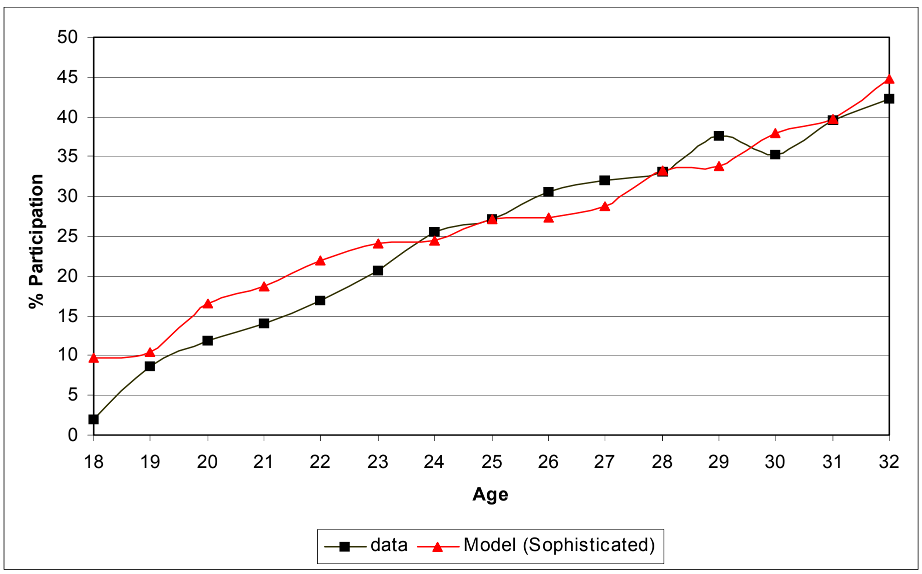

Figure 3: Age-Work Participation Profiles: Data vs. Model Simulation for Sophisticated Agents.

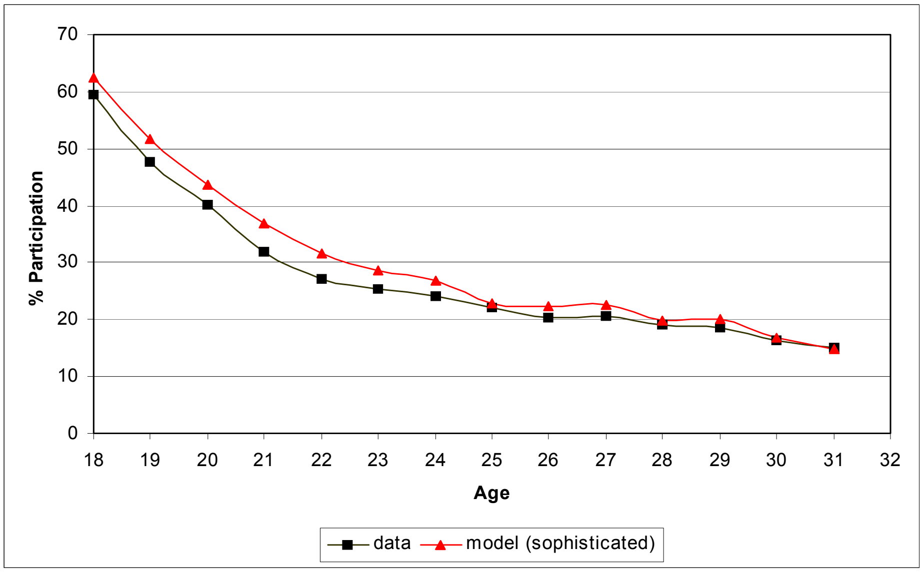

Figure 4: Age-Home Choice Profiles: Data vs. Model Simulation for Sophisticated Agents.

