tapraid1/z3y-aer/z3y-aer/z3y00506/z3y2008d06a | yodert | S=42 | 10/12/06 | 13:56 | Art: 20031016 | Input-dld(dld)

Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility

By Petra E. Todd and Kenneth I. Wolpin*

This paper uses data from a randomized social experiment in Mexico to estimate and validate a dynamic behavioral model of parental decisions about fertility and child schooling, to evaluate the effects of the PROGRESA school subsidy program, and to perform a variety of counterfactual experiments of policy alternatives. Our method of validation estimates the model without using post-program data and then compares the model's predictions about program impacts to the experimental impact estimates. The results show that the model's predicted program impacts track the experimental results. Our analysis of counterfactual policies reveals an alternative subsidy schedule that would induce a greater impact on average school attainment at similar cost to the existing program. (*JEL* I21, I28, J13, O15)

This paper studies a large-scale government program in Mexico, the PROGRESA program, designed to foster investment in children's human capital and to alleviate poverty.[1] One of the program's major goals is to increase schooling levels by providing substantial payments to parents which are contingent on their children's regular attendance at school. To evaluate the PROGRESA program, the Mexican government conducted a randomized social experiment, in which 506 rural villages (in 7 of the 31 states) were randomly assigned to either participate in the program or serve as controls. Previous studies that assessed the impact of the program by comparing outcomes for treatments and controls demonstrated significant gains in school attendance.[2] Experimental treatment effects, measured as the difference in average attendance rates of children in the treatment and control villages one year after the program, ranged from about 5 to 15 percentage points depending on age and sex.

In the PROGRESA experiment, all eligible treatment group households were offered the same school attendance subsidy schedule, which depended on the child's grade level and sex. For that reason, prior evaluations of the program based on a straightforward comparison of the treatment and control groups assessed the impact only of the particular subsidy scheme that was implemented. Obviously, the choice of subsidy scheme is a critical design element of the program, inherently limiting what can be learned from the experiment. For example, it is not possible from the PROGRESA experiment alone to determine the size and structure of the subsidy that achieves the policy goals at least cost, nor is it possible to assess the many alternative policy tools available to achieve the same goals.

* Todd: Department of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104 (e-mail: petratodd@gmail.com); Wolpin: Department of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104 (e-mail: wolpink@ssc.upenn.edu). We thank the National Science Foundation for support under grant SES-0111593 and the Population Studies Center at the University of Pennsylvania for computer support. We are grateful to PROGRESA (now Oportunidades) and to INSP for making the data available to us. We thank Monica Orozco, Daniel Hernandez, Santiago Levy, Susan Parker, and Iliana Yaschine for help in answering questions about the data. We also thank Jere Behrman, the editor, and the referees for helpful comments.

[1] PROGRESA stands for Programa de Educacion, Salud, y Alimentacion (Program of Education, Health, and Nutrition). The program initially targeted poor families living in rural areas, but has since, under the new name of Oportunidades, expanded into semi-urban and urban areas and now serves about five million families. Similar programs exist in Bangladesh, Brazil, Chile, Colombia, Guatemala, Nicaragua, and Pakistan.

[2] See, for example, Jere Behrman et al. (2005), José Gomez de Leon and Susan Parker (2000), Parker and Emmanuel Skoufias (2000), and T. Paul Schultz (2000).

1 LINE LONG

tapraid1/z3y-aer/z3y-aer/z3y00506/z3y2008d06a | yodert | S=42 | 10/12/06 | 13:56 | Art: 20031016 | Input-dld(dld)

The main aim of this paper is to compare the efficacy of the PROGRESA program with that of alternative policies that were not implemented as part of the experiment. An ex ante evaluation of these alternative policies requires a theoretical framework that can be used to extrapolate from existing sample variation. Our evaluation is based on the structural estimation of a dynamic behavioral model of parental decision-making about children's schooling and family fertility. Lacking consensus on what is a reasonable model structure, we pursue a strategy that specifically exploits the experiment that was conducted as part of the PROGRESA program.[3] In particular, we estimate our model using households that did not receive the subsidy, and evaluate the performance of the model in forecasting the behavior of the treated households. Thus, a second aim of the paper is to explore the usefulness of social experiments as a tool for model validation.[4]

The schooling-fertility model developed in this paper builds on several literatures, including the static quality-quantity fertility models of Robert J. Willis (1973) and Gary S. Becker and H. Greg Lewis (1973), static models of intrahousehold allocation of resources to children, as in Behrman et al. (1986), and dynamic fertility models, as in Wolpin (1984) and Joseph Hotz and Robert A. Miller (1993).[5] In our model, married couples are assumed to make sequential decisions over a finite horizon about the time allocation of all of their children age 6 through 15, including their school attendance and labor market participation, and about the timing and spacing of births. Parents receive utility from the stock of children and their current ages, their current schooling levels and attendance, and from their leisure time (home production). Household consumption, which also yields utility, is enhanced by children's earnings. Parents' income and child wage offers are taken as exogenous. Children's wages are assumed to depend on distance to the nearest largest city, which provides an important source of identification. Preferences, parental income, and child wages differ permanently across households according to unobservable types and are subject to time-varying stochastic shocks.

The model is quite rich, allowing for complex dynamic interactions in parental decision-making. For example, the value of having older girls at home may be greater if there are very young children in the household. The model also allows for a psychic cost of attending school that may be higher when a child is behind in school for his age. The existence of this psychic cost implies that forward-looking parents may forego having a child work when faced with a high child wage offer that is transitory.

Our out-of-sample validation first compares the actual post-program school attendance rates of the children in treated households to the rates predicted by the model based on simulating the introduction of the subsidy schedule. Comparisons for several subgroups, distinguished by age, sex, and grade level, show that the predicted and actual attendance rates are close, ranging from within 1 percentage point for the 12–15-year-olds of either sex to 7 to 8 percentage points for children age 13 to 15 who had completed 6 or more years of school and who were behind in school. As a further validation, we predict what the treatment effect would have been for the nontreated households and compare it to the experimental estimate based on the treated households. The model performs well in estimating the treatment effect for girls, but less satisfactorily for boys.

In our judgment, the results of the out-of-sample validation provide enough confidence in the model to use it to perform an evaluation of the benefits and costs of alternative programs as a comparison to PROGRESA. Our evaluation of the PROGRESA program itself differs from existing studies in that we forecast the impact of the program on completed schooling for households that would be subject to the program beginning at marriage and extending over their entire lifetimes. Our estimates take into account both changes in family size induced by the

Fn3 [3] Exemplifying this lack of consensus, Orazio Attanasio et al. (2001) estimate a very different behavioral model using the same data.

Fn4 [4] A similar strategy of using a social experiment as a validation tool was previously adopted by David A. Wise (1985), who compared predicted impacts of a housing subsidy program based on a model of housing demand to impacts derived from a randomized experiment. Relatedly, Daniel McFadden and Antti P. Talvitie and Associates (1997) compared the predictions of a model of travel mode choice, estimated prior to the introduction of a new rapid transit system, to actual usage after its introduction. More recently, Jeremy Lise et al. (2003) use a Canadian welfare experiment to validate a model of job search.

Fn5 [5] For recent surveys of these literatures, see Behrman (1997) and Hotz et al. (1997).

program and short-run rigidities in adjustment arising, for example, from psychic costs of attending school when a child has fallen behind.

With the current PROGRESA subsidy, we estimate that completed schooling for both boys and girls will increase, on average, by about one-half year and that the program will cost about 26,000 (1997) pesos per eligible family. Because attendance, in the absence of any subsidy, is almost universal through the elementary school ages, subsidizing attendance at the lower grade levels, as under the existing program, is essentially an income transfer. We determine that a change in the subsidy schedule that eliminates the subsidy to attending grades 3 to 5 and increases the amount of the subsidy to grades 6 to 9 by about 50 percent leaves the overall cost of the program unchanged and produces an increase in average completed schooling by about an additional 0.1 years. We also calculate the effects on completed schooling of bonus payments for completing graduation milestones, building schools to reduce transportation costs, providing a pure income transfer, and strictly enforcing child labor laws.

In the next section, we provide further details about the PROGRESA program, followed in Section II by a description of the data used in the estimation. Section III presents the model and estimation method and Section IV the results of the estimation, including an assessment of the model's within-sample fit. Section V provides evidence on the model's ability to forecast the impact of the program. Section VI evaluates the impact of the PROGRESA program and alternative programs on completed schooling and fertility. The last section concludes with a summary and a discussion of broader methodological issues.

I. The PROGRESA Program

The data gathered as part of the PROGRESA experiment contain information on household demographics, income, school attendance, and on the employment and wages of children. Data are available for all households located in the 320 villages randomly assigned to the treatment group and in the 186 villages assigned to the control group.[6] The data we analyze were gathered through two baseline surveys administered October 1997 and March 1998, and through three follow-up surveys administered October 1998, May 1999, and November 1999. Supplemental data were also gathered at the village level, including the travel distance to the nearest secondary school and to the nearest city. Data collection was exhaustive within each village and included children from both eligible and ineligible families. In the baseline survey, there were 9,221 separate households in the control villages and 14,856 in the treatment villages. Fn6

Within treatment localities, only households that satisfy program eligibility criteria receive the school subsidies, where eligibility is determined on the basis of a "marginality index" designed to identify the poorest families within each community.[7] The benefit levels that families receive under the program represent about one-fourth of average family income (Gomez de Leon and Parker, 2000). Given its generosity, most families deemed eligible decide to participate in the program to some extent.[8] Under program rules, parents receive subsidies for each grade-eligible child who attends school at least 85 percent of the time, up to a family maximum. Fn7 Fn8

The PROGRESA subsidy increases with the child's grade level, up to grade 9, to offset the greater opportunity costs of schooling for older children who are more likely to engage in household production or market work. As seen in Table 1, the subsidy is greatest for children in junior secondary school (grades 7 through 9) and is slightly higher for girls, who traditionally have lower secondary school enrollment levels.[9] T1 Fn9

[6] The 506 localities were selected in a stratified random sampling procedure from high-poverty localities. Behrman and Todd (2000) provide evidence that the treatment and control groups are highly comparable prior to the initiation of the program.

