Supplementary Material

Supplement to "Peer effects in sexual initiation: Separating demand and supply mechanisms"

(*Quantitative Economics*, Vol. 6, No. 3, November 2015, 663–702)

Seth Richards-Shubik
Department of Economics, Lehigh University

Appendix A: Approximations

A.1 *Arrival rate for model solution*

As noted in Section 2.2, it would be difficult to make an exact calculation for the expected arrival rate at randomly selected solution points so as to construct the value functions in (4). Instead, I approximate the decision rules among the opposite gender and use this approximation to simulate an expected value for the arrival rate.

This procedure uses auxiliary regressions that relate the probability of search among both virgins and nonvirgins in a group to the lagged nonvirginity rate for that group. These regressions are made separately for each gender and quarterly "age" in high school. To construct the regression coefficients, I start with initial values that allow the model to be solved and thus generate search probabilities for the entire sample, and then iterate with regressions of these search probabilities on the observed lagged nonvirginity rates.

To calculate the expected arrival rate ($\mathrm{E}_t \lambda_{it}$) to go into (4) or (7) with this approach, I first use the appropriate age-specific regression to assign a probability of search (not conditional on virginity status) to each person in the supply groups, based on the value of $Y_{k,t-1}$ for their group (k). Then I use a series of uniform draws to simulate their behavior several times, which gives a number of realizations for N_{it}. Finally I average over the resulting values of $\lambda(N_{it})$ to calculate an expected value for λ_{it}. The remainder of the solution algorithm for the individual problem is standard.

A.2 *Arrival rate for estimation*

I use Monte Carlo integration to approximate the expected arrival rate in (9), because the search decisions of virgins in the supply groups $S(i)$ are unobserved and depend on their individual shocks. The procedure is as follows. For each simulation round, $r \in 1, \dots, R$, and for each virgin, j, in the three supply groups, the preference type ω_j^r is drawn from the appropriate distribution. Then the search decision, d_{jt}^r, is simulated by comparing

Seth Richards-Shubik: sethrs@lehigh.edu

DOI: 10.3982/QE249

the type-specific search probability for individual j, given by $\Phi[\cdots]$ in (9), against a pseudorandom uniform draw. Combining these simulated search decisions of virgins with the known search behavior of nonvirgins (i.e., they all search) yields N^r_{it}. Then averaging $\lambda_{a_{it}}(N^r_{it})$ across simulation rounds produces an approximation for the expected arrival rate.

A.3 *Type distribution for virgins observed after ninth grade*

As noted in Section 4.1, I need to update the distribution of ω for cohorts that are first observed after ninth grade. To do this, I create approximate, type-specific hazard rates, which combine to give the probability, for each type, of still being a virgin when the individuals are first observed. Then I can use Bayes rule to update the distribution of ω, originally specified for ninth graders, to the appropriate grade.

The approximate, type-specific hazard rates are created with data from the younger cohorts at the individuals' schools, which relies on a steady state from one cohort to the next.[56] I regress the type-specific transition probabilities ($L_{it}(\omega)$) of the younger cohorts on their relevant state variables, which include the lagged nonvirginity rates by gender and grade ($Y_{m,t-1}$). I then use these regressions to predict type-specific hazard rates for the older cohorts before the observation period (i.e., when they were in earlier grades in high school). In these predictions, the current nonvirginity rates among younger cohorts substitute for the unobserved rates among older cohorts in previous years. The predicted hazard rates yield the probability of remaining a virgin for each type, and then I use Bayes rule to update the initial distribution of ω for each individual in the older cohorts.

The exact procedure is the following:

(i) Regress $L_{it}(\omega)$ on $Y_{m,t-1}$ and $\bar{x}_{s(i)}$, with separate approximations for each gender and age (i.e., quarter within grade).

(ii) Project Y_{m0} forward using the approximation ψ to create a sequence as long as the unobserved time span. For example, for someone first observed at the beginning of eleventh grade, Y_{m0} would be projected for 2 years (eight periods).

(iii) Predict $\widehat{L}_{it}(\omega)$ for the unobserved periods using the regressions from step (i) and the generated sequence of Y_{mt}, $t < 1$, from step (ii).

(iv) Define the approximate, type-specific probabilities of still being a virgin in the initial observation period as $\widehat{P}^0_i(\omega) \equiv \prod_{t=1-a_{i0}}^{0}[1 - \widehat{L}_{it}(\omega)]$.

