Habit Persistence and Teen Sex: Could Increased Access to Contraception have Unintended Consequences for Teen Pregnancies?*

Peter Arcidiacono, Department of Economics, Duke University

Ahmed Khwaja, Fuqua School of Business, Duke University

Lijing Ouyang, Centers for Disease Control and Prevention

June 1, 2009

Abstract

We develop a dynamic discrete choice model of teen sex and pregnancy that incorporates habit persistence. Habit persistence has two sources here. The first is a 'fixed cost' of having sex which relates to a moral or psychological barrier that has been crossed the first time one has sex. The second is a 'transition cost' whereby once a particular relationship has progressed to sex, it is difficult to move back. We estimate significant habit persistence in teen sex, implying the the long run effects of contraception policy may be different from their short run counterparts. Programs that increase access to contraception are found to decrease teen pregnancies in the short run but increase teen pregnancies in the long run.

Keywords: habit persistence, access to contraception, teen pregnancy, dynamic discrete choice

JEL: J13, C33, C51

*We thank Jacob Klerman, Shannon Seitz, Jim Walker, and seminar participants at Carnegie Mellon, Clemson University, the IRP Summer Research Workshop, the University of Maryland, Northwestern University, and the 2003 Winter and Summer Meetings of the Econometric Society. This research was partially funded by the National Institute for Child Health and Human Development (R03 HD042817-01A1). The views expressed are those of the authors and do not represent the opinions of the Centers for Disease Control and Prevention.

1 Introduction

Teenage pregnancy rates, although steadily declining since 1990, are still very high in the United States with 83.6 pregnancies per 1000 teenage women in 2000 (Alan Guttmacher Institute 2004). This rate is substantially higher than Canada and Western Europe (Singh and Darroch 2000). Furstenberg (1998) argues that the higher rate in the United States is in part due to the lack of availability of contraception compared to Western Europe. On the other hand, it may be argued that increased availability of contraceptives will decrease the rate of unprotected sex, *and* lead some individuals to choose sex when they otherwise would have abstained due to the lower risk of pregnancy because of lowered costs of contraception.

Previous research has shown that regulations and the availability of contraception affect sexual behavior. Using aggregate data, researchers have found that restrictions on Medicaid funding of abortions or access to clinics reduced the number of adolescent abortions but either had no effect or reduced the number of teen births (for example, Kane and Staiger (1996) and Levine, Trainor, and Zimmerman (1996)). For abortions to fall with no effect on teen births, there must be a strong behavioral response in sexual actiivity. Similarly, using data from the United Kingdom, Paton (2002) found no evidence that nearness to family planning clinics reduced either the pregnancy rate or the abortion rate, with some evidence that family clinics increased the pregnancy rate, while Girma and Paton (2006) found no effect of the availability of emergency contraceptives on teen pregnancies. Should contraception become more available, those who switch from unprotected sex to protected sex will lower the teen pregnancy rate, while those who move from abstaining to protected sex will increase the teen pregnancy rate due to contraception failure.

The effects of contraception policy may differ between the long and the short run if there is habit persistence in teen sexual behavior. While habit persistence is generally associated with addictive goods, such as alcohol or cigarettes, it can result from other sources as well. For example, if there is a moral or psychological barrier which is crossed the first time one has sex (a fixed cost), once an individual has sex they will be more likely to have sex in the future.[1] The importance of virginity is implicitly stressed by much of the sociological literature

[1]Evidence, even if anecdotal, that a fixed cost to sex may exist is to be found in the growth of social movements for 'virginity pledges' where promises are made to wait until marriage before having sex. Virginity movements have been associated with significant decreases in teen sex and pregnancy rates, particularly for

on adolescent sex through the attention paid to time of first intercourse.[2] This emphasis on virginity suggests that the costs to abstaining from sex are higher for those who have been sexually active in the past.[3]

Data from the 1997 National Longitudinal Survey of Youth (NLSY97) also speak to the possibility of the cost of abstaining from sex being higher once one has become sexually active. This panel data set follows individuals who were between the ages of 12 and 16 in 1997. Table 1 shows the sex patterns for women who answered the questions on whether they had sexual intercourse in the first four waves of the survey and who were between the ages 14 and 16 in wave 1. Besides abstaining in all four waves, the most populous cells for each age are those where once one has become sexually active, one is sexually active from that point forward. Further, the least populous cells are those in which the individual transitions in and out of sexual activity. The patterns in the data suggest that there may be a fixed cost (a moral or psychological barrier that has been crossed) and a transition cost (which may be relationship-specific).

Note that person-specific effects cannot be completely driving the persistence. For example, consider an individual who had sex in 1997 but not in 1998. This individual is considerably *less* likely to have sex in 1999 than someone who had sex in 1998 but not in 1997. Having sex in one year and not the next is rare. In fact, the least likely outcome for all ages was a pattern of two stops in sexual activity; having sex, then not, then having sex, then not.

In this paper we examine the difference between the short and long run effects of contraception policy that may arise due to the presence of habit persistence in sexual activity. We

those under the age of 18 (Bearman and Brückner 1997). This same study finds that those who take pledges and break them are more likely to have unprotected sex. This result also parallels the results here in that programs that lower the utility of contraceptive sex lead to less sex but also less contraception conditional on sex. Our policy simulations suggest that the former effect dominates the latter in the long run.

[2]See, for example, Day (1992), Furstenberg, Brooks-Gunn, and Morgan (1987), Widmer (1997), Zabin et al. (1986), and Zelnik and Shaw (1983) among many others.

[3]Bearman and Brückner use the Add Health data (http://www.cpc.unc.edu/projects/addhealth) for their analysis. There is a burgeoning literature using this data that examines the relationship between various social and economic factors and teen sexual activity, e.g., Bearman, Moody and Stovel (2004) analyze the effects of social networks; Adamczyk and Felson (2006) study the effects of the religious propensity of friends; Ashby, Edmonson and Arcari (2006) investigate the impact of exposure to television programming. None of these studies examine the persistence of sexual behavior.

Table 1: Persistence in Female Sexual Activity†

				Age in 1997		
Sex in 97	Sex in 98	Sex in 99	Sex in 00	14	15	16
No	No	No	No	36.7%	26.8%	17.7%
No	No	No	**Yes**	12.0%	11.3%	8.0%
No	No	**Yes**	**Yes**	15.3%	13.9%	12.8%
No	**Yes**	**Yes**	**Yes**	16.9%	18.0%	20.3%
Yes	**Yes**	**Yes**	**Yes**	10.6%	18.2%	28.9%
No	No	**Yes**	No	1.4%	2.1%	2.3%
No	**Yes**	No	No	0.6%	1.2%	0.9%
No	**Yes**	No	**Yes**	3.4%	2.1%	1.9%
No	**Yes**	**Yes**	No	0.6%	0.7%	1.4%
Yes	No	No	No	0.4%	0.7%	0.3%
Yes	No	No	**Yes**	0.4%	0.6%	0.5%
Yes	No	**Yes**	No	0.0%	0.4%	0.2%
Yes	No	**Yes**	**Yes**	0.4%	1.8%	1.4%
Yes	**Yes**	No	No	0.2%	0.6%	0.8%
Yes	**Yes**	No	**Yes**	0.8%	1.2%	1.4%
Yes	**Yes**	**Yes**	No	0.2%	0.6%	1.2%

† All females who had valid answers for the sex questions in first four waves. Sample sizes are 498, 727, and 640 for ages 14, 15, and 16 respectively.

estimate a dynamic model of sex and fertility for teenage women which allows for habit persistence in both the choice to have sex and the choice to contracept. In each year, individuals decide whether to be sexually active and, if so, whether to use contraception that requires advanced planning (for example, the pill) or to use contraception which can be implemented at the time of the act itself (for example, condoms). Should an individual choose to engage in sex, she becomes pregnant with a probability that depends upon the choice of contraception. As in Hotz and Miller (1993) and Carro and Mira (2006), contraception will reduce, but not eliminate, pregnancy risk. We use the model estimates to forecast the short run and long run effects of increased access to contraception on the rates of teen sex and pregnancy.