[7] Program eligibility is based in part on discriminant analysis applied to the October 1997 household survey data. The discriminant analysis uses information on household composition, household assets (such as whether the house had a dirt floor), and children's school attendance, among other factors.

[8] The program also provides monetary aid and nutritional supplements for infants and small children that are not contingent on schooling. More than 75 percent of the transfer is due to the school subsidy.

[9] Prior to 1992, Mexico had compulsory schooling that required children to complete at least six years. In 1992, the law was changed to require nine years. As our data show, however, the vast majority complete fewer than nine years.

TABLE 1—MONTHLY TRANSFERS FOR SCHOOL ATTENDANCE UNDER THE PROGRESA PROGRAM

School level	Grade	Monthly payment in pesos: Females	Males
Primary	3	70	70
	4	80	80
	5	105	105
	6	135	135
Secondary	1	210	200
	2	235	210
	3	255	225

Source: Schultz (1999a, Table 1). Corresponds to first term of the 1998–1999 school year.

II. Variable Definitions and Descriptive Statistics

A. *Variable Definitions*

Our estimation sample consists of landless nuclear households in which there was a woman under the age of 50 reported to be the spouse of the household head.[10] The restriction to landless households reduced the original sample to 5,602 households, and the further restrictions to 3,410 households. An additional 209 households were dropped for data-related problems, such as the appearance of additional adult household members in later rounds of the survey. Of the remaining 3,201 households, 1,316 households are in the control villages and 1,885 are in the treatment villages. As of 1997, there were 4,012 children born to the control households and 5,561 to the treatment households. Of these, 1,958 children in the control villages and 2,694 in the treatment villages were between the ages of 6 and 15 as of the October 1997 survey. In comparison to the entire sample, landless households tend to be poorer and, therefore, are more likely to be eligible for the program. As of the 1997 survey, 52 percent of all households were eligible to participate, while 62 percent of the landless households were eligible. In estimating the behavioral model, we use data on both program eligible and ineligible households. This avoids a choice based sampling problem, because the eligibility criteria depended on the number of children attending school, which is a choice variable in our model.[11]

The PROGRESA data include information on the highest grade completed for all children ever born to the couple and on school attendance and work for all children in the household at the survey dates. The discrete choice decision-making model that we estimate assumes that parents allocate their children each school year (1997–1998 and 1998–1999) to one of three mutually exclusive activities: school, market work, or home. We used the following rules to allocate children to these activities based on the information in the data: (a) a child was considered as having attended school for the entire year if a child was enrolled in at least one of the two surveys during each school year and completed at least one grade level; (b) a child was considered as having not attended if the child was not enrolled in both surveys during each school year and did not complete a grade level; (c) any other cases were hand-edited to provide a consistent sequence of attendance and grade completion. A child who attended school, but did not complete a grade level, was deemed to have failed that school year.

A child was defined as working during the school year if the child did not attend school using the criteria above and had been working for salary (for 1997 in the October 1997 survey, and for 1998 in the October 1998 survey). Weekly wages are available in the data for those who reported working in the week prior to the survey. A child was designated as being home if

[10] A landless household is defined as a household that reported producing no agricultural goods for market sale. This restriction was adopted for three reasons. First, because PROGRESA is in part intended as an antipoverty program, the landless sample, being much poorer, is interesting in its own right. Second, the model is computationally burdensome to estimate, and restricting attention to landless households creates a smaller and more homogeneous sample. Third, to the extent that family child labor is not a perfect substitute for hired labor, the opportunity cost of family child labor is not the market wage, but rather their marginal product. We thus avoid having to estimate an agricultural production function.

[11] Another potential sample-selection problem concerns the migration of families into and out of the experimental villages. Families that migrated into the treatment villages between 1997 and 1998 were not included in the program, even if they satisfied the eligibility criteria. But, estimates of the impact of the program, based either on the experiment or on a model, could be biased if families that would have left the treatment villages in 1998 stayed because of the program.

TABLE 2—MEANS AND STANDARD DEVIATIONS OF SELECTED VARIABLES

	Mean	Standard deviations
Wife's age in 1997	30.5	8.1
Husband's age in 1997	34.4	9.5
Wife's age at marriage	18.1	3.4
Number of children ever born (1997)	3.01	1.92
Number of children ever born to women age 35–49 (1997)	4.05	2.14
Highest grade completed of children		
Age 7–11	2.39	1.41
Age 12–15	5.79	1.76
Age 16 or older		
All	6.60	2.81
Males	6.64	2.82
Females	6.56	2.81
Pct. with secondary school in village	26.7	—
Distance to secondary school if not in village (km)	2.82	1.60
Distance from city	136	73.6
Parent's income (pesos) in 1997	12,030	13,072
Percentage of children age 12–15 who worked for pay (1997)	9.1	—
Market income in 1997 of working children age 12–15	7,782	5,893

neither attending school nor working. We obtained parents' weekly income from the October surveys; it includes market earnings of both parents as well as self-employment income.[12] Both the children's weekly wage and the parents' weekly income were multiplied by 52 to obtain an annual equivalent.[13]

B. *Descriptive Statistics*

Table 2 presents basic sample statistics. The mean age of the wives in the sample as of 1997 is 30.5, and that of their husbands is 34.4. On average, the families had 3 children as of 1997. The average number of years of schooling completed was 2.4 years for children age 7 to 11, 5.8 years for children age 12 to 15, and 6.6 years for children age 16 and older. For the last group, there is almost no difference by sex in completed schooling. Parents' income over the two survey years was, on average, about 12,000 pesos (approximately 1,100 U.S. dollars). Approximately 9 percent of children between the ages of 12 and 15 were working for pay in 1997. Among those who worked, average income was about 8,000 pesos. Only about one-quarter of the villages have a junior secondary school located in the village, and among those villages that do not, the average distance to a secondary school is approximately three kilometers. The villages are also generally quite distant from major cities, with an average distance of 136 kilometers.[14]

Table 3 provides more detail concerning the time allocation of children in 1997. The first two columns compare the reported school attendance rates (percentage) of girls and boys age 6 through 15, in column 1 based on the raw data of whether the child was enrolled as of the October 1997 interview date, and in column 2 based on the revisions according to the rules described above. The third column shows the percentage of children working for pay and the last column the percentage at home (based on revised attendance rates).

A comparison of the first two columns shows that the revised attendance rates are slightly higher than the raw attendance rates, particularly for girls. Regardless of the measure, school attendance is almost universal from ages 7 to 11; attendance rates are somewhat lower at age 6, particularly for boys, reflecting a tendency for delayed entry into school. At age 12, an age by which, with normal progression, children would have completed primary school (grade 6), attendance rates fall to 89 percent for boys and to 90

[12] It is rare, as reported in the survey, for children of these landless households to have contributed to the self-employment income of the household.

[13] Weeks worked during the year were not reported in the data.

[14] We thank T. Paul Schultz for making the data on distance available to us.

TABLE 3—PERCENT OF CHILDREN ATTENDING SCHOOL, WORKING, AND HOME BY AGE AND SEX[a]

	Attends school (Oct. 1997)		Attends school (revised)		Works		At home	
Age	M	F	M	F	M	F	M	F
6	91.0	94.9	92.9	95.3	—	—	7.1	4.3
7	97.8	97.4	98.9	97.8	—	—	1.1	2.2
8	97.5	97.3	98.6	99.2	0	0	1.4	0.8
9	99.6	98.4	99.6	99.2	0	0	0.4	0.8
10	97.2	97.9	97.6	98.8	0	0	2.4	1.2
11	97.7	95.9	98.6	96.9	0	0	1.4	3.1
12	89.2	89.3	88.7	90.0	2.5	1.1	8.8	8.9
13	78.1	67.5	78.1	70.9	8.6	4.0	13.4	25.2
14	66.9	58.8	67.3	60.4	16.1	10.1	16.7	29.5
15	48.7	38.5	47.7	40.2	27.5	15.6	24.8	44.3

[a] Control and treatment groups in 1997. Only if data on school attendance and work are not missing.

percent for girls. After age 12, attendance rates continue to decline, more rapidly for girls. By age 15, only 48 percent of boys and 40 percent of girls attend school. At age 12, few children work for pay (2.5 percent of boys and 1.1 percent of girls); but by age 15, 28 percent of boys and 16 percent of girls are working.[15]

Girls progress through the early grades somewhat faster than boys, but ultimately complete about the same amount of schooling. As of October 1997, 12-year-old girls completed about 0.3 more years of schooling on average than boys of the same age. By age 16, this difference disappears, with both girls and boys completing, on average, 6.6 years of schooling. Girls are more likely to complete sixth grade, but are also more likely to drop out of school after completing it. As seen in Table 4, among children age 15 or 16 in 1997, 22 percent of boys and 17 percent of girls have fewer than 6 years of schooling, 32 percent of boys and 39 percent of girls have exactly 6 years, and 46 percent of boys and 44 percent of girls have more than 6 years. Failure rates for boys are higher than for girls in the primary grades, 15.7 percent versus 14.0 percent.

Table 5 shows fertility patterns, in particular, the Kaplan-Meier Survivor functions for the first ten years of marriage with respect to the birth of each of the first three children. Fertility occurs rapidly after marriage. About 50 percent of the women had their first birth within a year of marriage.[16] Two-thirds of the women had a first birth within the first two years of marriage and three-quarters within the first three years. About one-half of the women had a second birth within four years and one-half had a third birth within seven years. By the tenth year of marriage, all but a quarter of the women had three births.