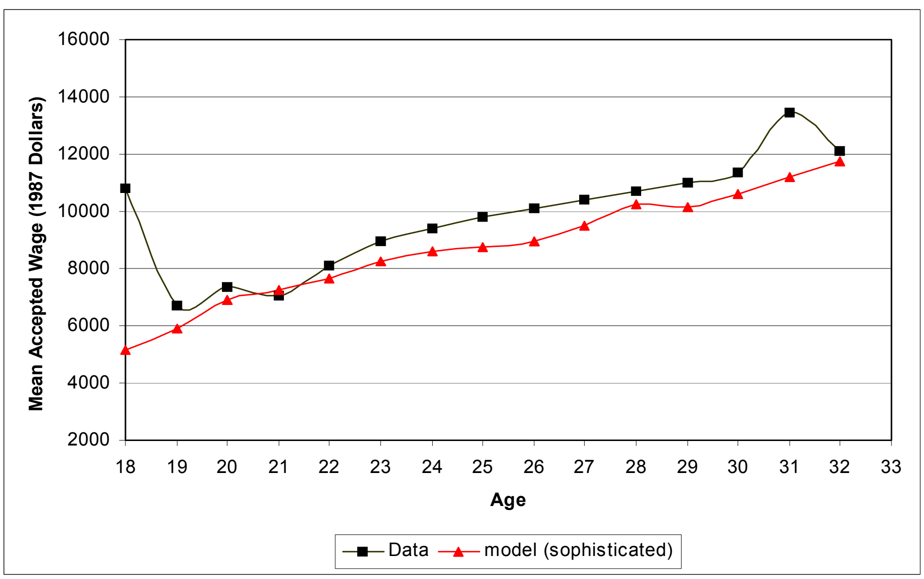

Figure 5: Mean Accepted Wages for Workers by Age: Data vs. Model Simulation for Sophisticated Agents.

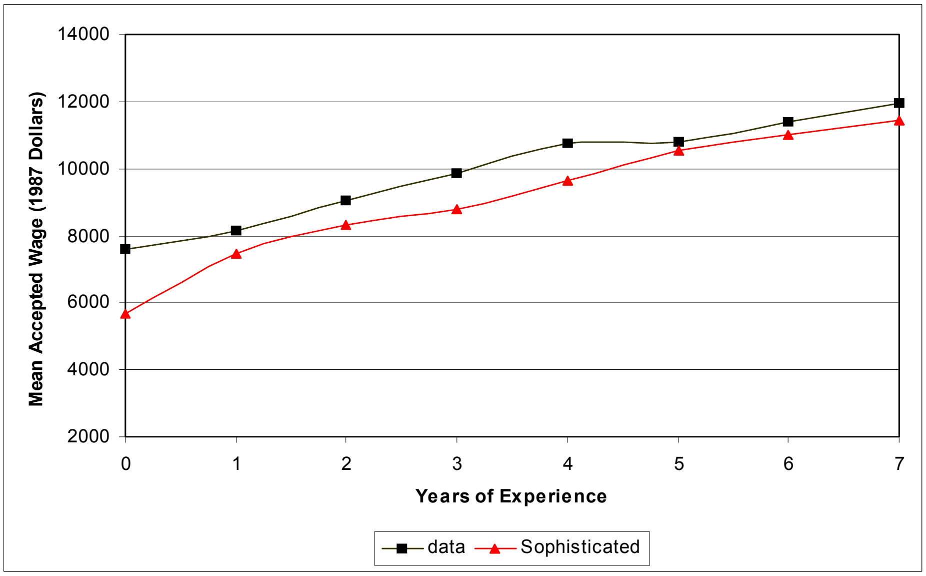

Figure 6: Mean Wage-Experience Profiles: Data vs. Model Simulation for Sophisticated Agents.