(v) Finally, update the individual's type distribution with

$$\Pr\big(\omega_i = \omega^k | Y_{m0}, z_i, a_{i0}, y_{i0} = 0\big) = \frac{\widehat{P}^0_i(\omega^k) \cdot \pi_{k|Y_{m0},z_i}}{\sum_{l=1}^{\kappa} \widehat{P}^0_i(\omega^l) \cdot \pi_{l|Y_{m0},z_i}}.$$

[56]The steady-state assumption appears elsewhere, notably in the use of an aggregate law of motion estimated from current data to function as beliefs about the future.

The circularity in this procedure is resolved by starting with some initial guess for the regressions that produce $\widehat{L}$ and then iterating. In practice, these approximations converge very quickly (within three iterations).

Appendix B: Additional details on identification

B.1 *Stable preferences and the reflection problem*

Here I show how the assumption of stable preferences over time provides an identifying restriction that can address the reflection problem. First we need to modify the model so that there is a linear dependence between the effects of Y_{t-1} and $\psi(Y_{t-1})$ on behavior. Accordingly, suppose ψ is a linear function ($\psi(Y_t) = \psi Y_t$) and consider a once-and-for-all decision to initiate sex in period 2 with the arrival rate set equal to 1. This eliminates the nonlinearities. Then, for simplicity, let Y_t be a scalar and specify the heterogeneity as $\omega_i = a'x_i + b'\bar{x} + cY_0$, so that flow utility for a nonvirgin is $a'x_i + b'\bar{x} + cY_0 + dY_{t-1} + e_{it}$.[57] Then the expected payoff to searching in period 2 is the discounted sum

$$\sum_{t=2}^{\infty} \beta^{t-2}\big(a'x_i + b'\bar{x} + cY_0 + d\mathrm{E}_2[Y_{t-1}] + e_{it}\big),$$

where $\mathrm{E}_2[Y_{t-1}]$ denotes the fully rational expectation for Y_{t-1} given the information set at period 2 (which includes Y_1). The approximation replaces $\mathrm{E}_2[Y_{t-1}]$ with $\psi^{t-2}Y_1$, and so the expression above simplifies to

$$(1-\beta)^{-1}\big(a'x_i + b'\bar{x} + cY_0\big) + (1-\beta\psi)^{-1}dY_1$$

(plus the mean-zero e_{it}). This determines the initiation hazard in period 2. Therefore, with ψ identified from the distribution of (Y_1, Y_2) and an assumed value for β, the parameters a, b, c, and d are identified.

The difference with Manski's (1993) result is that the relationship between the effects of Y_1 and $\psi^{t-2}Y_1$ on behavior are known. The latter equals the former scaled by powers of β under the assumption of stable preferences. By contrast, in Manski's static model, and the models considered in Brock and Durlauf (2001b) and Blume et al. (2011), the common group variable is essentially arbitrary, call it X_g. Because there is no theory to inform the relationship between the effects of X_g and $\mathrm{E}[Y|X_g]$ on behavior, there is another parameter to estimate, which creates the identification problem.

B.2 *Use of initial nonvirginity rates versus school fixed effects*

To see that conditioning on initial nonvirginity rates is equivalent to including a school (i.e., location) fixed effect in the types distribution, it is easiest to work with a simplified model that expresses only the composite effect of social interactions from a single reference group. Suppose the expected value of the *outcome* (not utility, as in the structural model) is

$$\mathrm{E}[y_{imt}|a_{it}, Y_{imt}, \omega_i] = g\big(\alpha a_{it} + \gamma\mathrm{E}[Y_{imt}|\mathbf{x}_m, \eta_m] + \omega_i\big),$$

[57] Blume et al. (2011) consider similar payoffs in a dynamic linear model (pp. 869–870).

where g is a known strictly increasing function, m denotes the school or, equivalently, the community where it is located, and η_m is the unobserved location-based effect.[58] This says that the expected value of an individual's virginity status at time t is a function of their age (a_{it}), the expected nonvirginity rate in their reference group ($\mathrm{E}[Y_{imt}|\mathbf{x}_m, \eta_m]$), and the combined effect of some permanent factors (ω_i).[59] Further suppose that $\omega_i \in \{\omega^L, \omega^H\}$ and $\Pr(\omega_i = \omega^H | x_i, \eta_m) = h(x_i'\beta + \eta_m)$, where h is also strictly increasing.