TABLE 4—DISTRIBUTION OF HIGHEST GRADE COMPLETED AT AGES 15 AND 16[a]

Years of schooling	Boys	Girls
0	2.9	2.3
1	1.0	0.8
2	2.3	1.6
3	3.6	1.2
4	4.5	3.1
5	7.8	8.2
6	32.0	38.5
7	10.7	5.8
8	12.3	12.1
9+	23.0	26.5

[a] Control and treatment group in 1997.

Once children leave school, they rarely return. As seen in Table 6, only 12.5 percent of boys age 13 to 15 who worked one year attended school the next year. Similarly, only 14.8 percent of those who were at home returned to school. Comparable figures for girls are 40.0 percent (although the sample size is

[15] Child labor laws prohibit children under the age of 14 from working, but enforcement is lax in rural areas.

[16] When a woman's first birth occurred before or at her age at marriage (about 25 percent of the cases), the marriage was assumed to have occurred one year prior to the birth.

tapraid1/z3y-aer/z3y-aer/z3y00506/z3y2008d06a | yodert | S=42 | 10/12/06 | 13:56 | Art: 20031016 | Input-dld(dld)

TABLE 5—KAPLEIN-MEIER SURVIVOR FUNCTION FOR THE DURATIONS FROM MARRIAGE TO FIRST, SECOND, AND THIRD BIRTHS[a]

Duration from marriage (years)	To first birth	To second birth	To third birth
1	50.8	—	—
2	33.4	89.8	—
3	24.3	66.4	98.4
4	20.0	48.6	90.0
5	16.8	37.7	75.0
6	14.4	29.6	60.3
7	13.0	23.3	48.6
8	11.7	19.2	39.0
9	10.5	16.5	31.5
10	9.4	14.7	25.2

[a] Control and treatment groups in 1997.

TABLE 6—ONE-PERIOD TRANSITION RATES BY SEX: AGE (a) 13 TO 15

	Boys		
	Home (a)	Work (a)	School (a)
Home ($a-1$)	44.4	40.7	14.8
Work ($a-1$)	25.0	62.5	12.5
School ($a-1$)	8.3	5.5	86.2
	Girls		
	Home (a)	Work (a)	School (a)
Home ($a-1$)	92.5	7.5	0.0
Work ($a-1$)	40.0	20.0	40.0
School ($a-1$)	21.5	1.5	76.9

only 5) and 0 percent. The school-to-school transition also exhibits substantial permanence, with 86.2 percent of the boys and 76.9 percent of the girls who attended school in one year also attending the following year. Lastly, girls who are at home one year are likely to be at home the next year (92.5 percent), and boys who work tend to remain at work (62.5 percent).

III. The Model

A. *An Illustrative Model and Identification of Subsidy Effects*

Given that there is no direct cost of schooling through junior secondary school, and thus no variation from which to extrapolate the impact of a subsidy to attendance, it is useful to consider an illustrative model to demonstrate what information in the data would enable one to forecast the impact of the subsidy program. Consider a household with one child making a single period (myopic) decision about whether to send the child to school or to work, the only two alternatives. Let utility of the household be separable in consumption (C) and school attendance (s), namely $u = C + (\alpha + \varepsilon)s$, where $s = 1$ if the child attends school and 0 otherwise, with ε a preference shock that is normally distributed with mean zero and variance σ^2. The family's income is $y + w(1 - s)$, where y is the parent's income and w is the child's earnings if working. Under utility maximization, the family chooses to have the child attend school if and only if $\varepsilon \geq w - \alpha$. The unknown parameters of the model are thus α and σ. In this simple model, the probability that family i's child attends school is $1 - F_\varepsilon((w_i - \alpha)/\sigma)$. To obtain separate estimates of α and σ, it is both necessary and sufficient that child wages vary among families and that we observe those wages.[17] Fn17

Now, suppose the government is contemplating a program to increase school attendance of children through the introduction of a subsidy to parents of amount b if they send their child to school. Under such a program, the term b times s is added to the budget constraint and the probability that a child attends school will increase by $F_\varepsilon((w_i - \alpha - b)/\sigma) - F_\varepsilon((w_i - \alpha)/\sigma)$. As this expression indicates, knowledge of α and σ, estimated as above without the program, is necessary to forecast the impact of the program. Variation in the opportunity cost of attending school, the child market wage, thus serves as a substitute for direct variation in the tuition cost of schooling (the negative of b). The magnitude of the effect of the subsidy depends on the size of the subsidy, the child wage level, and the strength and variability in parental preferences about child schooling.

[17] More precisely, to use the probability statement above to estimate the parameters, we need to observe child wage offers. If we observe only accepted wages, that is, the wages of children who work, then we need also to be able to identify the parameters of the offered wage distribution together with α and σ. Given normality, standard selection arguments would hold for the identification of the wage offer parameters (James J. Heckman, 1979). Identification of α and σ requires an exclusion restriction, a variable that affects the offered wage but not the family's preference for child schooling.

B. *Model Description*

We next describe in detail the model that we estimate. In each discrete time period, a married couple makes a fertility decision and a time allocation decision for each child age 6 to 15. Specifically, the couple decides whether to have the woman become pregnant and have a child in the next period, and whether to send each child to school, to have the child work in the labor market (an option only after age 12), or to let the child remain at home. At age 16, children are assumed to become independent and make their own schooling and work decisions. A woman can become pregnant beginning at the age of marriage and ending at some exogenous age when she becomes infecund (assumed to be age 43). The contribution of the husband and wife to household income is exogenous (there are no parental labor supply decisions) and stochastic, and the household cannot save or borrow. The contributions to household income from working children under the age of 16 are pooled with parental income in determining household consumption.

To aid in the presentation of the structure of the model, Table 7 contains definitions of each of the model's variables. The utility function is given by T7

$$(1)\quad U(t) = U(C(t), p(t), \boldsymbol{n}(t), \boldsymbol{s}_b(t), \boldsymbol{s}_g(t), \boldsymbol{S}_b(t), \boldsymbol{S}_g(t), \boldsymbol{l}_b(t), \boldsymbol{l}_g(t), z_s; \boldsymbol{\varepsilon}(t), \boldsymbol{\mu}).$$

Parents receive a utility flow from household consumption (C), from a current pregnancy (p), from the history of births (n), from their boys' and girls' school attendance (s) and cumulative schooling (S), and from the set of children at home (l). The utility function also incorporates permanent household heterogeneity in the form of discrete types (μ) and time-varying shocks (ε) that affect the marginal utilities of some of the arguments in (1). The precise functional form of the utility function is shown in the Appendix. Here, we highlight a few of its characteristics. The utility function is constant relative risk aversion (CRRA) in consumption. With respect to schooling, parents derive utility in each period from the current average level of schooling completed by their children, from the current number of children who graduated from elementary school (grade 6), and from the current number who graduated from junior secondary school (grade 9).[18] The utility function also incorporates a psychic cost of attending a junior secondary school which varies with the distance to the nearest village with a secondary school (z_s); a utility loss from sending a child to school who is lagging behind in grades completed; and a utility loss if a child attends grade 10, which often requires living away from home. Another feature is that the value of keeping an older girl at home is allowed to depend on whether there are very young children in the household. The exact specification of the utility function was determined in part using model fit criteria. Fn18

Family consumption at t is equal to total family income. Family income is the sum of parental income (y_p) and the earnings of children (y_o) who work in the market.[19] Thus, the family's budget constraint is given by Fn19

$$(2)\quad C(t) = y_p(t) + \sum_n y_o(t, \tau_n)h(t, \tau_n).$$

The earnings (offer) of a child depends on the child's age and sex and the distance of the household's village from a city. As with parental income, there is a time-varying shock ($\varepsilon_{y_o}(t)$) and a permanent unobservable component (μ_{y_o}), both of which are family-specific. The distance from a city (z_c) is assumed to affect wage offers because of urban-rural differences in skill prices. Distance serves, in the child wage equation, as an identifying exclusion restriction, as previously discussed.[20] Parental income at t depends on the age of the husband ($a_p(t)$), the distance of the household's village from a city, a random shock ($\varepsilon_{y_p}(t)$), and a permanent parent-specific unobservable component (μ_{y_p}).[21] Specifically, log earnings of parents and children are Fn20 Fn21

[18] Note that parents continue to receive utility throughout their lifetime from the level of schooling a child has completed by age 16.

[19] Child rearing costs are essentially indistinguishable from the psychic value of children of different ages, which is included in parental utility rather than in the budget constraint.

[20] A second source of identification is that the child wage is modeled as a more flexible function of child age than in the utility function.

[21] Parental education enters the income function through its relationship to the unobservable parental type.