As a precursor to the model, we show that fixed costs and transition costs can be accurately estimated even when there are permanent unobserved components to an individual's utility from sex and the panel is short. We conduct a series of Monte Carlos with different data generating processes to illustrate the biases that arise from not controlling for combinations of fixed and transition costs as well as unobserved preferences. The Monte Carlos show that separately identifying fixed costs, transition costs, and unobserved heterogeneity is indeed feasible even with a short panel.[4]

The estimates of the model reveal strong habit persistence in teen sexual behavior. These effects are so strong that policies which increase access to certain types of contraception such as condoms, while lowering teen pregnancy rates in the short run, may raise teen pregnancy rates in the long run. Consider a sixteen year old exposed to a policy that increases access to contraception, ceteris paribus. If the policy is a surprise, our simulations reveal that this individual will be less likely to become pregnant at age sixteen. However, a fourteen year old exposed to the same policy from the ages of fourteen through sixteen will actually have a higher probability of becoming pregnant at age sixteen. The differences in the long and short run effects are driven by the habit persistence. Individuals who are sixteen at the time the policy was implemented have already established certain sexual behaviors. Individuals exposed to the policy from age fourteen are more sexually active due to 'moral hazard' arising from the lower contraception costs (which lowers the risks of pregnancy). This increased sexual

[4]Distinguishing between unobserved heterogeneity and state dependence has received much attention in other contexts such as unemployment (e.g., Heckman 1981) and consumer choices (e.g., Keane 1997), with the specification of the structure of the unobserved heterogeneity being problem-specific.

activity is reenforcing due to habit persistence and results in higher long run pregnancy rates. Thus our results imply that well intentioned policies regarding teen access to contraception can have unintended consequences in the presence of habit persistence.[5]

Previous dynamic models of fertility decisions have not focused on the persistence of sexual behavior (for example Hotz and Miller 1993, Rosensweig and Shultz 1985, and Wolpin 1984). This is primarily because these studies focus on married couples and the optimal spacing of children; the act of sex itself is taken as given. Models of teen behavior have not focused on the dynamics of sex, in part because of inadequate data. Indeed, the two studies perhaps most related to our own, Lundberg and Plotnik (1995) and Oettinger (1999), used an earlier version of our data set which only asked at what age the individual first had sex; year by year questions on sexual behavior were not asked.

Lundberg and Plotnick (1995) estimate a sequence of choices in a static context using data from the 1979 National Longitudinal Survey of Youth (NLSY79). They estimate a nested logit model where the sequence of decisions is whether to have a premarital pregnancy, conditional on pregnancy, whether to have an abortion, conditional on not having an abortion, whether to become married. Lundberg and Plotnick find that the behavior of whites responds to the incentives of state welfare, abortion, and family planning policies. While we do not model the decision to marry or abort, we do model the choice to have sex and the choice of contraception and we explicitly account for the dynamics of the decisions.

Oettinger (1999) is one of the few studies in the economics literature to actually examine the decision to have sex as opposed to fertility outcomes. He also looks at fertility outcomes as well, but does not link the model of fertility to the model of sex. He estimates a hazard model of the time to first sex and time to first pregnancy using the NLSY79. Persistence in sexual activity, however, cannot be taken into account because the survey only asked when the respondent first engaged in sexual activity.

[5] Another way to capture these trends in the data would be to use time-varying heterogeneity. The data are not rich enough to distinguish between habit persistence and time-varying heterogeneity. We discuss this in more detail in the last section of the paper.

2 Monte Carlos

Before proceeding with the actual model and estimation, we first present evidence that we can separately identify habit persistence from unobserved preferences for sexual activity. The Monte Carlos have individuals choose whether or not to have sex, $c_t = 1$ and c_t=0 respectively, in each of five periods, $T = 5$.

We specify the flow utility from engaging in sex at time t to be a function of the following state variables:

1. whether the individual has ever had sex by period t, s_{ft},

2. whether, the individual had sex in the previous time period, s_{lt},

3. the individual's unobserved preference for sex which does not vary over time, s_i,

4. and the individual's time-varying unobserved preference for sex, ϵ_{1t}

We normalize the flow utility of sex of to zero except for the time-varying unobserved preference for no sex, ϵ_{0it}. Denote s_t as the set of state variables besides the ϵ's, $s_t = \{s_{ft}, s_{lt}, s_i\}$, and denote the flow utility of sex and no sex as $u_t(c_t = 1, s_t, \epsilon_t)$ and $u_t(c_t = 1, s_t, \epsilon_t)$. These flow utilities are given by:

$$u_t(c_t = 1, s_t, \epsilon_t) = \alpha_0 + \alpha_1 s_i - \alpha_2(1 - s_{ft}) - \alpha_3(1 - s_{lt}) + \epsilon_{1t} \quad (1)$$

$$u_t(c_t = 0, s_t, \epsilon_t) = \epsilon_{0t} \quad (2)$$

We further assume that the ϵ_t's are IID Type I extreme value and unknown to the individual until time t.

The individual chooses to have sex at T if $u_T(c_T = 1, s_T, \epsilon_T) > u_T(c_T = 0, s_T, \epsilon_T)$. The value function at T is then given by $V_T(s_T, \epsilon_T) = \max\{u_T(c_T = 1, s_T, \epsilon_T), u_T(c_T = 0, s_T, \epsilon_T)\}$. In time periods previous to T, individuals weigh the utility they receive today against how their choice today affects their expected utility in the future. Denote $v_{T-1}(c_{T-1} = j, s_{T-1}, \epsilon_{T-1})$ as the conditional value function for choosing j at time $T - 1$: the expected discounted utility of choosing j conditional on making the optimal choice at time T. With utility being time-separable and β denoting the discount factor, $v_{T-1}(c_{T-1} = j, s_{T-1}, \epsilon_{T-1})$ is given by:

$$v_{T-1}(c_{T-1} = j, s_{T-1}, \epsilon_{T-1}) = u_{T-1}(c_{T-1} = j, s_{T-1}, \epsilon_{T-1}) + \beta E\left(V_T(s_T, \epsilon_T | c_{T-1} = j, s_{T-1})\right) \quad (3)$$

where the expectation is taken over the ϵ_T's as the variables in s_t evolve deterministically given the choices. For all periods less than T we can then use backwards recursion to obtain the conditional value function for choosing j at time t:

$$v_t(c_t = j, s_t, \epsilon_t) = u_t(c_t = j, s_t, \epsilon_t) + \beta E\left(V_{t+1}(s_{t+1}, \epsilon_{t+1})|c_t = j, s_t\right) \quad (4)$$

where $v_T(c_T = j, s_T, \epsilon_T) = u_T(c_T = j, s_T, \epsilon_T)$ and $V_{t+1}(s_{t+1}, \epsilon_{t+1}) = \max\{v_{t+1}(c_{t+1} = 1, s_{t+1}, \epsilon_{t+1}), v_{t+1}(c_{t+1} = 0, s_{t+1}, \epsilon_{t+1})\}$.

Given the assumptions we have made on the data generating process, Rust (1987) derived the closed form expression for the expected value of future utility. Let $v_t(c_t, s_t) = v_t(c_t, s_t, 0)$ denote the conditional valuation function net of ϵ_{jt}. Then, for $t < T$ we have:

$$v_t(c_t = j, s_t, \epsilon_t) = u(c_t = j, s_t, \epsilon_t) + \beta \ln\left(\sum_{k=0}^{1} \exp\left[v_{t+1}(c_{t+1} = k, s_{t+1}|c_t = j, s_t)\right]\right) \quad (5)$$

We further assume that s_i follows a finite mixture distribution and can take on only two values, zero and one. The probability of a simulated individual having $s_p = 1$ is π_1, implying that $\pi_0 = 1 - \pi_1$. In addition to fixed and transition costs and unobserved heterogeneity, we also control for time fixed effects.[6] Although all the time effects are set at zero, controlling for these effects may soak up some of the habit persistence when unobserved heterogeneity is ignored and some the unobserved heterogeneity when habit persistence is ignored.[7]

We then generate data in three ways:

1. Fixed and transition costs are non-zero but there is no unobserved heterogeneity. That is, α_2 and α_3 are non-zero but $\alpha_1 = 0$.

2. Fixed and transition costs are set at zero but there is unobserved heterogeneity. That is, $\alpha_2 = \alpha_3 = 0$ and $\alpha_1 \neq 0$.

3. Fixed and transition costs are non-zero and there is unobserved heterogeneity.

For each of these data generating processes we estimate models that include and do not include fixed and transition costs as well as unobserved heterogeneity. We can then see whether fixed

[6]These are equivalent to age fixed effects given that in the Monte Carlos all individuals are the same age in each time period.

[7]Not including time dummies exacerbated the biases as well as made the predictions of misspecified models even further from those of the data generating process.

and transition costs can be separately identified from unobserved heterogeneity given a short panel and also assess the biases that result from not controlling for one or the other. The log likelihood for the most general model then follows a mixture distribution as s_i is unobserved:

$$L(\alpha, \pi) = \sum_{n=1}^{N} \log \left(\sum_{k=0}^{1} \pi_k \left[\prod_{t=1}^{5} \mathcal{L}(c_{nt}|s_{nt}, \alpha, k) \right] \right)$$

where:

$$\mathcal{L}(c_{nt}|s_{nt}, \alpha, k) = \frac{\exp(v_t(c_{nt}, s_{nt}|k)}{\sum_{j=0}^{1} \exp(v_t(j, s_{nt}|k)}$$

Table 2 shows the average parameter estimates for 2000 simulated data sets of 3000 individuals each as well as the parameters of the data generating process. The value of β is set at 1, though we have also performed all the simulations with β set at 0 and the qualitative results were the same. The first column shows the model estimates with no unobserved heterogeneity, the second column has unobserved heterogeneity but no habit persistence, and the third column has both habit persistence and unobserved heterogeneity.

The first set of rows shows the case when the data generating process has no unobserved heterogeneity but the fixed and transition costs are significant. Not surprisingly, when the estimated model is specified as the data generating process the parameter estimates match those of the data generating process. Further, estimating a model with both habit persistence and unobserved heterogeneity when the true model does not have unobserved heterogeneity still yields correct estimates of habit persistence with the unobserved heterogeneity parameters not significantly different from zero: the log likelihoods in the first and third columns are identical. The second column leads to significant estimates of unobserved heterogeneity as the unobserved heterogeneity approximates (poorly) the habit persistence.