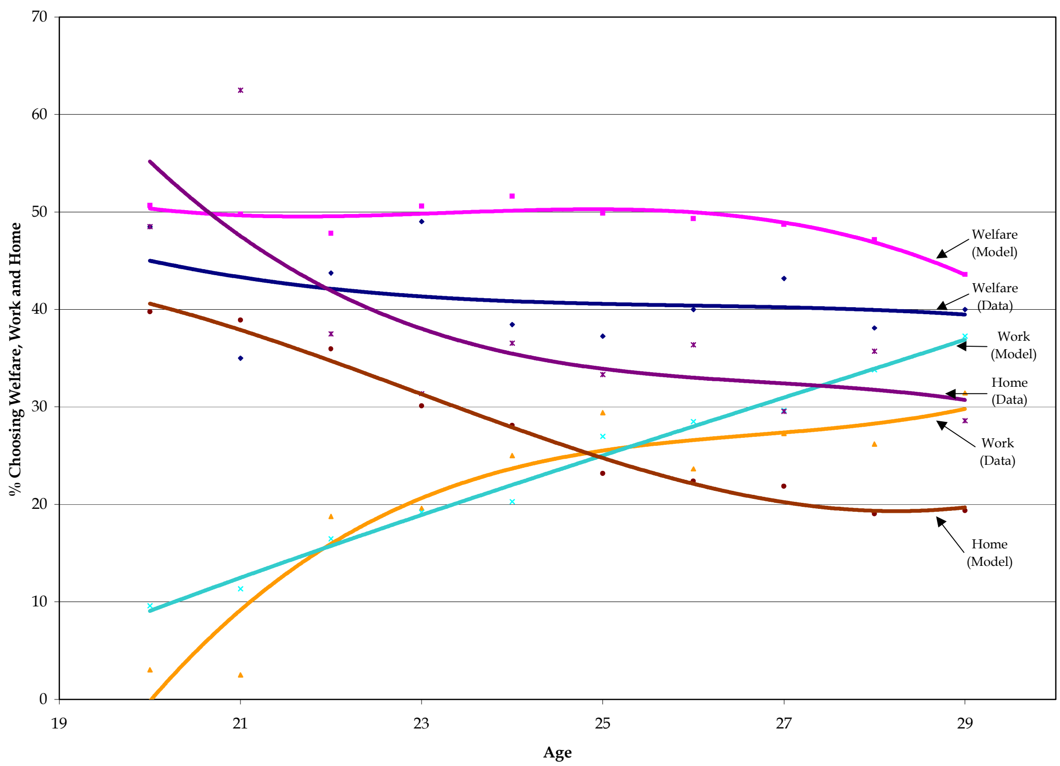

Figure 7: Age-Decision Profile: Comparison of Out-of-Sample Data and Simulation with Estimated Parameters (Sophisticated Agents).

A Appendix: Additional Tables and Estimates

Descriptive Statistics of the Selected Women and All Women Table A1 compares the statistics of our selected subsample (never married women with at least one child) with those of the entire sample of women in the NLSY. It shows that the subsample has on average more children at every age. By age 32, the gap is relatively small with the subsample having on average 2.1 children, and the entire sample 1.6. At every age the subsample has an average of 1.25 fewer years of work experience, and 2.01 more years of AFDC receipt; and at every age older than 19, full-time workers in the subsample earn on average $1,456 less than their counterparts in the entire sample. On average, the subsample has also completed fewer years of schooling (10.9) than the entire sample (12.6).

Estimates of Welfare Benefit Function G_j Table A2 presents the estimates of welfare benefit functions for the twenty States used in our estimation.

Fertility Function ρ Table A3 presents the estimates of the fertility function ρ.

Within-Sample Goodness of Fit Test Table A4 presents the χ^2 goodness-of-fit test of the within-sample choice distributions by age. The column labelled by "Row" is the χ^2 statistic for the overall choice distribution for the particular age in that row.

	Number of Children		Yrs. of Work Experience		Yrs. of Schooling*		Earnings for Workers**		Yrs. Received AFDC	
Age	All Women	Our Sample	All Women	Our Sample	All Women	Our Sample	All Women	Our Sample	All Women	Our Sample
16	0.06 (0.01)	1.23 (0.08)	0.00 (0.00)	0.00 (0.00)	9.38 (0.02)	8.81 (0.17)	5410.34 (617.97)	n.a.	0.00 (0.00)	0.13 (0.06)
17	0.10 (0.01)	1.24 (0.06)	0.03 (0.00)	0.00 (0.00)	10.28 (0.02)	9.37 (0.14)	5808.70 (166.49)	n.a.	0.01 (0.00)	0.24 (0.06)
18	0.17 (0.01)	1.28 (0.04)	0.10 (0.01)	0.00 (0.00)	11.14 (0.02)	10.04 (0.12)	7001.72 (167.58)	10822.56 (2254.07)	0.02 (0.00)	0.34 (0.06)
19	0.25 (0.01)	1.36 (0.04)	0.26 (0.01)	0.05 (0.01)	11.75 (0.02)	10.41 (0.10)	7723.65 (122.65)	6715.04 (766.20)	0.04 (0.00)	0.49 (0.06)
20	0.33 (0.01)	1.43 (0.04)	0.54 (0.01)	0.14 (0.03)	12.19 (0.02)	10.63 (0.09)	8301.49 (102.97)	7361.80 (623.75)	0.07 (0.01)	0.78 (0.07)
21	0.44 (0.01)	1.49 (0.04)	0.87 (0.02)	0.25 (0.03)	12.48 (0.02)	10.77 (0.09)	8819.52 (111.21)	7040.99 (594.02)	0.10 (0.01)	1.09 (0.08)
22	0.54 (0.01)	1.59 (0.05)	1.26 (0.02)	0.43 (0.05)	12.71 (0.03)	10.86 (0.08)	9676.16 (110.01)	8097.39 (505.02)	0.15 (0.01)	1.45 (0.09)
23	0.66 (0.01)	1.70 (0.05)	1.72 (0.03)	0.69 (0.07)	12.87 (0.03)	10.92 (0.08)	10405.64 (107.29)	8929.34 (374.76)	0.21 (0.01)	1.86 (0.10)
24	0.77 (0.01)	1.79 (0.05)	2.24 (0.03)	0.95 (0.08)	12.95 (0.03)	10.96 (0.09)	11086.97 (113.59)	9376.78 (394.42)	0.27 (0.01)	2.30 (0.12)
25	0.89 (0.02)	1.84 (0.05)	2.79 (0.03)	1.34 (0.10)	13.01 (0.03)	11.02 (0.09)	11719.11 (129.55)	9787.95 (411.93)	0.32 (0.02)	2.67 (0.13)
26	1.01 (0.02)	1.90 (0.05)	3.34 (0.04)	1.68 (0.12)	13.07 (0.03)	11.06 (0.09)	12280.13 (136.35)	10099.91 (381.09)	0.38 (0.02)	3.02 (0.14)
27	1.12 (0.02)	1.94 (0.05)	3.85 (0.04)	2.05 (0.14)	13.14 (0.03)	11.10 (0.09)	12674.95 (165.59)	10392.56 (393.77)	0.43 (0.02)	3.38 (0.16)
28	1.23 (0.02)	1.95 (0.06)	4.39 (0.05)	2.44 (0.17)	13.18 (0.04)	11.22 (0.10)	13379.60 (220.25)	10692.78 (448.43)	0.47 (0.02)	3.61 (0.19)
29	1.33 (0.02)	1.99 (0.07)	4.91 (0.06)	2.66 (0.21)	13.20 (0.04)	11.37 (0.11)	13651.76 (278.03)	11004.52 (497.69)	0.50 (0.03)	3.89 (0.23)
30	1.41 (0.02)	2.08 (0.08)	5.48 (0.08)	2.89 (0.26)	13.24 (0.04)	11.44 (0.13)	13531.17 (293.62)	11360.18 (632.93)	0.52 (0.03)	4.48 (0.29)
31	1.54 (0.03)	2.11 (0.10)	5.90 (0.10)	3.22 (0.32)	13.22 (0.05)	11.56 (0.15)	13614.89 (364.01)	13455.48 (1091.30)	0.55 (0.04)	4.67 (0.33)
32	1.63 (0.03)	2.08 (0.13)	6.40 (0.12)	3.99 (0.45)	13.24 (0.05)	11.70 (0.21)	14301.09 (827.82)	12091.61 (871.85)	0.58 (0.04)	4.49 (0.43)