TABLE 7—MODEL VARIABLE DEFINITIONS

Variable	Definition
T	Duration of marriage ($t = 0$ at marriage)
t_m, τ_n	Age of the woman at marriage (t_m)
	Age of mother at birth of nth child (τ_n)
$t - \tau_n$	Age of nth child at t
$\boldsymbol{p(t)}$, $p(t)$	Vector of pregnancies up to t, $\boldsymbol{p(t)} = (p(0), p(1), \ldots, p(t))$
	$p(t) = 1$ if pregnancy occurs, at t and 0 otherwise
$\boldsymbol{n(t)}$, $n(t)$	Vector of births up to t, $\boldsymbol{n(t)} = (n(1), n(2), \ldots, n(t))$
	$n(t) = 1$ if a child is born at t and 0 otherwise ($n(t) = 1$ implies $p(t - 1) = 1$)
$\boldsymbol{b(t)}$, $b(t)$	Vector of births of boys up to t, $\boldsymbol{b(t)} = (b(1), b(2), \ldots, b(t))$
	$b(t) = 1$ if $n(t) = 1$ and the child is a boy, $= 0$ otherwise
$\boldsymbol{g(t)}$, $g(t)$	Vector of births of girls up to t, $\boldsymbol{g(t)} = (g(1), g(2), \ldots, g(t))$
	$g(t) = 1$ if $n(t) = 1$ and the child is a girl, $= 0$ otherwise
$N(t)$	Number of children born through t (= number of pregnancies through $t - 1$)
	$N(t) = N(t - 1) + n(t)$
$B(t)$	Number of boys born through t ($B(t) = B(t - 1) + b(t)$)
$G(t)$	Number of girls born through t ($G(t) = G(t - 1) + g(t)$)
$\boldsymbol{s(t)}$, $s(t, \tau_n)$	Vector of children's school attendance at t, $\boldsymbol{s(t)} = (s(t, \tau_1), \ldots, s(t, \tau_n))$,
	$s(t, \tau_n) = 1$ if the nth child, of age $t - \tau_n$, attends school at t and 0 otherwise
$\boldsymbol{s_j(t)}$	Vector of school attendance at t of children of sex $j = b, g$
	$\boldsymbol{s_j(t)} = (s(t, \tau_1)j(\tau_1 - t_m), \ldots, s(t, \tau_n)j(\tau_n - t_m))$
$S(t, \tau_n)$	Cumulative schooling at t of the nth child
	$c(t - 1, \tau_n) = 1$ if the year of schooling is successfully completed, $= 0$ if not
	$S(t, \tau_n) = S(t - 1, \tau_n) + s(t - 1, \tau_n) \times c(t - 1, \tau_n)$
$\boldsymbol{S(t)}$	Vector of children's cumulative schooling at t, $\boldsymbol{S(t)} = (S(t, \tau_1), \ldots, S(t, \tau_n))$
$\boldsymbol{S_j(t)}$	Vector of cumulative schooling at t of children of sex $j = b, g$,
	$\boldsymbol{S_j(t)} = (S(t, \tau_1)j(\tau_1 - t_m), \ldots, S(t, \tau_n)j(\tau_n - t_m))$
$\pi_c(t)$	Probability of completing a grade conditional on attendance
$\boldsymbol{h(t)}$, $h(t, \tau_1)$	Vector of children at work at t, $\boldsymbol{h(t)} = (h(t, \tau_1), \ldots, h(t, \tau_n))$
	$h(t, \tau_n) = 1$ if the nth child works at t, $= 0$ otherwise
$\boldsymbol{h_j(t)}$	Vector of children at work of sex $j = b, g$, $\boldsymbol{h_j(t)} = (h(t, \tau_1)j(\tau_1 - t_m), \ldots, h(t, \tau_n)j(\tau_n - t_m))$
$\boldsymbol{l(t)}$, $l(t, \tau_n)$	Vector of children at home at t, $\boldsymbol{l(t)} = (l(t, \tau_1), \ldots, l(t, \tau_n))$
	$l(t, \tau_n) = 1$ if the nth child is at home, i.e., neither attends school nor works, at t, $= 0$ otherwise
$\boldsymbol{l_j(t)}$	Vector of children at home of sex $j = b, g$, $\boldsymbol{l_j(t)} = (l(t, \tau_1)j(\tau_1 - t_m), \ldots, l(t, \tau_n)j(\tau_n - t_m))$
z_s, z_c	Distance to a secondary school (z_s), distance to a city (z_c)
$y_p(t)$	Parents' income at t
$y_o(t, \tau_n)$	nth child's income at t
$a_p(t)$	Age of male parent at t
$C(t)$	Household consumption at t
$U(t)$	Parents' utility at t
$\boldsymbol{\varepsilon(t)}$, $\varepsilon_p(t)$, $\varepsilon_b(t)$, $\varepsilon_g(t)$, $\varepsilon_{yp}(t)$, $\varepsilon_{yo}(t)$	Vector of time-varying preference shocks, $\boldsymbol{\varepsilon(t)} = (\varepsilon_p(t), \varepsilon_b(t), \varepsilon_g(t), \varepsilon_{yp}(t), \varepsilon_{yo}(t))$
	$\varepsilon_p(t)$: parents' i.i.d. shock to the marginal utility of a pregnancy
	$\varepsilon_b(t)$, $\varepsilon_g(t)$: parents' i.i.d. shock to the marginal utility of having an additional (age-weighted) boy/girl at home
	$\varepsilon_{yp}(t)$: i.i.d. shock to parents' income
	$\varepsilon_{yo}(t)$: i.i.d. shock to children's income
$\boldsymbol{\mu}$, μ_{Nk}, μ_{Sk}, $\mu_{lb,k}$, $\mu_{lg,k}$, $\mu_{yp,k}$, $\mu_{yo,k}$	Vector of permanent components of utility, $\boldsymbol{\mu} = (\mu_{Nk}, \mu_{Sk}, \mu_{lb,k}, \mu_{lg,k}, \mu_{yp,k}, \mu_{yo,k})$
	μ_{Nk}: permanent components of the marginal utility of the number of children for the kth type of household
	μ_{Sk}: permanent components of the marginal utility of children's schooling for the kth type of household
	$\mu_{lb,k}$: permanent components of the marginal utility of having boys at home for the kth type of household
	$\mu_{lg,k}$: permanent components of the marginal utility of having girls at home for the kth type of household
	$\mu_{yp,k}$: permanent component of parents' income for the kth type of household
	$\mu_{yo,k}$: permanent component of children's income for the kth type of household
δ	Discount factor

$$(3)\quad \log y_p(t) = y_p(a_p(t), z_c, \varepsilon_{y_p}(t); \mu_{y_p}),$$

$$\log y_o(t, \tau_n) = y_o(t - \tau_n, \mathrm{I}(b(\tau_n) = 1), z_c, \varepsilon_{y_o}(t); \mu_{y_o}).$$

Finally, the probabilistic grade completion function depends on the child's grade level, age, and a permanent unobservable family component, namely

$$(4)\quad \pi_c(t, \tau_n) = \pi(t - \tau_n, S(t, \tau_n)|s(t, \tau_n) = 1, \mu_c).$$

The five time-varying ε-shocks are assumed to be jointly normally distributed and serially uncorrelated, with a density $t(\varepsilon(t))$.[22] The permanent components of parental preferences and income and of child earnings and grade completion are also assumed to be jointly distributed according to $g(\mu)$, where g is discrete with k support points or household types. These permanent components are assumed to be known to parents from the beginning of the marriage.

At any t, the couple maximizes the present discounted value of remaining lifetime utility with respect to $s(t)$, $l(t)$, and $p(t)$. Thus, in any period, the family faces $K(t)$ mutually exclusive alternatives, where K varies over time with the number of children eligible to attend school and work and with the woman's age (whether she is fecund). Define, $d_k(t) = 1$ if the kth alternative is chosen at t, and 0 otherwise. Further, define $\Omega(t)$ to be the state space at t, consisting of all the relevant factors affecting current or future utility or the distributions of future shocks, that is, $b(t)$, $g(t)$, $S_b(t)$, $S_g(t)$, $a_p(t)$, $\varepsilon(t)$, μ, t_m, z_s, z_c.

The maximized present discounted value of lifetime utility at t, the value function, is given by

$$(5)\quad \mathrm{V}(\Omega(t), t) = \max_{d_k(t)} \mathrm{E}\left\{\sum_{\tau = t_m}^{\bar{T}} \delta^{\tau - t_m}\mathrm{U}(t)|\Omega(t)\right\},$$

where $\bar{T}$ is the terminal decision period (woman's age 59) and the expectation is taken over the distribution of parental preference and income shocks, the children's earnings shock, and the implicit shocks to grade completion for choices that involve school attendance.[23] The solution to the optimization problem is a set of decision rules that relates the optimal choice at any t, from among the feasible set of alternatives, to the elements of the state space at t. Recasting the problem in a dynamic programming framework, the value function can be written as the maximum over alternative-specific value functions, $V^k(\Omega(t), t)$, i.e., the expected discounted value of alternative $k \in K(t)$, which satisfies the Bellman equation, namely

$$(6)\quad \mathrm{V}(\Omega(t), t) = \max_{k \in K(t)} [\mathrm{V}^k(\Omega(t), t)],$$

$$\mathrm{V}^k(\Omega(t), t) = \mathrm{U}^k(t, \Omega(t)) + \delta\,\mathrm{E}(\mathrm{V}(\Omega(t + 1), t + 1|d_k(t) = 1, \Omega(t))) \quad \text{for } t < \bar{T},$$

$$= \mathrm{U}^k(\bar{T}, \Omega(\bar{T})) \quad \text{for } t = \bar{T}.$$

C. *Model Solution*

Given that the solution of the optimization problem is in general not analytic, we solve the model numerically. Its solution consists of the values of $E(V(\Omega(t + 1), t + 1|d_k(t) = 1, \Omega(t)))$ for all k and elements of $\Omega(t)$. For convenience, we call this function *Emax*. The solution method proceeds by backward recursion beginning with the last decision period.

There are two complications in solving the model numerically. First, at any fecund period in which all children are of school and work age, the choice set is of order $2 \cdot 3^{N_1(t)}$, where the first term represents the choice of whether to have a pregnancy and the second reflects the number of joint school attendance–work choices (of which there are 3) and $N_1(t)$, the number of children age 12 to 15. For example, if there are three children between the ages of 12 and 15, there are 54 possible choices. One way to reduce the size of the choice set in a way that is for the most part consistent with the data is to assume that for each sex, a child may attend school only if all younger children attend

[22] The implicit time-varying shock to grade completion is assumed to be independent of all other shocks in the model.

[23] The integration is also performed over whether a birth is a boy or a girl. We assume the probability of each sex outcome to be 0.5.

school and, independent of sex, a child may work for pay only if all older children work for
Fn24 pay.[24] In the case of three children within the 12- to 15-year age range of the same sex, the number of alternatives, for example, is reduced to 20.