The second set of rows shows the case when the data generating process has unobserved heterogeneity but no habit persistence. Estimating the model with just habit persistence shows that the habit persistence parameters pick up some of the effects of the unaccounted for unobserved heterogeneity. Estimating with just unobserved heterogeneity or with both unobserved heterogeneity and habit persistence yields estimates indistinguishable from the true values, with the latter yielding habit persistence parameters that were very small and insignificant.

The most interesting results are in the third set of rows. Here we show that when the true data generating process has both unobserved heterogeneity and habit persistence estimating

Table 2: Parameter Estimates from the Monte Carlos†

		True	Model 1		Model 2		Model 3	
		Values	Coeff.	Std. Dev.	Coeff.	Std. Dev.	Coeff.	Std. Dev.
	Intercept	0	0.002	0.046	0.176	0.068	-0.181	0.731
DGP	Fixed Cost	0.4	0.402	0.042			0.392	0.044
Model 1	Trans. Cost	0.6	0.598	0.060			0.589	0.061
	$s_i = 1$				1.758	0.071	0.398	0.982
	Prob. $s_i = 1$				0.354	0.038	0.495	0.197
	Log likelihood		-9589		-9750		-9589	
	Intercept	0	-0.595	0.044	0.015	0.130	0.012	0.166
DGP	Fixed Cost		0.082	0.044			0.001	0.047
Model 2	Trans. Cost		0.1806	0.0507			0.001	0.057
	$s_i = 1$	-1			-1.035	0.081	-1.041	0.103
	Prob. $s_i = 1$	0.5			0.5	0.118	0.495	0.130
	Log likelihood		-9970		-9948		-9947	
	Intercept	0	-0.640	0.043	1.455	0.096	0.016	0.145
DGP	Fixed Cost	0.4	0.528	0.051			0.401	0.056
Model 3	Trans. Cost	0.6	0.792	0.067			0.597	0.073
	$s_i = 1$	-1			-2.101	0.0576	-1.0459	0.1161
	Prob. $s_i = 1$	0.5			0.686	0.0254	0.498	0.114
	Log likelihood		-8748		-8860		-8727	

† Estimates were taken form 2000 sets of 3000 simulated individuals. The discount factor is set at 1. Model 1 refers to the estimated model where unobserved heterogeneity is unaccounted for, DGP model 1 refers to the data generating process where there is no unobserved heterogeneity. Model 2 refers to the estimated model where habit persistence is unaccounted for, DGP model 2 refers to the data generating process where there is no habit persistence. Model 3 and DGP model three refer to the estimated model and the data generating process respectively when both unobserved heterogeneity and habit persistence are present.

a model with both will yield estimates indistinguishable from those of the data generating process even when the number of observations per individual is small. Consistent with the first and and second set of rows, not accounting for unobserved heterogeneity when it is present yields too much habit persistence while not accounting for habit persistence yields parameter estimates that attribute too much of the variation in the data to unobserved heterogeneity.

Table 3 presents some of the patterns that result from the data generating process that help to separate out unobserved heterogeneity from habit persistence. Here we show probabilities of various courses of action where the true data generating process has both unobserved heterogeneity and habit persistence. In particular, we examine the probability of having sex conditional on three histories: never having had sex, had sex in the past but not in the previous period, and had sex in the previous period.

Not surprisingly, a model which accounts for both unobserved heterogeneity and habit persistence yields estimated behavior similar to the trends in the data. Accounting for just unobserved heterogeneity yields too low of a sex rate for those who had sex in the previous period and too high of a sex rate for those who have never had sex— particularly in the earlier time periods. In contrast, in the first few time periods, accounting for just habit persistence yields too high of a sex rate for those had sex in the previous period and too low of a sex rate for those who have never had sex. These patterns flip in later time periods with sex rates that are too high when an individual has never had sex and when we do not account for unobserved heterogeneity. The Monte Carlos then suggest that the patterns of sexual behavior over time allow us to separately identify both habit persistence and unobserved heterogeneity if either are present.

3 Model and Estimation

What distinguishes our work from much of the previous literature is that we do not model fertility as a single choice, but as a sequence of choices integrated with uncertainty about pregnancy outcomes. Individuals make decisions regarding sex knowing that there is some probability of becoming pregnant, with the probability being lower if contraceptives are used. Individuals may still engage in unprotected sex even though *ex post* they may regret the decision if they become pregnant. This can still be consistent with a rational expectations

Table 3: Predicted Probabilities of Sex Conditional on Sex Histories†

		Probability of Sex				
		t=1	t=2	t=3	t=4	t=5
	Data	0.510	0.459	0.431	0.437	0.512
Mean |	Model 1‡	0.510	0.448	0.437	0.459	0.543
History§=1	Model 2	0.510	0.540	0.479	0.454	0.501
	Model 3	0.510	0.459	0.431	0.437	0.513
	Data			0.653	0.642	0.671
Mean |	Model 1			0.647	0.639	0.668
History=2	Model 2			0.696	0.663	0.673
	Model 3			0.653	0.642	0.670
	Data		0.808	0.802	0.800	0.820
Mean |	Model 1		0.818	0.802	0.796	0.816
History=3	Model 2		0.730	0.762	0.782	0.816
	Model 3		0.808	0.802	0.800	0.820

† Simulations are conducted using the parameter values from the third set of rows in table 2. Each simulation is conducted on 2000 times using 3000 simulated individuals.

‡ Model 1 accounts for habit persistence but not unobserved heterogeneity, model 2 accounts for unobserved heterogeneity but not habit persistence, model 3 accounts for both.

§ History=1 refers to those who have never had sex. History=2 refers to those who have had sex in the past but not at $t-1$. History=3 refers to those who have had sex at $t-1$.

framework. The key is that, given the probability of getting pregnant without contraception, the expected utility was still higher for having unprotected sex. Hence, the flow utility of having sex without contraception compensated the individual for the increased probability of becoming pregnant. We first discuss the model without including unobserved heterogeneity and then follow with how unobserved heterogeneity is incorporated.

3.1 Base Model

We propose a dynamic discrete choice model of sex and contraception decisions. Throughout the model, we want to distinguish 'flow utility', utility in the period, from the full consequences of having sex which include the utility of various pregnancy outcomes. Although decisions with regard to sex are joint decisions, since women have to bear the consequences of a pregnancy through carrying the child and have the exclusive right to abort the child, we model the decisions from the perspective of the woman.

In each period of T periods women choose whether to engage in sexual activity. Those who engage in sex must also decide whether to contracept and, if so, what type of contraception to use. We distinguish between two types of contraception: *scheduled contraception*, which requires advanced planning such as the pill, and *episode-specific contraception*, where the choice to use it can be postponed until the act itself. Define c_t as the sex and contraception combination chosen at time t where:

$$c_t = \begin{cases} NS & \text{if abstains} \\ NC & \text{if sex, but no contraception} \\ EC & \text{if sex with episode-specific contraception} \\ SC & \text{if sex with scheduled contraception} \end{cases} \tag{6}$$

Individuals receive flow utility from having sex that may vary across observable characteristics such as the family environment. The flow utility may also depend upon past sexual decisions. We allow there to be both a fixed and transition cost for sex itself, as well as fixed and transition costs for the two types of contraception. The fixed and transition costs for sex are included to capture a moral or psychological barrier associated with losing one's virginity (the fixed cost) or through having sex the first time in a particular relationship (the transition cost). With regard to contraception, in addition to a per-period cost to using it, there may be a fixed cost to learning or acquiring particular contraception and it may also be easier to access particular forms of contraception if they were used recently. Let s_{ft} represent the possible fixed cost states and s_{lt} the lagged choices which are over the same set as c_t. Define s_{ft} as:

$$s_{ft} = \begin{cases} NS & \text{if never had sex in the past} \\ NC & \text{if had sex in the past, but never used contraception} \\ EC & \text{if had sex in the past, but only used episode-specific} \\ & \text{contraception or no contraception} \\ SC & \text{if had sex in the past, but only used scheduled} \\ & \text{contraception or no contraception} \\ BC & \text{if had sex in the past and have used both scheduled} \\ & \text{and episode-specific contraception} \end{cases}$$

The utility of sex may also be a function of the family environment (e.g., mother works, two parent family, religion). We denote these variables as s_{xt}. Normalizing the flow utility of abstaining to zero, the flow utility for sex without contraception, with episode-specific contraception, and with scheduled contraception are specified as follows:

$$u_t(NC, s_t, \epsilon_t) = s_{xt}\alpha_0 - [1 - (s_{ft} \neq NS)]\alpha_1 - [1 - (s_{lt} \neq NS)]\alpha_2 + \epsilon_t(NC) \quad (7)$$

$$\begin{aligned} u_t(EC, s_t, \epsilon_t) = s_{xt}\alpha_0 & - [1 - (s_{ft} \neq NS)]\alpha_1 - [1 - (s_{lt} \neq NS)]\alpha_2 \\ & - [1 - (s_{ft} \in \{EC, BC\})]\alpha_{31} - [1 - (s_{lt} = EC)]\alpha_{41} - \alpha_{51} + \epsilon_t(EC) \end{aligned} \quad (8)$$

$$\begin{aligned} u_t(SC, s_t, \epsilon_t) = s_{xt}\alpha_0 & - [1 - (s_{ft} \neq NS)]\alpha_1 - [1 - (s_{lt} \neq NS)]\alpha_2 \\ & - [1 - (s_{ft} \in \{SC, BC\})]\alpha_{32} - [1 - (s_{lt} = SC)]\alpha_{42} - \alpha_{52} + \epsilon_t(SC) \end{aligned} \quad (9)$$

The ϵ's refer to the unobserved preference for particular sex and contraception combinations. As in the section 2, we assume that the ϵ's are distributed Type I extreme value and IID across time and choices.[8] Note that the coefficients on s_{xt} are common across the sex choices. Similarly, the fixed and transition costs for sex itself does not vary across the contraception choices.