The expected value of Y_0 (i.e., when $a_{i0} = 0$) in school m is

$$\begin{aligned}\mathrm{E}[Y_0|m] = \mathrm{E}\big[&g\big(\gamma \mathrm{E}[Y_0|m] + \omega^L\big) + h\big(x'\beta + \eta_m\big)\big[g\big(\gamma \mathrm{E}[Y_0|m] + \omega^H\big) \\ &- g\big(\gamma \mathrm{E}[Y_0|m] + \omega^L\big)\big]|m\big],\end{aligned}$$

where m is shorthand for the available information (so the outer expectation on the right-hand side is taken over the distribution of x at school m). From the monotonicity of h, this can be solved uniquely for η_m given $\mathrm{E}[Y_0|m]$ and the observable distribution of x (and of course the common parameters).[60] The intuition is fairly clear. The distribution of x at a school would predict a certain value for Y_0. If the observed value of Y_0 is higher or lower, this must be due to a common unobserved factor (apart from random noise, which vanishes as the sample gets large). The conclusion is that observation of Y_0 and the distribution of x at each school identifies an unobserved location-based effect in the distribution of preference types. As a consequence, the function for the probability of being "high" type, $h(x_i'\beta + \eta_m)$, can be replaced with a function of these observables, say $f(x_i, \mathbf{x}_m, Y_{m0})$. The third version of the logit in (6) serves as an approximation to this function, which follows the strategy proposed in Heckman (1981).

[58] This assumes there is only one high school per community in the data, which is the case.

[59] Notice that this uses contemporaneous group outcomes rather than lagged group outcomes. This is a common specification, and I use it here to simplify the expressions. As explained in Section 4.2.2, this timing does not impact the identification of my model.

[60] The other parameters would be identified with data from a second time period.

Table A.1. Logit hazard models for sexual initiation with same-gender peer group alone.

	Boys						Girls					
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Nonvirginity rates												
Peer group lag	1.37	0.59	0.53	0.62	0.54	0.42	1.25	0.57	0.51	0.49	0.29	−0.36
($Y_{i,t-1}$)	(0.22)	(0.42)	(0.42)	(0.42)	(0.43)	(0.32)	(0.19)	(0.32)	(0.32)	(0.33)	(0.33)	(0.29)
Initial rates incl. ($Y_{i,0}$ and $Y_{s(i),0}$)	no	yes	yes	yes	yes	no	no	yes	yes	yes	yes	no
Individual characteristics												
Age	0.09	0.04	0.06	0.04	0.05	0.18	0.01	−0.02	0.00	0.00	0.02	0.21
	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)
Black			0.18	0.29	0.28	0.33			0.02	0.10	0.10	0.10
			(0.09)	(0.10)	(0.10)	(0.11)			(0.07)	(0.09)	(0.09)	(0.09)
Parent is college-educ.			−0.30	−0.30	−0.31	−0.36			−0.30	−0.28	−0.28	−0.29
			(0.07)	(0.08)	(0.08)	(0.08)			(0.07)	(0.07)	(0.07)	(0.07)
Younger child			0.03	0.04	0.04	0.04			−0.02	−0.02	−0.02	−0.01
			(0.07)	(0.07)	(0.07)	(0.07)			(0.07)	(0.07)	(0.07)	(0.07)
Only child			0.17	0.17	0.17	0.14			0.41	0.40	0.40	0.42
			(0.10)	(0.10)	(0.10)	(0.10)			(0.08)	(0.08)	(0.08)	(0.08)
Peer means incl.	no	no	no	yes	yes	yes	no	no	no	yes	yes	yes

(*Continues*)

Table A.1. *Continued.*

	Boys						Girls					
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
School policies												
Sex ed in grades 9/10					−0.09						0.01	
					(0.13)						(0.11)	
Sex ed in grades 11/12					0.03						0.22	
					(0.07)						(0.07)	
Fam. planning counseling at school					−0.17						−0.25	
					(0.09)						(0.09)	
No referrals for fam. planning					−0.13						−0.03	
					(0.09)						(0.08)	
Day care for children of students					0.11						−0.01	
					(0.13)						(0.11)	
For-credit courses in parenting					−0.00						0.01	
					(0.07)						(0.07)	
School fixed effects	no	no	no	no	no	yes	no	no	no	no	no	yes
Observations	21,671	21,671	21,671	21,671	21,671	21,504	24,116	24,116	24,116	24,116	24,116	24,045

Note: Standard errors are given in parentheses.