Notes: Standard errors in parentheses. Means are calculated using the NLSY's 1979 sample weights. Members of te poor white and military oversamples excluded.

*: years of schooling at the birth of the first child. **: earnings are full-time equivalent in 1987 dollars.

Table A1: Descriptive Statistics for All Women and Selected Sample (Never Married Women with At Least One Child): Ages 16-32.

States*	$\hat{\theta}_{j0}$	$\hat{\theta}_{j1}$	annual benefit for 1 child	annual benefit for 2 children	Percent on Welfare**
1	2380.45	1238.01	3618.46	4856.48	39.6
2	2467.68	1301.31	3768.99	5070.30	50.0
3	2962.66	1203.84	4166.50	5370.34	32.9
4	2979.62	1280.44	4260.06	5540.50	22.5
5	3128.33	1340.02	4468.35	5808.38	39.2
6	3493.63	1186.81	4680.45	5867.26	29.6
7	3541.08	1251.03	4792.11	6043.13	50.3
8	3985.20	1212.98	5198.18	6411.15	46.6
9	4348.62	1098.98	5447.60	6546.58	28.2
10	4358.47	1318.76	5677.23	6995.99	71.0
11	4279.58	1419.96	5699.54	7119.50	51.2
12	4509.59	1368.62	5878.21	7246.83	29.4
13	4183.05	1539.27	5722.32	7261.59	13.6
14	4592.94	1343.95	5936.89	7280.83	20.2
15	4511.30	1411.63	5922.93	7334.57	66.8
16	5005.98	1480.68	6486.65	7967.33	52.5
17	4988.00	1577.07	6565.07	8142.15	27.0
18	5634.63	1661.86	7296.49	8958.35	61.7
19	5317.42	1851.81	7169.23	9021.04	69.7
20	6264.03	1613.01	7877.04	9490.05	68.5
Mean	4146.61	1385.00	5531.62	6916.62	43.5
Std. Dev.	1042.30	187.46	1182.51	1334.38	17.9

*: States are left unnamed to maintain the anonyminity of survey respondents.