[8]More general correlation structures are possible using a GEV framework (for example, Arcidiacono 2005 and Khwaja 2001). We experimented with more flexible correlation structures and were unable to reject the Type I extreme value assumption.

Those who engage in sex weigh the flow utility of sex against the probability and consequences of becoming pregnant. Conditional on having sex, the probability of becoming pregnant will depend upon variables related to fecundity, s_{zt}, as well as the choice of contraception. We assume that the probability of becoming pregnant follows a logit form:

$$p_{ct} = \frac{\exp(s_{zt}\gamma_0 + \gamma_c)}{1 + \exp(s_{zt}\gamma_0 + \gamma_c)} \tag{10}$$

The γ_c's refer to indicator variables for each of the possible contraception choices (no contraception, episode-specific, or scheduled). We assume that individuals know these probabilities when they make their sex decisions.[9]

Given that the decision as to whether to abort or give birth is often very traumatic, and given limited data on abortions,[10] we do not model the abortion decision. Rather, we model the lifetime utility from that point forward using a terminal value function. Since we are only interested in teenage sexual behavior, we also assign a terminal value function for individuals who arrive to age nineteen without becoming pregnant. Examples of variables which will affect this terminal utility but not the flow utility of having sex are ability measures and family income as the opportunity cost of a child will be higher for those who have better expected labor market outcomes. We write the terminal value of becoming pregnant at time t as a linear function of variables we expect to affect the costs and benefits of becoming pregnant, s_{pt}:

$$V_{pt}(s_t) = s_{pt}\alpha_6. \tag{11}$$

Normalizations must be made in order to identify the model. We normalize the utility of making it to age 20 (T+1) without a pregnancy to zero. The conditional value functions at T, $v(c_T, s_T, \epsilon_T)$, are then given by:

$$v_T(c_T, s_T, \epsilon_T) = u_T(c_T, s_T, \epsilon_T) + \beta p_{cT} V_{pT+1}(s_{T+1}) \tag{12}$$

where β is the discount factor. The value function at time T is then $V_T(s_T, \epsilon_T) = \max_{c_T}\{v_T(c_T, s_T, \epsilon_T)\}$.

[9] An alternative is to ask individuals about what they know regarding contraceptive efficay. Delavande (2008) examines what individuals know about contraceptive effectiveness and how their knowledge affects contraceptive choice. van der Klaauw (2000) shows how to incorporate these subjective expectations into a dynamic discrete choice model.

[10] See Jones and Forest (1992) for a detailed discussion of the quality of abortion data from surveys.

In addition to uncertainty regarding future pregnancy outcomes and on unobservable preferences, individuals also face uncertainty regarding the state variables themselves.[11] We assume that individuals know the stochastic processes governing the transitions of the demographic variables and these transitions do not depend upon decisions made by the individual regarding their sexual behavior. Since the transitions on the demographic variables are not the focus of the analysis, we discuss the estimation of the transitions of these variables in the appendix. The transition of sexual histories is completely determined by the individuals' choices. Let s_t denote the full set of state variables, including the state variables that affect sex, the transitions on the demographics, and the cost of pregnancy, the probability of pregnancy, We then specify the probability of moving from state s_t to s_{t+1} as $q(s_{t+1}|s_t, c_t)$ implying that s_{t-1} only affects s_{t+1} through s_t and c_t.[12]

The conditional valuation functions at $T-1$ then follow:

$$v_{T-1}(c_{T-1}, s_{T-1}, \epsilon_{T-1}) = u_{T-1}(c_{T-1}, s_{T-1}, \epsilon_{T-1}) + \beta E\left[p_{cT-1}V_{pT}(s_T) + (1-p_{cT-1})V_T(s_T, \epsilon_T)|c_{T-1}, s_{T-1}\right] \tag{13}$$

where the expectations are taken over pregnancies and the values of the observed and unobserved state variables. Using backwards recursion we can then write the conditional valuation functions at time t as:

$$v_t(c_t, s_t, \epsilon_t) = u_t(c_t, s_t, \epsilon_t) + \beta E\left[p_{ct}V_{pt+1}(s_{t+1}) + (1-p_{ct})V_{t+1}(s_{t+1}, \epsilon_{t+1})|c_t, s_t\right] \tag{14}$$

Given the assumptions on the ϵ's, closed form solutions for the conditional expectations of future utility exist (Rust 1987). Denote $v_t(c_t, s_t) = v_t(c_t, s_t, 0)$, the conditional valuation function net of the ϵ's. We can then write equation (**??**) as:

$$v_t(c_t, s_t, \epsilon_t) = u_t(c_t, s_t, \epsilon_t) + \beta \sum_{s_{t+1}} \left[p_{ct}V_{pt+1}(s_{t+1}) + (1-p_{ct}) \ln\left(\sum_{c_{t+1}} \exp(v_{t+1}(c_{t+1}, s_{t+1}))\right)\right] q(s_{t+1}|c_t, s_t) \tag{15}$$

In order to ensure that we are truly picking up habit persistence, we also control for a full set of age-choice interactions normalized with respect to the utility of no sex. Full sets of age

[11]While uncertainty also exists due to the probability of contracting a sexually transmitted disease, our data set includes no information on STDs.

[12]Note that the conditioning on c_t only matters through the sexual histories.

dummies come at a cost as it is now unclear what variation in the data identifies the discount factor and we therefore set the discount factor to 0.9.[13]

Probabilities of choosing particular sex and contraception combinations then yield multinomial logit probabilities where, instead of the term inside the exponential function being linear in the parameters, it is highly nonlinear. The log likelihood function is then the sum of three parts:

1. L_c–the log likelihood of the sex and contraception choices,

2. L_p–the log likelihood of the pregnancy outcomes

3. L_s– the log likelihood of the transitions on the demographics.

The full log likelihood function is then:

$$L(\alpha,\beta,\gamma)=\sum_{n=1}^{N}\sum_{t=1}^{T}L_c(c_{nt}|s_{nt},\alpha,\beta,\gamma)+L_p(Preg_{nt}|c_{nt},s_{nt},\gamma)+L_s(s_{nt}|s_{nt-1},\gamma) \tag{16}$$

where L_c, L_p, and L_s refer to the log likelihood contributions of the sex and contraception decision, the probability of becoming pregnant, and the transitions on the other state variables. Since the log likelihood is additively separable, it is possible to estimate the parameters of the model in stages. In particular, we estimate the pregnancy and transition parameters separately from the utility function parameters.

3.2 Unobserved Heterogeneity

There may, however, still be unobservable preferences for having sex. As shown in the Monte Carlos, not controlling for unobserved heterogeneity may attribute too much sexual activity to habit persistence.[14] In order to account for unobserved preferences, we estimate the model using mixture distributions to allow for unobserved heterogeneity in the taste for sex. Mixture distributions can be used to overcome this problem and control for 'dynamic selection.'[15]

[13] Arcidiacono, Sieg, and Sloan (2007) use age-behavior profiles to identify the discount factor. We have also estimated the model with the discount factor set at zero and the habit persistence parameters were large and significant under this alternative specification. This suggests that the presence of habit persistence is robust to the value of the discount factor.

[14] Gilleskie and Strumpf (2005) deal with a similar issue in the context of youth smoking.