TABLE A.2. Logit hazard models for sexual initiation with peer and supply groups combined.

	Boys						Girls					
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Nonvirginity rates												
Peer & supply lag	1.74	0.74	0.69	0.79	0.70	0.69	1.51	0.64	0.60	0.60	0.41	−0.18
($Y_{i,t-1}$)	(0.24)	(0.44)	(0.44)	(0.45)	(0.45)	(0.40)	(0.21)	(0.39)	(0.40)	(0.41)	(0.41)	(0.36)
Initial rates incl. ($Y_{i,0}$ and $Y_{s(i),0}$)	no	yes	yes	yes	yes	no	no	yes	yes	yes	yes	no
Individual characteristics												
Age	0.05	0.04	0.06	0.03	0.04	0.15	−0.01	−0.03	−0.00	−0.00	0.02	0.19
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.05)
Black			0.18	0.29	0.28	0.33			0.02	0.10	0.10	0.10
			(0.09)	(0.10)	(0.10)	(0.11)			(0.07)	(0.09)	(0.09)	(0.09)
Parent is			−0.30	−0.30	−0.31	−0.36			−0.30	−0.28	−0.28	−0.29
college-educ.			(0.07)	(0.08)	(0.08)	(0.08)			(0.07)	(0.07)	(0.07)	(0.07)
Younger child			0.03	0.04	0.04	0.04			−0.02	−0.02	−0.02	−0.01
			(0.07)	(0.07)	(0.07)	(0.07)			(0.07)	(0.07)	(0.07)	(0.07)
Only child			0.17	0.17	0.17	0.14			0.41	0.40	0.40	0.42
			(0.10)	(0.10)	(0.10)	(0.10)			(0.08)	(0.08)	(0.08)	(0.08)
Peer means incl.	no	no	no	yes	yes	yes	no	no	no	yes	yes	yes

(*Continues*)

Table A.2. *Continued.*

	Boys						Girls					
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
School policies												
Sex ed in grades 9/10					−0.10						0.01	
					(0.13)						(0.11)	
Sex ed in grades 11/12					0.02						0.22	
					(0.07)						(0.07)	
Fam. planning counseling at school					−0.16						−0.25	
					(0.09)						(0.09)	
No referrals for fam. planning					−0.13						−0.03	
					(0.09)						(0.08)	
Day care for children of students					0.11						−0.01	
					(0.13)						(0.11)	
For-credit courses in parenting					−0.00						0.01	
					(0.07)						(0.07)	
School fixed effects	no	no	no	no	no	yes	no	no	no	no	no	yes
Observations	21,671	21,671	21,671	21,671	21,671	21,504	24,116	24,116	24,116	24,116	24,116	24,045

Note: Standard errors are given in parentheses.

TABLE A.3. Logit hazard models for sexual initiation with peer and supply groups separately.

Variable	Boys						Girls					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Nonvirginity rates												
Peer group lag	1.00	0.42	0.36	0.45	0.38	0.35	0.95	0.55	0.47	0.43	0.22	−0.40
($Y_{i,t-1}$)	(0.24)	(0.46)	(0.46)	(0.46)	(0.46)	(0.32)	(0.22)	(0.35)	(0.35)	(0.36)	(0.36)	(0.29)
Supply grp. lag	0.75	0.34	0.33	0.35	0.33	0.34	0.58	0.05	0.11	0.15	0.19	0.19
($Y_{s(i),t-1}$)	(0.23)	(0.37)	(0.37)	(0.37)	(0.37)	(0.30)	(0.20)	(0.39)	(0.39)	(0.39)	(0.39)	(0.28)
Initial rates incl. ($Y_{i,0}$ and $Y_{s(i),0}$)	no	yes	yes	yes	yes	no	no	yes	yes	yes	yes	no
Individual characteristics												
Age	0.05	0.04	0.06	0.03	0.04	0.15	−0.02	−0.02	0.00	−0.00	0.02	0.19
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.05)
Black			0.18	0.29	0.28	0.33			0.02	0.10	0.10	0.10
			(0.09)	(0.10)	(0.10)	(0.11)			(0.07)	(0.09)	(0.09)	(0.09)
Parent is		−0.30	−0.30	−0.31	−0.36			−0.30	−0.28	−0.28	−0.29	
college-educ.			(0.07)	(0.08)	(0.08)	(0.08)			(0.07)	(0.07)	(0.07)	(0.07)
Younger child			0.03	0.04	0.04	0.04			−0.02	−0.02	−0.02	−0.01
			(0.07)	(0.07)	(0.07)	(0.07)			(0.07)	(0.07)	(0.07)	(0.07)
Only child			0.17	0.17	0.17	0.14			0.41	0.40	0.40	0.42
			(0.10)	(0.10)	(0.10)	(0.10)			(0.08)	(0.08)	(0.08)	(0.08)
Peer means incl.	no	no	no	yes	yes	yes	no	no	no	yes	yes	yes