**: Percent of Sample living in the corresponding State that choose welfare

Table A2: Estimated Annual Welfare Benefits Function, Summary Statistics (1987 dollars)

Parameter	Estimate	Std. Error
γ_0	-0.811	0.323
γ_1	-0.044	0.015
γ_2	-0.077	0.059
γ_3	0.094	0.115
γ_4	-0.494	0.162
Observations:	3911	
Likihood Ratio:	38.20	
Log Likelihood:	-1287.16	
Pseudo R^2	0.014	

Table A3: Logit Estimates of the Fertility Function.

	Choice			
Age	Welfare	Work	Home	Row
18	6.09^*	†	0.19	6.28^*
19	4.56^*	0.79	2.05	7.40^*
20	7.82^*	3.76	2.18	13.76^*
21	2.63	3.78	0.16	6.56^*
22	3.98^*	4.19^*	0.50	8.68^*
23	1.28	1.94	0.03	3.25
24	0.00	0.19	0.14	0.32
25	0.61	0.00	1.32	1.93
26	0.23	1.46	0.38	2.07
27	1.39	1.38	0.17	2.93
28	0.25	0.00	0.72	0.97
29	0.64	1.06	0.02	1.72
30	0.32	0.36	0.00	0.68
31	0.19	0.00	0.66	0.85
32	0.77	0.12	4.63^*	5.52^*

$*$: Significant at the 5% level.

† : Fewer than 5 obervations.

Table A4: χ^2 Goodness-of-Fit Tests of the Within-Sample Choice Distribution By Age, Model with Sophisticated Agents.

B Appendix: Simulation Results for Completely Naive Present-Biased Agents

We present the simulation results for naive present-biased agents in this appendix.

	Choice at t		
Choice at $t-1$	Welfare	Work	Home
Welfare			
Row %	84.4	4.2	11.4
Column %	78.5	7.4	19.7
Work			
Row %	10.7	74.5	14.8
Column %	5.6	72.8	14.3
Home			
Row %	25.7	17.0	57.2
Column %	16.0	19.8	66.0

Table B5: Simulated Transition Probability Matrix for Naive Present-Biased Agents.

	Initial Conditions Cell							
	1	2	3	4	5	6	7	8
changes in % working	9.41	-2.79	17.95	-3.07	22.58	-2.16	23.80	11.30
% change in lifetime utility	2.82	2.39	4.23	2.43	4.95	2.47	5.00	3.17

Table B6: Simulated Effects of the Ability to Commit for Naive Present-Biased Agents, by Initial Conditions.

Time Limits		Initial Conditions Cell							
		1	2	3	4	5	6	7	8
7 years	% change in lifetime util.	−1.82	−9.25	−0.99	−7.62	−0.20	−6.24	−0.16	−1.69
	changes in % working	11.31	22.76	7.39	20.77	2.13	17.77	1.55	9.27
5 years	% change in lifetime util.	−2.63	−12.70	−1.37	−10.48	−0.21	−8.78	−0.21	−2.43
	changes in % working	17.60	33.45	12.51	31.82	4.60	28.69	3.25	16.29
3 years	% change in lifetime util.	−3.87	−17.17	−2.15	−14.10	−0.38	−11.89	−0.32	−3.36
	changes in % working	25.48	43.49	19.18	42.47	8.22	40.24	6.38	24.42
1 year	% change in lifetime util.	−6.25	−23.23	−3.62	−19.32	−0.77	−16.08	−0.47	−4.56
	changes in % working	34.03	54.34	28.20	54.78	14.13	52.74	11.69	35.06
0 year	% change in lifetime util.	−6.51	−24.60	−3.65	−20.30	−0.24	−16.80	0.00	−4.08
	changes in % working	39.27	60.10	33.14	60.67	18.83	59.13	15.43	40.89

Table B7: Simulated Effects of Time Limits of Varying Lengths for Naive Present-Biased Agents, by Initial Conditions.

Policy		Initial Conditions Cell							
		1	2	3	4	5	6	7	8
Workfare version 1	% change in lifetime util.	−5.36	−14.64	−3.25	−12.87	−0.70	−11.39	−0.27	−3.49
	changes in % working	22.24	13.39	19.78	19.54	11.51	21.79	10.35	23.19
Workfare version 2	% change in lifetime util.	0.76	−3.50	1.96	−2.69	2.40	−1.73	1.85	2.02
	changes in % working	20.61	19.78	17.78	23.93	7.79	25.63	7.29	18.27
Workfare version 3	% change in lifetime util.	5.49	4.77	5.57	4.91	4.59	5.29	3.32	5.74
	changes in % working	13.02	12.98	12.06	16.30	2.21	18.52	2.56	9.21

Table B8: Simulated Effects of Workfare Policies for Naive Present-Biased agents, by Initial Conditions.