Second, the size of the state space makes a full solution of the problem computationally intractable, because the *Emax* functions must be calculated for all state values at each t. As long as the ages of children affect lifetime utility, as they must because of the age restrictions on children's eligibility for schooling and work, the state space includes the entire sequence of births by sex, and not simply the stock of children. In addition, at any t, the schooling level of each child affects expected lifetime utility at t. To solve the dimensionality problem, we adopt an approximation method developed in Michael P. Keane and Wolpin (1994, 1997, 2001) in which the *Emax* functions are evaluated at a random subset of the state points, and the values are used to fit a global polynomial approximation in the state variables. To further limit the size of the state space, we also assume that women can have no more than eight chil-
Fn25 dren.[25] As in Keane and Wolpin, the multivariate integrations necessary to calculate the expected value of the maximum of the alternative-specific value functions at those state points are performed
Fn26 by Monte Carlo integration over the ε-shocks.[26]

D. *Model Estimation*

The solution to the couple's maximization problem serves as input into estimating the parameters of the model. The numerical solution method described above provides the *Emax* functions that appear on the right-hand side of (6). The alternative-specific value functions, $V^k(t)$ for $k = 1, \ldots, K(t)$, are known, except for the parental random preference and income shocks and the earnings shock of the children. Thus, conditional on the deterministic part of the state space, the probability that a couple chooses option k takes the form of an integral over the region of a subset of the random shocks, such that k is the preferred option.

Specifically, in the decision model presented above, the observed outcomes at each period include: (a) the choice (from the feasible set) made by the couple of whether or not to have a pregnancy, which children to send to school, which to have work in the market, and which to remain at home; (b) the wages received by the children who work in the market; (c) the success or failure of those children who attend school to complete a grade level; and (d) parental income. Let the outcome vector at t be denoted by $O(t) = \{d^k(t), y_o(t), c(t), y_p(t)\}$. Suppose we observe these outcomes for a sample of N households beginning at marriage, $t = \underline{t}_{mn}$ for household n, and ending at some $t = t_n$. Then, the likelihood for this sample is

$$(7) \qquad \prod_{n=1}^{N} \Pr(\mathrm{O}(\bar{t}_n), \ldots, \mathrm{O}(t_{m+1,n}), \mathrm{O}(t_{mn}))|\bar{\Omega}(t_{mn}), \mu),$$

where $\bar{\Omega}(t_{mn})$ is the observable components of the initial state space at the time of marriage, that is, the state space net of the family's type and stochastic shocks at $t = t_{mn}$. The observable part of the state space at marriage consists only of the age of the woman and of the man at t_{mn} and distances from a secondary school and from a city. Because type is unobserved, it must be integrated out. Thus, the sample likelihood is

$$(8) \qquad \prod_{n=1}^{N} \sum_{j=1}^{J} \Pr((\mathrm{O}(\bar{t}_n), \ldots, \mathrm{O}(t_{m+1,n}), \mathrm{O}(t_{mn}))|\bar{\Omega}(t_{mn}), \text{type} = j)\Pr(\text{type} = j|\bar{\Omega}(t_{mn})).$$

[24] We do not impose these restrictions on six- and seven-year-old children to accommodate the fact that school entry is sometimes delayed. Violations of the assumption in the 1997 survey occur in about 5 percent of the households in the case of schooling, and in about 1 percent of the households in the case of working. Such households were dropped.

[25] Only about 3 percent of women in our sample report having more than eight children. In the empirical implementation, we assume that children of birth order greater than eight were not born.

[26] We used 2,500 state points for the estimation of the *Emax* approximations and 50 draws for the numerical integrations. The *Emax* approximations did not appear to be sensitive to increases in these parameters, up to 10,000 state points and 300 draws. There were approximately 150 variables used in the *Emax* approximation, which includes interactions among the state variables. The R-squares were above 0.99 in all model periods.

tapraid1/z3y-aer/z3y-aer/z3y00506/z3y2008d06a | yodert | S=42 | 10/12/06 | 13:56 | Art: 20031016 | Input-dld(dld)

We assume that the initial conditions—the ages of marriage of both parents and the distances—are exogenous conditional on type.

There are two additional considerations in computing the likelihood. Because we assume that the child wage shock is family-specific, having an observation on the wage for two children in the same family working in the same period who have different wages (conditional on the relevant observable determinants of child earnings, child age, and sex as in (3)) will lead to a degenerate likelihood. We therefore assume that the children's wages are measured with error, which seems like a reasonable assumption.[27] Assuming a multiplicative normal measurement error, observed child earnings is given by $y_o^{\text{obs}}(t) = y_o(t)\exp(\eta(t))$.

Another difficulty arises because, for most of the families, we do not observe the decisions from the start of marriage. In particular, we have a complete fertility history but do not have a complete school attendance and work history for children. For example, consider a family with 3 children whose ages are 6, 12, and 16 as of the October 1997 survey date and whose marriage occurred in 1980, when the woman was age 19. For this family, we observe fertility outcomes at every t between 1980 and 1997, when the woman was age 19 through 36. However, we are missing the history of school attendance for the 12- and 16-year-olds, and the work history for the 16-year-olds. Although it is conceptually straightforward to accommodate the missing information into the likelihood function (8), it is computationally infeasible to perform the necessary integrations over all of the feasible unobserved choice paths.

To avoid having to deal with missing data on the schooling and work histories of children, one could restrict the sample to marriages that occurred between 1989 and 1997 for which there are complete data. But for the earliest marriages in this range, the oldest age a child could be at the time of the survey is 6. It is not possible to identify all of the model parameters solely from those observations, because children do not start work for pay until age 12.

For children age 7 to 15 in 1997, we have data on outcomes at the survey years (1997 and 1998), but we are missing their outcome history. We can incorporate the available information into the likelihood, but need to allow for the fact that the state variables for the decision problem in 1997 and 1998 are not strictly exogenous; they include the histories of past decisions about pregnancies and schooling.[28] Our assumption of serial independence in the shocks implies that the state variables at any time t are exogenous with respect to decisions at t conditional on type. Thus, the likelihood for the observations in 1997 and 1998, conditioning on the 1997 state space, can be written, analogous to (8), as

$$(9) \qquad \prod_{n=1}^{N} \sum_{j=1}^{J} \Pr(\mathrm{O}(t_n^{98}), \mathrm{O}(t_n^{97}) | \bar{\Omega}(t_n^{97}), \text{type} = j) \Pr(\text{type} = j | \bar{\Omega}(t_n^{97})),$$

where t_n^{97} and t_n^{98} are the marriage durations as of 1997 and 1998 and where $\bar{\Omega}(t_n^{97})$ is the state space as of 1997 (inclusive of the initial conditions at the time of marriage). Equation (9) requires that we specify how the type distribution is related to the state variables. The form of the conditional probability is, in principle, derivable from the structure of the behavioral model, together with the relationship between type and the initial state variables, i.e., the second term in (8). The form is not analytical, however, and is not numerically tractable. The alternative we adopt is to approximate the conditional type probability using a multinomial logit specification.[29]

[27] We follow this strategy as opposed to allowing for child-specific wage shocks, to avoid having to integrate over all of the child shocks in calculating the *Emax* functions. The problem of degeneracy exists more generally, namely that with family-level shocks, some choices may not be generated by the model. Restricting the choice set as we have reduces the likelihood of this event, but does not eliminate it necessarily. Estimation is feasible when such events occur because our procedure smooths over zero likelihood events (see below). After estimating the model, we verified that simulations of the model could generate all of the outcomes that were observed in the data, so none of these outcomes has zero probability of occurrence.

[28] This is the initial conditions problem in discrete choice models as discussed in Heckman (1981).

[29] Furthermore, the parameters of the approximation are not really free, being themselves functions of the structural parameters. The estimation method is thus not fully efficient.

To summarize, in estimating the model, we use (8) for the families with complete decision histories, that is, couples who have been married 7 years or less as of October 1997. We use (9) for the families with incomplete decision histories (couples married 8 or more years as of October 1997), which means that we ignore the information about pregnancy decisions made prior to 1997.[30] Given the assumption of joint serial independence of the vector of shocks (conditional on type), both (8) and (9) can be written as the product of within-period outcome probabilities conditional on the corresponding state space and type. Each of these conditional probabilities is of dimension equal to the number of contemporaneous shocks in $\varepsilon(t)$.

To illustrate the calculation of the likelihood, it is sufficient to consider a specific outcome at some period. Suppose that the kth alternative that is chosen at period t is to send at least some children to work. The children who work are observed to have wages given by $y_{oj}(t)^{\text{obs}}$, where j signifies the jth working child and the superscript "obs" distinguishes the observed wage from the true wage, $y_{oj}(t)$. The likelihood contribution for such an observation is (for a given type)

$$
\begin{aligned}
(10)\quad & \Pr(d^k(t) = 1, \tilde{y}_o(t)^{\text{obs}} \,|\, \Omega(t), \text{type}) \\
& = \int_{\tilde{y}_o(t)} \Pr(d_t^k = 1 \,|\, \tilde{y}_o(t), \Omega(t), \text{type}) \\
& \quad \cdot \Pr(\tilde{y}_o^{\text{obs}}(t), \tilde{y}_o(t) \,|\, \Omega(t), \text{type}),
\end{aligned}
$$

where "$\sim$" signifies the vector of child wages over j and the integration is of the same order as the number of children who work.[31] Notice that it is necessary to integrate over the vector of true wages in (10) because the choice probability depends on true wages, which we observe only with error. Probability statements for other alternative choices are obtained similarly. We calculate the right-hand side of (10) by a smoothed frequency simulator.[32]

The entire set of model parameters enters the likelihood through the choice probabilities that are computed from the solution of the dynamic programming problem. Subsets of parameters enter through other structural relationships as well, e.g., child wage offer functions, the parents' income function, and the school failure probability function. The estimation procedure, i.e., the maximization of the likelihood function, iterates between the solution of the dynamic program and the calculation of the likelihood.