(*Continues*)

Table A.3. *Continued.*

	Boys						Girls					
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
School policies												
Sex ed in grades 9/10					−0.09						0.01	
					(0.13)						(0.11)	
Sex ed in grades 11/12					0.02						0.22	
					(0.07)						(0.07)	
Fam. planning counseling at school					−0.16						−0.25	
					(0.09)						(0.09)	
No referrals for fam. planning					−0.13						−0.03	
					(0.09)						(0.08)	
Day care for children of students					0.11						−0.01	
					(0.13)						(0.11)	
For-credit courses in parenting					−0.00						0.01	
					(0.07)						(0.07)	
School fixed effects	no	no	no	no	no	yes	no	no	no	no	no	yes
Observations	21,671	21,671	21,671	21,671	21,671	21,504	24,116	24,116	24,116	24,116	24,116	24,045

Note: Standard errors are given in parentheses.

TABLE A.4. Parameter estimates with alternative specifications.

	Boys			Girls		
	Var's in Types Dist.		Alt. Match	Var's in Types Dist.		Alt. Match
Parameter	$z=(\cdot)$ (1)	$z=(x,\text{bar})$ (2)	(3)	$z=(\cdot)$ (4)	$z=(x,\text{bar})$ (5)	(6)
α (age)	0.082	0.063	0.098	0.167	0.166	0.173
	(0.063)	(0.059)	(0.082)	(0.039)	(0.037)	(0.081)
γ (peer norms)	0.203	0.227	0.224	0.204	0.213	0.213
	(0.071)	(0.067)	(0.080)	(0.054)	(0.052)	(0.082)
Arrival rate						
λ_{01} (9th grade)	−2.627	−2.644	−2.247	−2.368	−2.327	−2.179
	(0.321)	(0.314)	(0.201)	(0.248)	(0.236)	(0.177)
λ_{02} (10th grade)	−2.867	−2.880	−2.270	−2.414	−2.388	−2.144
	(0.315)	(0.295)	(0.150)	(0.252)	(0.239)	(0.149)
λ_{03} (11th grade)	−2.906	−2.944	−2.211	−2.439	−2.440	−2.247
	(0.340)	(0.321)	(0.145)	(0.246)	(0.234)	(0.129)
λ_{04} (12th grade)	−2.850	−2.938	−2.208	−2.382	−2.383	−2.297
	(0.349)	(0.334)	(0.149)	(0.251)	(0.240)	(0.131)
λ_{11} (same grade)	0.689	0.782	−0.007	0.252	0.185	0.061
	(0.403)	(0.382)	(0.075)	(0.289)	(0.269)	(0.082)
λ_{12} (below/above)	0.202	0.104	−0.003	−0.003	0.051	−0.080
	(0.258)	(0.254)	(0.023)	(0.160)	(0.156)	(0.085)
λ_{13} (above/2 above)	0.182	0.171	0.049	0.197	0.166	0.069
	(0.187)	(0.184)	(0.073)	(0.157)	(0.153)	(0.088)
Type values						
ω^L (low type)	−0.260	−0.243	−0.301	−0.284	−0.302	−0.318
	(0.124)	(0.058)	(0.118)	(0.054)	(0.048)	(0.090)
ω^H (high type)	−0.118	−0.145	−0.115	−0.103	−0.099	−0.037
	(0.079)	(0.060)	(0.074)	(0.055)	(0.048)	(0.089)
$\nu(\omega^L)$ (terminal val.)	−1.575	−1.436	−1.556	−0.139	−0.327	−0.473
	(1.161)	(0.745)	(0.931)	(0.420)	(0.377)	(0.832)
$\nu(\omega^H)$ (terminal val.)	0.158	−1.030	0.021	1.329	0.671	1.383
	(1.131)	(0.806)	(1.118)	(0.637)	(0.747)	(1.628)
Type probabilities (π^H)						
Constant term	0.694	−6.303	0.154	−0.722	−0.572	−0.754
	(1.365)	(2.873)	(0.893)	(0.635)	(0.710)	(0.617)
Y_0: 9th grade (own gender)	1.074	16.912	1.632	1.773	0.565	1.478
	(2.482)	(8.453)	(2.307)	(1.419)	(1.423)	(1.582)
Y_0: 9th grade (opposite gender)	−0.191	−7.237	1.194	3.794	4.436	3.860
	(2.274)	(5.492)	(1.939)	(1.523)	(1.708)	(1.689)