[15] In addition to Gilleskie and Strumf, mixture distributions have been used to account for dynamic selection in dynamic discrete choice models by Keane and Wolpin (1997, 2000, 2001), Eckstein and Wolpin (1999), and

Namely, let there be K types of people with π_k being the population probability of being the kth type. Preferences are common across types and types are known to the individuals. We assume that, conditional on observed characteristics and one's unobserved type, the unobserved preferences are serially uncorrelated. Treating type as a random effect, it is possible to integrate out the probability of being a particular type. The log likelihood function is then:

$$L(\alpha, \beta, \gamma, \pi) = \sum_{n=1}^{N} \log \left(\sum_{k=1}^{K} \pi_k \left[\prod_{t=1}^{T} \mathcal{L}_c(c_{nt}|s_{nt}, \alpha, \beta, \gamma, k) \mathcal{L}_p(Preg_{nt}|c_{nt}, s_{nt}, \gamma) \mathcal{L}_s(s_{nt}|s_{nt-1}, \gamma) \right] \right) \tag{17}$$

where the π's are the population probabilities of being a particular type[16] and the $\mathcal{L}$'s are the likelihoods rather than the log likelihoods. Since the probability of a pregnancy conditional on sexual behavior as well as the other transition processes does not depend upon type, we can rewrite the log likelihood as:

$$L(\alpha, \beta, \gamma, \pi) = \sum_{n=1}^{N} \log \left(\sum_{k=1}^{K} \pi_k \left[\prod_{t=1}^{T} \mathcal{L}_c(c_{nt}|s_{nt} \alpha, \beta, \gamma, k) \right] \right) + \sum_{t=1}^{T} L_p(Preg_{nt}|c_{nt}, s_{nt}, \gamma) + L_s(s_{nt}|s_{nt-1}, \gamma). \tag{18}$$

Estimation can again proceed in stages. As before, we first estimate the probability of becoming pregnant conditional on the choice of sexual activity as well as the transitions on the demographic variables. Taking these parameters as given, we estimate the parameters of the utility function.

4 Data

We use the 1997 National Longitudinal Survey of Youth (NLSY97) to estimate the model. The NLSY97 data contain surveys of youths born during the years 1980 to 1984. The first survey was conducted in 1997, when the individuals were between the ages of 12 and 16. Participants are interviewed each year, with six waves of the data currently available. Blacks and Hispanics were oversampled.

Cameron and Heckman (1999, 2001) among many others. In their analysis of married couples, Carro and Mira (2006) model the unobserved heterogeneity on pregnancy probabilities rather than sex.

[16] In principle, these population probabilities can vary with state variables. See, for example, Keane and Wolpin (1997, 2000, 2001). This is done to account for initial conditions, something which is unnecessary as we have the full sex history.

In each wave those surveyed by the NLSY97 answer questions about their sexual activity. They are also asked what percentage of the time they used contraception and what their primary form of contraception was. Over 55% report using birth control every time they had sexual intercourse. We classify those who use contraception less than 100% of the time under the unprotected (no contraception) category. Those individuals who reported using contraception every time but whose primary method of birth control was withdrawal or the rhythm method were classified as unprotected as well. Episode-specific contraception was defined as condoms, foam, jelly, sponges, and diaphragm. Scheduled contraception was defined as the pill, intrauterine devices (IUDs), Norplant, Depo-Provera, and injectables.[17] We only use continuous sex histories beginning from age 14. For example, if an individual is 14 in wave 1 and answers the sex questions in wave 1 and 2 but not wave 3, no survey answers after wave 2 would be used regardless of whether or not answers were given in waves 4 and 5. For more detailed descriptive statistics on the sex rates and the use of contraception, see Walker (2003).

Means conditional on the choice of sex and contraception type are given in Table 4 using the first five waves of the survey.[18] Roughly seventeen percent of the sample was classified as unprotected, with twelve and eight percent using episode-specific and scheduled contraception respectively. Those who engage in sex tend to be older, particularly those who choose scheduled contraception.

The variables listed either affect the flow utility directly or affect decisions through the terminal value function. All independent variables are taken from wave one of the survey with the exception of mother working, two-parent family and whether the individual was living with their biological mother. The mother work variable takes on a value of one if the mother

[17] In earlier waves, individuals were asked what fraction of the time they used birth control and, if they were protected what the primary method of birth control was. In later waves, the individuals were first asked what percent of the time a condom was used when they had sex. They were also asked what percent of the time birth control was used as well as the primary birth control method besides a condom. If the woman reported using a condom 100% of the time, than the birth control method was classified as episode-specific. If all the acts were protected but a condom was used less than 100% of the time, we used the primary method besides a condom to determine whether to classify the birth control as episode-specific or scheduled.

[18] Only individuals fourteen and older were asked the sex questions, while only individuals 14 and younger in 1997 were asked about parental religious practices, limiting our sample sizes. We also eliminate all individuals who did not report a family income in wave 1.

works full-time.[19] Two-parent family refers to the family structure where the teen lives with both biological parents. While coming from a two-parent family is associated with less sex, having a working mother or no longer living with one's biological mother is associated with higher sexual activity.

Having a mother who prays more than once a day is associated with abstaining.[20] Conditional on using contraception, a praying mother is associated with a higher probability of choosing episode-specific contraception. Higher incomes and test scores are also associated with abstaining and, conditional on sex, higher rates of contraception suggesting that the cost of a pregnancy may be higher for those with better labor market options.

The NLSY97 contains detailed information on the timing of births, abortions, and miscarriages. For the pregnancy data, we date all births, abortions, and miscarriages back to when the sex act would have taken place. A birth reported in wave 2 may have resulted from intercourse in either wave 1 or wave 2. To determine whether pregnancy resulted from sex in wave 1 or wave 2, the date a birth takes place is dated back nine months. This latter date is then linked to the sex decisions for the relevant wave. Similarly, the NLSY97 reports the date of miscarriages and abortions as well as how far along the pregnancy was at the time of the miscarriage or abortion. Pregnancies are then the sum of births, miscarriages, and abortions.

Table 5 presents means conditional on pregnancy outcomes. Because there are so few variables used here and since by assumption the pregnancy parameters can be estimated outside of the model, we are able to use a much larger sample. The overall pregnancy rate in the sample is 12% for those who are sexually active. Those who become pregnant are older than those who did not. Both kinds of contraception are associated with much lower pregnancy rates than unprotected sex. However, even with 100% reported contraception pregnancies still result.

[19] In wave 1, we do not observe whether the mother worked full-time or part-time. For wave 1, we classify a mother as working full-time if she also worked full-time in wave 2. For those who did not work full-time in wave 2 but reported working in wave 1, we set the probability of working full-time in wave 1 to match the transitions from work to not work, and work to work in the future waves. This probability was 0.38.

[20] We also investigated the effects of denominations and found no relationship between denomination and sexual behavior.

Table 4: Means Conditional on Sex and Contraception Choices†

		Contraception Conditional on Sex		
			Episode-	
	No Sex	None	Specific	Scheduled
Black	0.221	0.298	0.342	0.179
Mother Works Full-time‡	0.601	0.726	0.660	0.715
Two Parent Family§	0.613	0.417	0.472	0.480
No Longer Living with Mother	0.137	0.327	0.260	0.261
Mother Prays Regularly (1997)	0.597	0.524	0.586	0.528
Math Percentile	52.2	44.3	43.9	51.1
	(33.8)	(33.0)	(33.5)	(32.2)
Household Income (1997)	47765	36551	38913	38279
	(40768)	(30677)	(34563)	(28043)
Age	16.0	16.9	16.7	17.2
	(1.3)	(1.3)	(1.3)	(1.1)
Observations	3080	834	707	341

†Standard deviations in parenthesis

‡Conditional on living with one's biological mother and after wave 1. In wave 1, no distinction was made between part and full-time work.

§ Conditional on living with one's biological mother. No updated information is available on these variables when the individual leaves home.

Table 5: Means Conditional on Pregnancy State for Those Who Were Sexually Active†

	Not Pregnant	Pregnant
Episode-Specific Contraception	0.353	0.185
Scheduled Contraception	0.227	0.168
Age	18.0	18.4
	(1.7)	(1.6)
Observations	7437	1012

†Standard deviations in parenthesis. Sample includes all females between the ages of 14 and 22 who had sex in waves 1-5. All pregnancies are dated back to the sex decisions in these waves. The sample also include pregnancies reported in wave 6 that resulted from sex acts in wave 5.

5 Results

We now proceed to the estimates of the model, beginning with the pregnancy parameters. Estimates of the transition parameters on family status, mother working, and living with one's biological mother are reported in the appendix. Recall that the pregnancy parameters are only estimated for those who chose to engage in sexual intercourse. Table 6 presents the logit estimates of the probability of a pregnancy. As individuals age, they are more likely to become pregnant conditional on their choice of protection. The coefficients on both episode-specific and scheduled contraception are significantly negative.

To get a sense for the magnitude of the effects of age and contraception on the probability of becoming pregnant, Table 7 shows the estimated pregnancy probabilities conditional on age and contraception choice. The probability of becoming pregnant increases substantially as individuals age, doubling from age 14 to age 19 regardless of the type of contraception employed. Both contraception methods are effective but not foolproof. Unprotected pregnancy rates are about two times that of episode-specific and scheduled contraception.[21]

[21] These differences may seem small but recall that an individual was classified as having protected sex only if the individual reported using contraception one hundred percent of the time. The unprotected category then includes individuals who often used protection but did not report using contraception every time.

Table 6: Probability of Becoming Pregnant†

	Coefficient	Standard Error
Episode-Specific Contraception	-1.038	0.087
Scheduled Contraception	-0.805	0.092
Age	0.169	0.021
Constant	-4.632	0.384
Observations	8449	

† Logit estimates. Sample includes all females between the ages of 14 and 22 who had sex in waves 1-5. All pregnancies are dated back to the sex decisions in these waves. The sample also include pregnancies reported in wave 6 that resulted from sex acts in wave 5.