IV. Results

A. *Parameter Estimates*

The precise functional forms of the model's structure are provided in Appendix A. Parameter estimates, and their standard errors, are given in Appendix Table B1. The model was fit with three household types.[33] Recall that types differ with respect to their underlying preferences (for fertility, child schooling, and child leisure), school failure rates, parental income

[30] Some households are missing a subset of choices. For example, we may know that a child did not attend school but do not know whether the child was at home or at work. If there are missing data in 1997, then the state space for 1998 choices is also missing some elements. We account for households with missing data in the likelihood function by integrating over all possible choices the household could have made and appropriately updating the state space. Some households are also missing parental income or child wages for working children. The likelihood accounts for these missing elements by numerically integrating over the appropriate densities.

[31] For ease of exposition, we have ignored parents' income in the formulation of the likelihood function as well as whether the children who were sent to school failed to progress to the next grade level. The modifications of (9) to account for these additional observable variables are straightforward and we take them into account in evaluating the likelihood.

[32] The kernel smoothed frequency simulator we adopt was proposed in McFadden (1989). For each of K draws of the error vector, $\varepsilon_p(t)$, $\varepsilon_{1b}(t)$, $\varepsilon_{1g}(t)$, $\varepsilon_{y_p}(t)$, $\eta(t)$, noting that $\varepsilon_{y_o}(t)$ is chosen to satisfy the observed wage for each child, that is, inclusive of the measurement error, the kernel of the integral is $\exp[\mathrm{V}^k(t) - \max(\mathrm{V}^j(t))/\tau]/\Sigma_i \exp[\mathrm{V}^i(t) - \max(\mathrm{V}^j(t))/\tau]$ times the joint density of the observed and true wage, where the j superscript denotes the vector of value functions over all alternatives. The first term in the kernel is the smoothed simulator of the probability that $d_k(t) = 1$, with τ, the smoothing parameter, set equal to 10, which provided sufficient smoothing given the magnitudes of the value functions. See Keane and Wolpin (1997) and Zvi Eckstein and Wolpin (1999) for further applications.

[33] We settled on three types because there were significant improvements in model fit beyond two types. Given the computational burden, we did not attempt to fit the model with four types.

TABLE 8—PREDICTED SELECTED CHARACTERISTICS BY UNOBSERVED TYPE

	Type 1		Type 2		Type 3	
	Girls	Boys	Girls	Boys	Girls	Boys
Percent of children age 6–11 in school	98.5	99.4	97.6	99.9	78.7	64.2
Percent of children age 12–15 in school	37.3	50.2	84.6	86.9	44.5	36.8
Percent of children age 12–15 at home	55.9	31.0	11.3	7.0	33.5	30.9
Percent of children age 12–15 at work	6.8	18.8	4.1	6.1	21.9	32.3
Mean wage of children 12–15	2,675	3,599	2,600	3,499	2,739	3,666
Mean parental income	9,953		11,944		10,107	
Percent becoming pregnant	15.0		5.6		14.8	
Percent of sample	38.8		52.0		9.2	

potential, and child earnings potential. The three types have distinctly different behaviors. T8 As seen in Table 8, type 1 households, comprising 39 percent of the sample, and type 3s, comprising 9 percent of the sample, value schooling less than type 2s. However, type 1s and type 3s also differ; the percentage of the youngest children, age 6 to 11, from type 3 households who attend school is considerably lower than those from type 1 households. Moreover, in terms of schooling overall, type 1 households seem to favor boys and type 3 households girls, with type 2 households exhibiting little sex bias. Children age 12 to 15 from type 2 households are least likely to work. Although school attendance rates of children age 12 to 15 are similar for type 1 and type 3 households, those from type 3 households are considerably more likely to work and, concomitantly, less likely to be at home. Child offered wages, on the other hand, differ very little among the types and are only about one-third as large as mean accepted wages, while parental income is 20 percent higher for type 2s than for type 1s or 3s. Type 2 households, in addition to sending their children to school at a higher rate, are about two-thirds less likely to have an additional pregnancy during the year than either of the other types.

B. *Internal Validation: Within-Sample Fit*

We next present evidence on the within-sample fit of our model along various dimensions of the T9 data. Table 9 compares the model's prediction of the distribution of child activity allocations (school, work, or home) at individual ages by sex to the actual distribution. The table also reports the chi-square statistic associated with a test of the null that the predicted and actual distributions are the same.[34] Fn34 At younger ages, when school attendance is nearly universal, the model predicts an attendance rate nearly identical to the actual rate. Between ages 11 and 12, when attendance drops as children finish primary school, the model captures and, in fact, overstates the drop. It predicts an 11.8-percentage-point drop for boys compared to an actual drop of 9.2 and a 9.4-percentage-point drop for girls compared to an actual drop of 7.3. The model also fits the choices between working for pay and staying at home. For example, it captures the pattern in the data that teenage girls are twice as likely as teenage boys to be at home at age 15, while teenage boys are more likely to work for pay. As seen in the table, the null that predicted and actual rates are the same is never rejected at the 5-percent level.

Table 10 compares the actual and predicted school attendance rates for children whose T10 schooling attainment differs from their maximum potential, defined as the level they could have achieved had they enrolled at age 6 and attended school continuously without repeating grades. Later, we use these subgroups for out-of-sample validation. The predicted rates for the subgroups that are not behind in school are about 5 percentage points too low (the null is rejected at the 5-percent level), but the attendance rates for the other subgroups are within 1 to 2 percentage points of the actual rates. Table 11 compares the observed wages of children T11 who are working to the wages for working children predicted under the model. The model's predicted (accepted) wages tend to be high relative to the observed (accepted) wages. Av-

[34] These tests do not correct for the fact that the predicted distributions are based on estimated parameters.

TABLE 9—ACTUAL AND PREDICTED CHOICE DISTRIBUTION BY CHILD AGE AND SEX
(*Pooled* 1997 *and* 1998)

Boys							
	Actual			Predicted			
Age	School	Work	Home	School	Work	Home	χ^2
6	0.933	—	0.066	0.923	—	0.077	0.58
7	0.981	—	0.019	0.980	—	0.020	0.02
8	0.987	—	0.013	0.980	—	0.020	0.99
9	0.994	—	0.006	0.979	—	0.021	3.49
10	0.982	—	0.018	0.974	—	0.026	0.86
11	0.977	—	0.023	0.964	—	0.036	1.45
12	0.885	0.021	0.094	0.846	0.039	0.115	3.99
13	0.780	0.084	0.136	0.736	0.078	0.186	4.51
14	0.677	0.157	0.166	0.619	0.191	0.190	3.41
15	0.490	0.276	0.235	0.520	0.251	0.229	0.88
Girls							
6	0.965	—	0.035	0.942	—	0.058	3.84
7	0.976	—	0.024	0.968	—	0.032	0.77
8	0.989	—	0.011	0.976	—	0.024	1.96
9	0.991	—	0.009	0.975	—	0.025	3.26
10	0.979	—	0.021	0.970	—	0.030	0.93
11	0.969	—	0.031	0.948	—	0.052	2.97
12	0.896	0.007	0.097	0.854	0.020	0.126	4.61
13	0.726	0.028	0.245	0.676	0.025	0.299	2.85
14	0.582	0.089	0.329	0.566	0.092	0.342	0.22
15	0.419	0.123	0.458	0.402	0.157	0.442	1.68

Note: χ^2 (0.05, 1) = 3.84, χ^2 (0.05, 2) = 5.99.

TABLE 10—ACTUAL AND PREDICTED SCHOOL ATTENDANCE RATES BY NUMBER OF YEARS LAGGING BEHIND IN SCHOOL: AGE 12–15

	Boys			Girls		
Age	Actual	Predicted	χ^2	Actual	Predicted	χ^2
Not behind	88.3	82.1	8.50	83.8	78.2	6.02
Behind one year	79.8	76.4	1.56	75.4	74.5	0.09
Behind two years	65.8	62.5	0.91	52.9	51.0	0.20
Behind three years or more	49.1	51.7	0.62	44.7	42.7	0.39

Note. χ^2 (0.05, 1) = 3.84.

eraged over the ages of 12 through 15, the mean accepted wage is approximately 10 percent higher for boys and 28 percent higher for girls.

We also looked at the fit of the model with respect to fertility. Given space limitations, we summarize the results without presenting additional tables. The model predicts that the mean number of children born to couples married seven years or less would be 1.73, when the actual mean number is 1.79. The model overpredicts the number of women in this group having zero children by about 7 percentage points, and underpredicts the number with two

TABLE 11—ACTUAL AND PREDICTED ANNUAL WAGE IF WORKING, BY CHILD AGE AND SEX[a]
(*Number of observations in parentheses*)

	Boys		Girls	
Age	Actual	Predicted	Actual	Predicted
12	6,234 (6)	9,298	3,720 (2)	7,301
13	7,064 (21)	7,618	5,460 (6)	6,908
14	7,644 (34)	10,218	8,726 (19)	9,306
15	10,188 (53)	10,313	6,386 (22)	9,848

[a] 1997 pesos.