(*Continues*)

TABLE A.4. *Continued.*

	Boys			Girls		
	Var's in Types Dist.		Alt. Match	Var's in Types Dist.		Alt. Match
Parameter	$z = (\cdot)$ (1)	$z = (x, \text{bar})$ (2)	(3)	$z = (\cdot)$ (4)	$z = (x, \text{bar})$ (5)	(6)
Individual characterisitics						
Black		13.960	2.774		0.679	0.073
		(6.669)	(2.078)		(0.540)	(0.409)
Younger child		1.449	0.612		−0.165	−0.055
		(1.163)	(0.533)		(0.364)	(0.324)
Only child		7.421	1.661		2.789	2.258
		(3.909)	(1.210)		(1.078)	(1.004)
Parent educ.		−6.369	−1.937		−2.117	−1.750
		(3.175)	(1.227)		(0.848)	(0.833)
Group means						
Black		−17.527			−1.485	
		(7.630)			(1.052)	
Younger child		8.372			0.044	
		(5.078)			(1.230)	
Only child		9.083			1.692	
		(6.414)			(1.660)	
Parent educ.		5.208			−0.495	
		(3.758)			(1.062)	

Note: Standard errors are given in parentheses.

TABLE A.5. Approximation of equilibrium beliefs (nonvirginity rates).

	Boys		Girls	
	Observed	Simulated	Observed	Simulated
Grade intercepts				
9th grade	0.014	0.019	0.022	0.024
	(0.003)	(0.003)	(0.004)	(0.004)
10th grade	0.013	0.022	0.023	0.022
	(0.004)	(0.004)	(0.005)	(0.004)
11th grade	0.014	0.028	0.018	0.017
	(0.005)	(0.005)	(0.005)	(0.005)
12th grade	0.015	0.023	0.024	0.017
	(0.005)	(0.005)	(0.004)	(0.005)
Peer group nonvirginity rate				
Linear term	1.048	1.019	1.047	1.050
	(0.014)	(0.015)	(0.015)	(0.016)
Squared term	−0.089	−0.050	−0.088	−0.086
	(0.014)	(0.015)	(0.016)	(0.015)
Supply groups nonvirginity rates				
Group 1	0.020	0.008	0.012	0.014
(same grade)	(0.006)	(0.006)	(0.007)	(0.007)
Group 2†	0.008	0.007	0.007	0.006
	(0.006)	(0.006)	(0.005)	(0.005)
Group 3†	0.001	−0.001	0.001	−0.002
	(0.004)	(0.004)	(0.006)	(0.005)
R^{2*}	0.965	–	0.960	–
N	2079	2079	2084	2084
SUR test statistic (obs. vs. sim.)		13.69		
P-value (χ^2, 18 d.f.)		0.75		

Note: Standard errors are given in parentheses. * R^2 calculated from separate regressions by gender, with constant term and no 9th grade dummy. † Group 2 is grade below for boys and grade above for girls. Group 3 is grade above for boys and two above for girls.

ADDITIONAL REFERENCE

Heckman, J. J. (1981), "The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process." In *Structural Analysis of Discrete Data with Econometric Applications* (Manski, C. F. and D. L. McFadden, eds.), 179–195, MIT Press, Cambridge. [4]

Submitted January, 2012. Final version accepted September, 2014.