Table 7: Estimated Pregnancy Probabilities

	Age					
	14	15	16	17	18	19
Unprotected	0.0937	0.1091	0.1266	0.1465	0.1689	0.1939
Episode-Specific	0.0353	0.0416	0.0488	0.0573	0.0671	0.0785
Scheduled	0.0442	0.0519	0.0608	0.0712	0.0832	0.0971

The parameters characterizing the flow utility of sex are given in Table 8. The first set of rows shows the coefficients on the demographic characteristics. Having a mother who prays regularly and coming from a two parent family both lower the probability of engaging in sex. In contrast, having a mother who works or not living with one's biological mother positively affects the probability of sex.

The second panel shows the parameters of the terminal value function conditional on becoming pregnant. Higher test scores and parental income then makes sex—particularly unprotected sex— less attractive as the opportunity cost of a child is increasing in both test scores and income. We parameterized the unobserved preferences for sex using a two-type mixture distribution.[22] The second type, which makes up a little less than half the population, is substantially less likely to have sex than the first type.

The final set of rows show the persistence parameters. For both contraception choices, we see no transition costs. However, both yield substantial fixed costs with the fixed cost of scheduled contraception roughly three times that of episode-specific contraception. There then may be a tradeoff between encouraging the use of the pill versus encouraging condoms. Condoms help prevent sexually transmitted diseases, but encouraging individuals to use the pill will make birth control more of a habit. Both fixed and transition costs are significant for sex itself, with the fixed cost being roughly three times that of episode-specific contraception and approximately three times the sex transition cost. Such large effects imply that the long run effects of policy on sexual behavior may be different from the short run effects.

The utility function parameters are difficult to interpret because of the nonlinearities in the choice function. To see how demographic characteristics and habit persistence affect the sex choices, we calculate the probabilities of each of the choices given different demographics and sex histories. In particular, we forecast the decisions of sixteen year olds given the characteristics of those who are fourteen. We then assign the different values for particular demographic characteristics and see how these affect the probability of choosing particular sex options at age sixteen. Results of these simulations are given in Table 9.

The first set of rows gives the unconditional probabilities of sex choices at age sixteen. Moving from having a mother who does not pray regularly to one who does decreases the probability of having sex by slightly under seven percentage points. Contraception is relatively more popular for those with mothers who do not pray regularly. Individuals who are more likely to have sex in the future also expect to receive higher benefits from paying the fixed costs associated with contraception. The effect of an intact family is stronger–moving from an intact family at age fourteen to a single parent family at age fourteen leads to an over eleven percentage point increase in the probability of having sex at age sixteen. In contrast, having a working mother at age fourteen pushes up the probability of having sex at fourteen by three and a half percent. For math test scores, moving form the 25th percentile to the 75th

[22] We experimented with more types but the results consistently yielded estimates such that additional types were indistinguishable from the first two types.

Table 8: Parameters of the Utility Function†

	Variable	Coefficient	Std. Error
	Black	0.283	0.136
	Mother Works Full-time	0.292	0.096
Demographics	Two Parent Family	-0.520	0.114
	No Longer Living with Mother	0.567	0.138
	Mother Prays Regularly	-0.256	0.088
	Type 2	-1.656	0.227
	Black	-0.770	1.296
Pregnancy	Math Score (0's)	-0.268	0.144
Costs	1997 Family Income ($0000)	-0.505	0.145
	Sex Transition Cost	-0.398	0.183
	Sex Fixed Cost	-1.425	0.262
Habit	Episode-Specific Contraception Transition Cost	0.104	0.239
Persistence	Episode-Specific Contraception Fixed Cost	-0.489	0.191
	Scheduled Contraception Transition Cost	-0.133	0.498
	Scheduled Contraception Fixed Cost	-1.593	0.438
	Prob. Type 2	0.434	0.058
	Observations	4962	

† Estimates from the dynamic discrete choice model on only those who have continuous sex histories. The discount factor is set at 0.9. The utility function included age interacted with each of the choices which allows for pregnancy costs and flow tastes for the various choices to vary by age.

percentile of the distribution increases the probability of sex by less than three percent.

The next set of rows conditions on history. That is, instead of forecasting what the history will be given particular demographic characteristics, we will instead assume a particular sexual history. The second set of rows assumes a history of no sex while the third set assumes the person had sex in the previous period (age 15) and used episode-specific contraception. The differences across the second and third set of rows are quite large. While an individual who had an intact family at age sixteen would abstain eighty-three percent of the time conditional on abstaining in the past, the probability that a similar individual who had sex with episode-specific contraception in the previous period is less than forty-four percent. Habit persistence is much more important in determining sexual activity than having a praying mother, an intact family, or a working mother.

These estimated effects and the policy simulations conducted in the next section are not informative if the model does a poor job of fitting the data. Using the sample of those aged fourteen, we forecast the sex choices and fertility outcomes and see how well this matches the trends in the data. The model predictions for ages 14 through 19 are shown in Table 10. Although we would expect to match the trends given the full set of age interactions, we are forecasting ahead with a particular subset of individuals. With the exception of slightly over-predicting sex for those who are fifteen or sixteen, the predictions match the data quite well.

6 Policy Simulations

Given the model matches the predicted choices of sex and contraception use reasonably well, we now turn to policy simulations that examine the effects of changes in access to contraception for teens. In particular, we forecast the sex and contraception decisions and consequent pregnancy outcomes for sixteen year olds both when the policy is initially put into place (and thereby surprising the current sixteen year olds) and in the next two years after the policy. We use the characteristics of the fourteen year olds for the simulations. Hence, in year three of the policy sixteen year olds will have been exposed to the policy since they were fourteen.

We focus on two hypothetical policies. The first policy simulates the effects of increases in access to episode-specific contraception. This could be through ad campaigns that encourage the use of condoms or through making condoms available in school bathrooms, both of which lower the effective costs of using condoms. This is accomplished by raising the utility of episode-specific contraception. The second policy simulates the effects of decreasing access

Table 9: Choice Probabilities For Those Aged 16 Under Different Demographics[†]

		Mother Prays		Family Intact		Mother Works		Math Score	
		Yes	No	Yes	No	Yes	No	75th	25th
Unconditional Probability	Abstaining	0.6652	0.5970	0.6872	0.5695	0.6241	0.6592	0.6539	0.6269
	Unprotected Sex	0.1445	0.1685	0.1358	0.1785	0.1594	0.1461	0.1443	0.1609
	Episode-Specific Contraception	0.1378	0.1696	0.1282	0.1821	0.1566	0.1410	0.1462	0.1534
	Scheduled Contraception	0.0525	0.0650	0.0487	0.0699	0.0599	0.0538	0.0600	0.0600
Conditional on $c_{fc} = NS, c_{15} = NS$	Abstaining	0.8147	0.7655	0.8301	0.7502	0.7884	0.8065	0.8076	0.7884
	Unprotected Sex	0.0791	0.0976	0.0727	0.1036	0.0892	0.0819	0.0799	0.0902
	Episode-Specific Contraception	0.0901	0.1162	0.0826	0.1240	0.1038	0.0947	0.0955	0.1029
	Scheduled Contraception	0.0160	0.0208	0.0146	0.0222	0.0185	0.0169	0.0171	0.0184
Conditional on $c_{fc} = EC, c_{15} = EC$	Abstaining	0.4203	0.3638	0.4366	0.3449	0.3885	0.4122	0.4100	0.3901
	Unprotected Sex	0.1963	0.2090	0.1917	0.2142	0.2041	0.1973	0.1940	0.2069
	Episode-Specific Contraception	0.3447	0.3838	0.3342	0.3969	0.3662	0.3511	0.3562	0.3621
	Scheduled Contraception	0.0387	0.0434	0.0374	0.0449	0.0412	0.0395	0.0399	0.0409

[†]Forecasted from the characteristics of fourteen year-olds. Effects for intact family and mother works are the effects from having these characteristics at age 14.

Table 10: Comparing Model Predictions with the Data†

	Age=14		Age=15		Age=16		Age=17		Age=18		Age=19	
Probability of:	Model	Data	Model	Data	Model	Data	Model	Data	Model	Data	Model	Data
Abstaining	0.861	0.861	0.739	0.792	0.638	0.674	0.518	0.534	0.406	0.415	0.310	0.315
Unprotected Sex	0.067	0.067	0.105	0.086	0.154	0.141	0.209	0.204	0.271	0.264	0.335	0.327
Episode-specific	0.067	0.067	0.133	0.103	0.151	0.134	0.171	0.165	0.197	0.195	0.214	0.210
Scheduled	0.006	0.006	0.023	0.018	0.058	0.052	0.102	0.097	0.126	0.126	0.141	0.148

†Model predictions are calculated using the characteristics of fourteen year olds.

to all contraception, both episode-specific and scheduled. In this case the utility of using birth control is decreased (increasing the effective cost), though the utility of sex itself is left *unchanged*. An example of this could be curbing the distribution of contraceptives on school premises. Table 11 shows the percentage change in the four sex activity measures (no sex, sex without contraception, episode-specific contraception, and scheduled contraception) and the percentage change in the probability of becoming pregnant after the each of the policies have been implemented. The first panel of Table 11 is for the population while the second panel focuses on those individuals whose parental income in 1997 places them in the lowest income quartile.