TABLE 12—ACTUAL AND PREDICTED SCHOOL ATTENDANCE RATES BY CHILD AGE, SEX, AND SCHOOL ATTAINMENT: CONTROL AND TREATMENT GROUPS BY YEAR[a]

	Girls				Boys			
	Control group		Treatment group		Control group		Treatment group	
	1997	1998	1997	1998	1997	1998	1997	1998
Age 6–11								
Actual	96.9	96.5	97.6	98.5	96.6	96.7	97.6	98.7
Predicted	96.1	96.2	96.4	97.1	96.4	96.4	96.3	97.1
No. obs.	449	431	632	600	471	460	671	678
Age 12–15								
Actual	65.3	66.5	62.9	74.4	68.8	72.5	69.5	76.3
Predicted	61.6	61.8	61.8	74.9	68.8	68.8	68.0	77.1
No. obs.	190	176	205	223	189	182	279	262
Age 12–15 behind in school								
Actual	58.3	58.7	56.9	71.4	64.0	67.4	64.2	71.6
Predicted	54.2	55.5	55.6	72.3	63.9	65.3	62.7	72.9
No. obs.	127	121	144	161	139	135	204	190
Age 13–15 HGC $\geq$ 6 behind in school								
Actual	40.9	44.4	30.3	51.5	59.0	57.1	52.6	58.3
Predicted	40.2	45.3	37.3	58.7	55.0	53.0	51.7	66.7
No. obs.	66	72	66	66	61	56	95	96

[a] Based on 200 simulation draws per family.

or more by about the same amount. The model predicts well the fall in the pregnancy rate with the number of prior pregnancies. In the data, 57 percent of these women without children become pregnant. The model predicts 53 percent. Both in the data and in the model, 18 percent with 2 children and 6 percent with 4 children become pregnant. For couples married 8 years or more, the model overpredicts the proportion having a pregnancy at most ages by less than 2 percentage points; for example at ages 25 to 29, the model predicts a pregnancy rate of 0.174 and the actual pregnancy rate is 0.151.

V. The Test of Model Validity: Comparison of Impacts Predicted under the Model to Experimental Impacts

Given the parameter estimates, it is straightforward to predict the impact of the school subsidy program on school attendance. A subsidy paid to the family for each child who attends school changes the family budget constraint (2). Resolving the optimization problem for each family taking into account the subsidy will lead to a different pattern of school attendance and fertility decisions. Comparing the decisions of the treatment group predicted under the model to their actual decisions (at the same stage in the life cycle and for the same state variables) provides a direct out-of-sample test of the model's validity.

We predict the subsidy effects in two different ways. The one-step-ahead prediction uses information on the state variables in 1997 or 1998 to forecast the effects of the program during the subsequent year. The *N*-step-ahead prediction makes use only of information on initial conditions, i.e., the age and education levels of the wife and husband at marriage, and the distances to schools and to the nearest city, to forecast choices at any point in the couple's lifetime. Comparing the *N*-step-ahead prediction to the actual data is a more severe test of the model and is useful in assessing the model's validity in making longer-term forecasts.

A. *One-Step-Ahead Predictions from Current State Variables*

Table 12 compares the actual and predicted school attendance rates for different categories of children in the control and treatment groups, defined by age, sex, and completed T12

tapraid1/z3y-aer/z3y-aer/z3y00506/z3y2008d06a | yodert | S=42 | 10/12/06 | 13:56 | Art: 20031016 | Input-dld(dld)

schooling. The only group to receive the subsidy, the treatment group in 1998, was, as noted, not used in fitting the model. Therefore, a comparison of the predictions shown in the last two columns of the table with the actual attendance rates represents an out-of-sample test of the model's validity.[35] The table also presents within-sample comparisons for the three nonsubsidy groups, for the control group in 1997 and 1998, and for the treatment group in 1997. As seen in the first two rows of the table, predicted attendance rates usually come within 1 to 2 percentage points of the actual attendance rate for all the groups in the 6–11 age category. In 1998, the model predicts an attendance rate for the treatment group equal to 97.1 percent for both boys and girls, compared to actual attendance rates of 98.5 percent and 98.7 percent. The fact that there is close to universal attendance for those ages clearly limits the power of the out-of-sample validation test.

In contrast to children in the 6–11 age category, there is a substantial difference in attendance rates for older children between the three nonsubsidy groups and the subsidy group. Specifically, the actual attendance rate for 12- to 15-year-olds in the 1998 treatment group exceeds that of the nonsubsidy groups by 8 to 12 percentage points for girls and 4 to 8 percentage points for boys. Nevertheless, the accuracy of the out-of-sample prediction of the subsidy group differs very little from that of the within-sample prediction for the nonsubsidy groups. For children age 12 to 15, the predicted attendance rates tend to be a few percentage points lower than the actual rates for the nonsubsidy groups and are within 1 percentage point of the actual rate for the 1998 treatment group, 74.9 versus 74.4 percent for girls and 77.1 versus 76.3 percent for boys.[36, 37]

To further assess the validity of the model, it is useful to consider subsamples for which there is even more room for prediction error. For this purpose, we consider the sample of children age 12 to 15 who are behind in school, having either failed at previous grade levels or not attended some time in the past. Attendance rates are lower for those who are behind in school than for those who are not behind, about 6 to 8 percentage points for girls and about 4 to 5 percentage points for boys. In that case as well, the predicted attendance rate for the 1998 treatment group is quite close to the actual, 72.3 versus 71.4 percent for girls and 72.9 versus 71.6 percent for boys. Further restricting this last sample to those who have completed the sixth grade (HGC $\geq$ 6), for whom there is an even lower attendance rate, does lead to a poorer prediction both for some of the nonsubsidy groups and for the 1998 treatment group. In the latter case, the model overpredicts attendance rates by 7.2 percentage points for girls and by 8.4 percentage points for boys. Although larger than previous prediction errors, attendance rates for the nonsubsidy groups are a lower bound for the model's prediction of the attendance rate for the subsidy group, and that 100 percent is the upper bound.

Without the experiment, the estimate of the subsidy effect would have to make use of the nonsubsidy groups alone. Table 13 compares the model's predicted impacts of the subsidy on attendance, obtained using those groups, to the experimental impact estimates. Two different ways of computing the experimental impacts are shown in the row labeled "experimental treatment effect." The cross-section effect is the average attendance rate for the treatment group minus the average rate for the control group in

[35] We did not look at any statistics for the treatment group in 1998 before we had completed the estimation of the model.

[36] For 12- to 13-year-old girls (145 children), the actual and predicted attendance rates were 86.9 versus 85.7 percent, and for boys (141 children), 89.4 versus 87.6 percent. For 14- to 15-year-old girls (78 children), the actual and predicted rates were 57.3 versus 51.3, and for boys (121 children) 61.2 versus 65.6 percent. Although the differences are somewhat larger than in the combined 12- to 15-year-old group, they are still quite close, especially considering that the predictions could range up to 100 percent attendance.

[37] Our strategy of using the treatment group to validate the model assumed that the control group did not expect to be brought into the program. PROGRESA administrative personnel reportedly were especially careful not to inform the control group families about the existence of the program or about future plans to incorporate them. If there were anticipatory effects, as suggested by Attanasio et al. (2001), we would expect them to be present in 1998 but not in 1997, because the baseline data were gathered before the initiation of the program. Thus, we would expect our model, estimated under the assumption of no anticipation, to fit the 1997 schooling patterns better than the 1998 patterns. The results above lead us to conclude that there is no strong evidence of anticipatory effects.

TABLE 13—ACTUAL VERSES PREDICTED SUBSIDY EFFECTS ON PERCENT ATTENDING SCHOOL

	Girls age 12–15			Girls age 12–15, behind in school			Girls age 13–15, HGC ≥ 6, behind in school		
	(1) Actual	(2) Pred. with Subsidy	(2)–(1)	(1) Actual	(2) Pred. with Subsidy	(2)–(1)	(1) Actual	(2) Pred. with Subsidy	(2)–(1)
97 Control	65.3	72.7	7.4	58.3	67.0	8.7	40.9	58.6	17.7
98 Control	66.5	72.9	6.4	58.7	66.9	8.2	44.4	60.6	16.2
97 Treatment	62.9	73.0	10.1	56.9	67.6	10.7	30.3	56.2	25.9
Experimental treatment effect:									
Cross section		8.0 (4.6)			12.8 (5.7)			7.1 (8.6)	
Difference-in-difference		10.3 (6.7)			14.1 (8.3)			17.7 (12.0)	

	Boys age 12–15			Boys age 12–15, behind in school			Boys age 13–15, HGC ≥ 6, behind in school		
	(1)	(2)	(2)–(1)	(1)	(2)	(2)–(1)	(1)	(2)	(2)–(1)
97 Control	68.8	79.6	10.8	64.0	75.8	11.8	59.0	72.7	13.7
98 Control	72.5	80.2	7.7	67.4	78.0	10.6	57.1	72.8	15.7
97 Treatment	69.5	79.4	9.9	64.2	75.8	11.6	52.6	71.6	19.0
Experimental treatment effect:									
Cross section		3.8 (4.2)			4.2 (5.2)			1.2 (8.4)	
Difference-in-difference		3.1 (6.1)			4.0 (7.4)			3.8 (11.7)	

the post-subsidy year, 1998. The difference-in-difference estimate subtracts from the cross-sectional impact estimate the presubsidy (1997) difference between the groups' attendance rates. Standard errors of these experimental effects are reported as well.

Predicted subsidy effects are shown for the three nonsubsidy groups separately, for the control group in 1997 and 1998, and for the treatment group in 1997. For example, the model predicts an impact of 10.1 percent for treatment group girls age 12 to 15 in 1997. That is, given the state space for households in the treatment group as of (October) 1997, this figure represents the difference between the attendance rate of girls during the 1997/1998 school year that is predicted by the model if the subsidy had been in force and the actual attendance rate. This predicted subsidy effect falls within the range of the experimental estimates (8.0 percent–10.3 percent) and is within one standard deviation of either experimental treatment effect estimate. The model's predicted effect using either the 1997 or 1998 control group also falls well within the one standard deviation. Similarly, the estimates of the subsidy effects for girls who are behind in school are also close to the actual treatment effects, falling within one standard deviation of the experimental treatment effect estimates in most cases.

Predicted subsidy effects are considerably less accurate for boys. The experimental impact estimates for boys are smaller than for girls and are not usually statistically significantly different from zero, although the model's predicted subsidy impacts are of a similar magnitude for girls and boys. For boys age 12 to 15, the predicted effects are 2 to 3 times greater than the experimental treatment effect estimates, although the prediction is within one standard deviation of the experimental effect for at least one group. The result is similar for boys age 12 to 15 and behind in school. The predictions are much worse, however, in terms of the magnitude of the error for boys age 13 to 15 who are behind and have completed the sixth grade.[38] Fn38

As a further evaluation of the model's performance, Table 14 presents evidence on the T14
model's ability to forecast the full school/work/

[38] Interestingly, in the full sample, which includes also landed households, the experimental impacts for boys tend to be of similar magnitude to those of girls (see Behrman et al., 2005).