We first focus on the simulations for the population. In the first year of the policy that increases access to episode-specific contraception, we see a drop in the teen pregnancy rate. This happens despite the substitution of some individuals from no sex to sex that occurs because of the 'moral hazard' generated by the lower costs of episode-specific contraception (which lower the risks of pregnancy). The overall rate of sex has increased but the drop in unprotected sex leads to lower pregnancy rates. However, this drop in teen pregnancy rates is only short-term. By year three of the policy teen pregnancy rates actually rise as more individuals at sixteen are having sex now due to the increased rate of sexual activity in the long run. These patterns are reversed when contraception is made less accessible. Namely, there is virtually no effect of the policy in year one on teen pregnancy rates. However, both abstinence and incidence of unprotected sex increase in the short run. The rates of abstinence increase due to the higher contraception costs, which is the converse of the moral hazard described earlier. Relative to year one, the rates of abstinence increase, while those of unprotected sex decrease, with each subsequent year due to habit persistence in teen sex. On the whole, individuals at age sixteen are less sexually active having been exposed to the policy for three years, so much so that lower pregnancy rates result.

One might suspect that optimal policies will depend upon the characteristics of the individuals at the school. There may be schools where the sex rate is so high that encouraging contraception lowers pregnancy rates both in the short and the long run because the induced entry into sex is small. The second panel performs the same simulation except now on those individuals who were below the 25th percentile of the income distribution.[23] Here we see

[23]The income distribution refers to the observed parental income distribution in the data from the first wave

Table 11: Short and Long Run Sex Behavior and Outcomes for 16 year olds†

			No Sex	Unprotected Sex	Episode-Specific Contraception	Scheduled Contraception	Pregnancy
	Increased	Year 1	-2.10%	-5.71%	17.35%	-8.62%	-0.46%
	Access to	Year 2	-3.18%	-4.74%	21.13%	-8.97%	1.04%
Population	Episode-Specific	Year 3	-3.73%	-4.09%	22.45%	-8.28%	1.86%
	Decreased	Year 1	2.66%	7.99%	-13.64%	-16.72%	-0.12%
	Access to	Year 2	3.92%	6.69%	-16.36%	-19.83%	-1.98%
	Contraception	Year 3	4.48%	5.65%	-17.35%	-20.86%	-3.00%
	Increased	Year 1	-2.22%	-6.23%	17.07%	-8.20%	-1.04%
	Access to	Year 2	-3.40%	-5.66%	20.41%	-8.93%	0.03%
Low Income	Episode-Specific	Year 3	-3.97%	-5.14%	21.42%	-8.35%	0.67%
	Decreased	Year 1	2.92%	8.72%	-12.78%	-15.52%	1.02%
	Access to	Year 2	4.27%	7.89%	-15.09%	-18.16%	-0.36%
	Contraception	Year 3	4.89%	7.11%	-15.88%	-19.03%	-1.15%

†Forecasted from the sample of 14 year-olds assuming the change in policy was a surprise. Percent changes use the forecasts with no policy as the base. For increased access to episode-specific contraception, we increase the intercept parameter in equation (6) by 0.2, leaving the fixed and transition costs unchanged. For decreased access to contraception, we lower the intercept parameters in equations (6) and (7) by 0.2. We experimented with different values for these changes and the qualitative results remained the same. Low income forecasts are done only for those individuals who had parents earnings at or below the 25th percentile of the income distribution in the data.

that the long run costs of increasing access to contraception are smaller than when looking at the population. However, the same patterns emerge: increasing access to episode-specific contraception, while attractive in the short run, increases pregnancies in the long run, with just the opposite occurring when access to contraception is decreased.

While our focus here has been on the crossing that may occur between short and long run pregnancy rates due to changes in contraception policy, other simulations resulted in lower pregnancy rates across the board. For example, we have also forecasted short and long run pregnancy rates when the effectiveness of birth control increases. Here, pregnancy rates fall both in the short run and in the long run. The fall, however, is smaller in the long run due to the increased sexual activity.

7 Conclusion

There is much persistence in teen sexual behavior. If this habit persistence arises from a moral or psychological barrier that has been crossed once an individual has sex for the first time (a fixed cost) or the first time in a relationship (a transition cost), programs that increase rates of teen sexual activity may lead to higher pregnancy rates in the long run than in the short run.

Our estimates show large transition and fixed costs to having sex. Persistence is also observed in using birth control methods like the pill, with smaller effects for condoms. The persistence in sexual activity is such that policies that affect access to contraception may have very different effects in the short run than the long run. Our results suggest that increasing access to contraception may actually increase long run pregnancy rates even though short run pregnancy rates fall. On the other hand, policies that decrease access to contraception, and hence sexual activity, may lower pregnancy rates in the long run. The primary purpose of our research is to illustrate the unintended consequences that my result if the dynamic aspects of teen decisions regarding sexual activity are ignored. In spite of the limitations that we discuss below, we believe that our work is important in showing that policy makers should be aware of such dynamic considerations when developing contraceptive policies.

It also needs to be emphasized that our focus is on teen sexual behavior and pregnancy

of the survey.

outcomes. Hence our conclusions are not necessarily applicable to older individuals. For example, Goldin and Katz (2002) provide evidence on the benefits of the availability to oral contraception to women of college going and older ages. In our analysis we also do not examine the effects of access to contraception on incidence of sexually transmitted diseases. This is another factor that could be important in determining appropriate policies regarding access to contraception, particularly condoms.

There are many other factors, however, that may also point towards increased access to contraception having negative consequences. For example, Akerlof, Yellen, and Katz (1996) argue that contraception and birth control changed the bargaining terms between men and women, and led to an increase in out-of-wedlock births. We also do not examine the effects of peer networks or multiplicity of sexual partners on teen sexual decisions and pregnancy outcomes, both of which may lead to greater access to contraception having negative effects in the long run. For example, we may see fixed costs in the form of a moral or psychological barrier the first time one has sex outside of a committed relationship. To the degree that increased access to contraception encourages experimentation outside the committed relationship, habit persistence may again lead to greater access to contraception increasing teen pregnancy rates. Future research that extends our analysis to incorporate factors such as sexually transmitted diseases, bargaining in relationships, and multiplicity of partners will improve our understanding of the consequences of increased access to contraception for teens.

An alternative explanation of why there is so much persistence in sexual activity in the data is that the individual-level heterogeneity that we model as permanent is actually time-varying. In this case, the persistence observed in the data would not be endogenous to past behaviors but would reflect exogenous taste shocks that may be persistent over time. The policy implications of these two explanations are very different. Under the time-varying heterogeneity only our short run policy simulations are relevant. Our data are not rich enough to distinguish between these two hypotheses. Moreover, with rare exceptions (e.g., Pakes 1987), the convention in the dynamic discrete choice literature has been to allow for serial correlation between observed variables but not between unobserved variables. As in our work, this is commonly done using the procedure proposed by Heckman and Singer (1983) that allows for permanent unobserved heterogeneity. In our context an additional empirical argument in favor of the permanent unobserved heterogeneity approach is that most of biological maturation has already occurred by age fourteen.[24]

Finally, in the spirit of Stigler and Becker[25] (1977) we have eschewed an explanation based on time varying unobserved heterogeneity. In making standard assumptions we find strong evidence of habit persistence. Although our findings that fixed costs are almost three times the time varying costs (Table 8) suggest that it is unlikely that introducing time varying unobserved heterogeneity will completely eliminate the role of fixed costs, a clear direction for future research would be to attempt to distinguish habit persistence from time-varying unobserved heterogeneity. If as our results suggest the persistence observed in the data is indeed behavior-driven, even if partially, then the long run implications of our simulations need to be considered seriously in the development of contraception policies given the potential for unintended consequences.

References

[1] **Adamczyk, A., and J. Felson.** "Friends' Religiosity and First Sex." *Social Science Research*, pp. 924-947, 35 (2006).

[2] **Akerlof, G.A., Yellen, J.L. and M.L. Katz** "An Analysis of Out-of-Wedlock Childbearing in the United States." *Quarterly Journal of Economics*, 111 (2:2002).

[3] **Arcidiacono, P.** "Affirmative Action in Higher Education: How do Admission and Financial Aid Rules Affect Future Earnings?" *Econometrica*, 73 (5:2005).

[4] **Arcidiacono, P., Sieg, H. and F. Sloan.** "Living Rationally Under the Volcano? An Empirical Analysis of Heavy Drinking and Smoking", *International Economic Review*, pp. 37-65, 48 (1:2007).

[5] **Ashby, S. L., Edmonson, M. B., and C. M. Arcari.** "Television Viewing and Risk of Sexual Initiation by Young Adolescents." *Archives of Pediatrics and Adolescent Medicine*, pp. 375-380, 160 (2006).

[24] Chumelea et al. (2003) found that 90% of U.S. girls had attained menarche by 13.75 years of age. Further, menarche is one of the late events of puberty (Zabin et al. 1986).

[25] E.g., Becker and Stigler state (p. 76) "On our view, one searches, often long and frustratingly, for the subtle forms that prices and incomes take in explaining differences among men and periods. If the latter approach yields more useful results, it is the proper choice."

[6] **Bearman, P.S. and H. Brückner.** "Promising the Future: Virginity Pledges and the Transition to First Intercourse." *American Journal of Sociology*, 106 (4:2001).

[7] **Bearman, P., Moody, J., and K. Stovel.** "Chains of Affection: The Structure of Adolescent Romantic and Sexual Networks." *American Journal of Sociology*, pp. 44-91, 110 (1: 2004).

[8] **Cameron, S. and J. Heckman.** "The Dynamics of Educational Attainment for Black, Hispanic, and White Males." *Journal of Political Economy*, 109 (3:2001).

[9] **Cameron, S. and J. Heckman.** "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males." *Journal of Political Economy*, 106 (2:1998).

[10] **Chumlea, W.C., Schubert, C.M., Roche, A.F., and H.E. Kulin.** "Age at Menarche and Racial Comparisons in U.S. Girls", *Pediatrics*, pp. 110-113, 111 (1: 2003).

[11] **Carro, J.M. and P. Mira.** "A Dynamic Model of Contraceptive Choice of Spanish Couples", *Journal of Applied Econometrics*, pp. 955-980, 21 (2006)

[12] **Day, R.D.** "The Transition to First Intercourse Among Racially and Culturally Diverse Youth" *Journal of Marriage and the Family*, pp. 749-62, 54(4:1992).

[13] **Delavande, A.** "Pill, Patch, or Shot? Subjective Expectations and Birth Control Choice" *International Economic Review*, pp. 999-1042, 49(3:2008).

[14] **Eckstein, Z. and K. Wolpin.** "Why Youths Drop out of High School: The Impact of Preferences, Opportunities, and Abilities" *Econometrica*, 67 (6:1999).

[15] **Furstenberg, F.F.** "When Will Teenage Childbearing Become a Problem? The Implications of Western Experience for Developing Countries" *Studies in Family Planning*, pp. 246-253, 29 (2:1998).

[16] **Furstenberg, F.F., Brooks,-Gunn, J., and S.P. Morgan.** "Race Differences in the Timing of Sexual Intercourse" *American Sociological Review*, pp. 511-18, 52 (1987).

[17] **Girma, S. and D. Paton.** "Matching estimates of the impact of over-the-counter emergency birth control on teenage pregnancy." *Health Economics*, pp. 1021-1032, 15 (10:2006).

[18] **Gilleskie, D. and K. Strumpf.** "The Behavioral Dynamics of Youth Smoking", *Journal of Human Resources* pp. 822-866, 40 (4:2005).

[19] **Goldin, C. and L. F. Katz.** "The Power of the Pill: Oral Contraceptives and Women's Career and Marriage Decisions." *Journal of Political Economy*, pp. 730-70, 110 (4:2002).

[20] **Heckman, J.** "Heterogeneity and State Dependence." *Studies in Labor Markets*, (Edited by S. Rosen). (University of Chicago Press, 1981).

[21] **Heckman, J. and B. Singer** "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica*, pp. 271–320. (52:1984)

[22] **Hotz, V. and R. Miller.** "Conditional Choice Probabilities and the Estimation of Dynamic Models." *Review of Economic Studies*, pp. 497-529, 60 (3:1993).

[23] **Jones, E.F. and J.D. Forrest** "Underreporting of Abortion in Surveys of U.S. Women: 1976 to 1988." *Demography*, pp. 113-126, 29 (1:1992).

[24] **Kane, T. and D. Staiger.** "Teen Motherhood and Abortion Access." *Quarterly Journal of Economics*, pp. 467-506, 111 (2:1996).

[25] **Keane, M.** "Modeling heterogeneity and State Dependence in Consumer Choice Behavior." *Journal of Business and Economic Statistics*, pp. 310-327, 15 (3:1997).

[26] **Keane, M. and K. Wolpin.** "The Career Decisions of Young Men." *Journal of Political Economy*, 105 (3:1997).

[27] **Khwaja, A.** "Health Insurance, Habits, and Health Outcomes: A Dynamic Stochastic Model of Investment in Health." Ph.D. Dissertation, University of Minnesota (2001).

[28] **Levine, P., Trainor, A., and D. Zimmerman.** "The Effect of Medicaid Abortion Funding Restrictions on Abortions, Pregannacies, and Births." *Journal of Health Economics*, pp. 555-578, 15 (5:1996).

[29] **Lundberg, S. and R. Plotnik.** "Adolescent Premarital Childbearing: Do Economic Incentives Matter?." *Journal of Labor Economics*, pp. 177-200, 13 (2:1995).

[30] **McFadden, D.** "Econometric Models of Probabilistic Choice." *Structural Analysis of Discrete Data with Econometric Applications*, (Edited by Manski, Charles and Daniel McFadden). (1981).

[31] **McFadden, D.** "Modelling the choice of Residential Location" in *Spatial Interaction Theory and Planning Models* (Edited by Karlqvist, Anders; Lundqvist, Lars; Snikcars, Folke; and Jorgen Weibull.) (New York, NY: North-Holland Publishing Company, 1978).

[32] **Oettinger, G.** "The Effects of Sex Education on Teen Sexual Activity and Teen Pregnancy." *Journal of Political Economy*, pp. 606-644, 107 (3:1999).

[33] **Paton, D.** "The Economics of Family Planning and Underage Conception." *Journal of Health Economics*, 21 (2:2002).

[34] **Rosenzweig, M. and Schultz, T.** "The Demand for and Supply of Births: Feritility and Its Life Cycle Consequences." *American Economic Review*, pp. 992-1015, 75 (5:1985).

[35] **Rust, J.** "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher." *Econometrica*, 55 (5:1987).

[36] **Singh, S. and Darroch, J.E.** "Adolescent pregnancy and childbearing: levels and trends in developed countries." *Family Planning Perspectives* pp. 14-23, 32 (1:2000).

[37] **Stigler, G. J. and G. S. Becker.** "De Gustibus Non Est Disputandum." *American Economic Review*, pp. 76-90, 67 (2: 1977).

[38] **The Alan Guttmacher Institute.** "U.S. Teenage Pregnancy Statistics: Overall Trends, Trends by Race and Ethnicity, and State-by-State Information." New York, NY: The Alan Guttmacher Institute, 2004.

[39] **van der Klaauw, W.** "On the Use of Expectations Data in Estimating Dynamic Behavioral Models", working paper (2000).

[40] **Walker, J.** "Pregnancy and Fertility Expectations: Estimates of Bounded Rationality and Unintended Births." working paper, 2003.

[41] **Widmer, E.D.** "Influence of Older Siblings on Initiation of Sexual Intercourse", *Journal of Marriage and the Family* pp. 928-938, 59 (4:1997).

[42] **Zabin, L.S., Smith, E.A., Hirsch, M.B., and J.B. Hardy** "Ages of Physical Maturation and First Intercourse in Black Teenage Males and Females", *Demography* pp. 595-605, 23 (4:1986).

[43] **Zelnik, M. and F.K. Shah** "First Intercourse Among Young Americans", *Family Planning Perspectives* pp. 64-70, 15 (2:1983).

Appendix

In this appendix we show the estimating equations and results for the transition parameters on whether or not one's mother works full-time, whether a divorce occurs, and whether the individual lives with his biological mother. This last measure is designed to capture whether the individual no longer lives at home without modeling every possible living arrangement. We assume that the state variables at time t depend only on the state variables at time $t-1$:

$$q(s_t|s_{t-1}) = q(s_t|s_{t-1}, s_{t-2}, ...)$$

We assume that each follows a logit process subject to the following restrictions:

1. Divorce is an absorbing state

2. No longer living with one's biological mother is an absorbing state

Since we cannot distinguish between full-time and part-time work for the mother in the first wave, we estimate the transitions using outcomes from waves 3-5 with the corresponding lagged values coming from waves 2-4.

Table 12 presents the estimates of the transition parameters. The most significant predictor of one's mother working full-time at time t is whether one's mother worked full-time at time $t-1$. Living with both biological parents reduces the probability of the mother working, though this effect is less than one-tenth the size of the lagged mother working effect. The effect of a praying mother is also negative, but smaller and only marginally significant. The coefficients on age and black are small and insignificant.

The probability of the biological family remaining intact at time t falls if the mother worked at time $t-1$. A mother who prayed regularly in 1997 increases the probability of the family remaining intact, while black families are significantly more likely to experience divorce. An intact family at time $t-1$ significantly lowers the probability an individual will leave home, as does begin black and having higher test scores. Not surprisingly, age has a strong positive effect on the probability of leaving home.

Table 12: Transition Parameters

	Mother Work		Intact Family		Leave Home	
	Coeff	Std Err	Coeff	Std Err	Coeff	Std Err
Lag Mother Work	3.351	0.083	-0.937	0.288	-0.164	0.118
Lag Intact Family	-0.263	0.087			-0.808	0.121
Black	-0.053	0.101	-0.816	0.294	-0.418	0.141
Age	-0.017	0.033	0.128	0.099	0.517	0.046
Mother Prays	-0.158	0.085	0.541	0.253		
Math Score (00's)					-0.465	0.177
Constant	-0.729	0.565	1.946	1.669	-10.78	0.839
Observations	4737		2642		5088	