TABLE 14—ACTUAL AND PREDICTED CHOICE DISTRIBUTION BY CHILD
(*Age, sex, and school attainment: post-subsidy treatment*)

	In school		Home		Work		In school		Home		Work	
	Act[a]	Pred[b]	Act.	Pred.	Act.	Pred.	Act.	Pred.	Act.	Pred.	Act.	Pred.
Age 12–15[c, d]												
All	74.8	74.9	21.2	22.3	4.1	2.8	76.3	77.1	14.9	14.9	8.8	7.9
Not behind	82.3	82.1	14.5	14.8	3.2	3.1	88.9	88.2	9.7	9.7	1.4	2.1
Behind	71.9	72.3	23.7	25.0	4.4	2.7	71.6	73.0	16.8	16.9	11.6	10.2
Behind and HGC $\geq$ 6	52.3	58.7	41.5	37.7	6.2	3.6	58.3	66.7	25.0	20.9	16.7	12.4

[a] Based on observations in which neither the school nor work choice is missing.
[b] Based in all observations, including those missing school or work.
[c] Numbers of observations for each of the four rows are 222, 62, 160, and 65 for girls, and 262, 72, 190, and 96 for boys.
[d] Based on 200 simulation draws per family.

home choice distribution, by sex, for the 1998 treatment group for all children age 12 to 15, for children of those ages not behind in school, those behind in school, and those behind who have completed sixth grade. The model performs well in predicting the rates of staying home or working for pay (usually within 2 percent of the actual rates) and it captures the differences in the work/home pattern between boys and girls, although as before not as well for the last subsample.

B. *N-Step Ahead Predictions from Initial Conditions*

The intent of the Mexican government was to make the PROGRESA program a permanent feature of the social welfare system. As is the case for social experiments in general, however, it is often politically infeasible to deny the control group access to the treatment for a long time period. The short-term nature of the PROGRESA experiment limits its usefulness for evaluating the effect of the policy change on a household's decision-making over the longer term. Indeed, newly formed households will be subject to the program over their entire lifetimes and it would be useful to assess the effect of the program on completed schooling levels of the children borne to those households and on completed fertility within those households. Unlike the households in the experiment whose response to the program was conditioned on the prior fertility and schooling decision that had already been made, households facing the program from the time they are formed would have greater flexibility. For example, as already noted, while it may be difficult over the short term to bring children who have dropped out back into school, the availability of the program from the beginning may prevent dropping out in the first place.

Before assessing the effect of the program in the long-term, however, we evaluate the performance of the model in making long-term forecasts. To do that, we use the model to predict school and fertility outcomes at the survey dates for the nonsubsidy groups using only information on initial conditions at the date of marriage. Given that marital durations at the time of the 1997 survey range from 8 to 38 years, this validation exercise is a more severe within-sample test of the model's performance than the previous test based on one-step-ahead predictions.[39] Fn39

Table 15 shows the actual and predicted school attendance rates obtained from that simulation exercise separately for the control sample in 1997 and 1998 and for the treatment sample in 1997. The predictions represent a within-sample fit test of the model's ability to make accurate long-term forecasts of attendance rates in the absence of the subsidy. The predictions of attendance rates are clearly not as accurate as the short-term predictions and tend to underpredict attendance rates. Nevertheless, the predictions are still reasonably good. For example, for those age 12 to 15, the predicted attendance rates are between 6 and 10 percentage points below the actual rates. In comparison, the difference between them based on the T15

[39] We did not use these forecasts to pre-test the model.

tapraid1/z3y-aer/z3y-aer/z3y00506/z3y2008d06a | yodert | S=42 | 10/12/06 | 13:56 | Art: 20031016 | Input-dld(dld)

TABLE 15—COMPARISON OF ACTUAL AND PREDICTED ATTENDANCE AND FERTILITY BASED ON N-YEAR PREDICTIONS USING INITIAL CONDITIONS

	Controls, 1997		Controls, 1998		Treatments, 1997	
	Actual	Predicted	Actual	Predicted	Actual	Predicted
Percent attending school						
Age 6–11						
Girls	96.9	95.3	96.5	95.4	97.6	95.3
Boys	96.6	93.3	96.7	93.5	97.6	93.2
Age 12–15						
Girls	65.3	58.2	66.5	58.5	62.9	56.6
Boys	68.8	62.5	72.5	62.7	69.5	61.2
Age 12–15, behind in school						
Girls	58.3	52.4	58.7	52.6	56.9	51.1
Boys	64.0	56.4	67.4	56.8	64.2	55.2
Age 12–15, HGC $\geq$ 6, behind in school						
Girls	40.9	41.3	44.4	41.0	30.3	39.6
Boys	59.0	51.1	57.1	50.1	52.6	48.9
Percent pregnant						
Age 20–24	18.0	21.2	17.3	19.6	17.8	20.8
Age 25–29	17.0	20.0	14.8	19.4	16.7	19.8
Age 30–34	13.1	10.8	9.4	10.8	13.1	11.0
Age 35–44	4.9	7.3	6.5	8.1	6.2	7.7

short-term forecasts ranges from 0 to 5 percentage points. The last four rows of the table show the model's predictions of pregnancy rates for different age ranges, which usually are within 2 to 3 percentage points of the actual rates.

VI. Long-Term Impacts and Counterfactual Policy Experiments

A. *The Long-Term Impact of the Program*

Given that the model's long-term forecasts of fertility and school attendance rates are reasonably accurate, we now use the model to predict the long-term impact of exposure to the PROGRESA subsidy program. That is, we predict the effect of the subsidy on family choices for the control and treatment groups at each survey date, assuming that the program had been available to them from the time of marriage.[40] Table 16 compares the short-term and long-term predictions of the program on school attendance rates for girls and boys age 12 to 15. As expected, long-run impacts are larger than the short-run impacts; however, they exceed the short-run effects by only 0.5 to 1.5 percentage points, which suggests that much of the effect of the program on attendance is observed over the short run. As seen below, this result is due, in part, to the unresponsiveness of fertility to the subsidy.

Tables 17 and 18 report estimated long-term impacts on completed schooling and fertility, which are obtained by simulating fertility and schooling outcomes from the mother's age at marriage through age 59, when all the children in the family would be at least 16 years of age. The model predicts that without the subsidy, girls will complete 6.29 years and boys 6.42 years of schooling. Had the program been in existence from marriage, given our estimates, children's mean years of completed education at age 16 would have increased by 0.54 years for both girls and boys. The model also predicts an increase in girls completing sixth grade by 6.4 percentage points and in boys by 4.5 percentage points. Increases in ninth grade completion rates are predicted by the model to be 6 percentage points for girls and 5.3 percentage points for boys.

One concern the Mexican government had in implementing the PROGRESA subsidy pro-

[40] Our long-term forecasts assume that the families are eligible for the program whenever they have grade-eligible children. In reality, a family could become ineligible, for example, by accumulating certain assets, such as a car. Given that our model does not incorporate asset accumulation, we do not take into account that eligibility may change with changes in assets. Our model does, however, allow families to change fertility decisions to become eligible for program subsidies.

Fn40 T16 T17-18

1 LINE LONG

tapraid1/z3y-aer/z3y-aer/z3y00506/z3y2008d06a | yodert | S=42 | 10/12/06 | 13:56 | Art: 20031016 | Input-dld(dld)

TABLE 16—SHORT-RUN AND LONG-RUN EFFECTS OF THE SUBSIDY ON THE PERCENT OF 12- TO 15-YEAR-OLDS ATTENDING SCHOOL

	Girls		Boys	
	Short-run effect[a]	Long-run effect[b]	Short-run effect	Long-run effect
Control group				
1997	10.9	11.9	10.7	12.0
1998	11.2	12.3	11.4	12.7
Treatment group				
1997	11.2	12.3	11.3	12.4
1998	11.7	12.7	12.1	12.4

[a] Predicted value with subsidy minus predicted value without subsidy, conditional on current state space.

[b] Predicted value with subsidy minus predicted value without subsidy, based on initial conditions.

TABLE 17—PREDICTED EFFECT OF THE SUBSIDY ON COMPLETED SCHOOLING OF CHILDREN BY AGE 16: ALL CHILDREN EVER BORN[a]

	Girls		Boys	
	No subsidy	Subsidy	No subsidy	Subsidy
Mean schooling	6.29	6.83	6.42	6.96
Percent completing grade six or more	75.8	82.2	78.8	83.3
Percent completing grade nine or more	19.8	25.9	22.8	28.0

[a] Completed schooling truncated at grade 10.

TABLE 18—PREDICTED EFFECT OF SUBSIDY ON COMPLETED FERTILITY: ALL CHILDREN EVER BORN

	No subsidy	Subsidy
Mean number of children ever born	4.24	4.28
Percent of families with		
Zero children	0.05	0.04
One child	1.16	1.12
Two children	9.23	8.75
Three children	22.97	22.49
Four children	24.43	24.64
Five children	21.54	21.55
Six children	14.78	15.23
Seven children	5.05	5.32

Fn41 gram was that it might induce higher fertility.[41] In fact, concerns about providing incentives for fertility were an important reason why subsidies began only at grade 3. When we evaluate the long-term effects of the program on fertility, we find that fertility outcomes are essentially invariant to the subsidies. Without the subsidy, the predicted long-run average number of children is 4.24, compared to 4.28 with the subsidy. The program also induces only minor changes in the distribution of numbers of children across families. For example, the prevalence of families having four or more children increases by 1 percentage point.

B. *Alternative Subsidy Programs*

Designing an optimal subsidy scheme to achieve some desired increase in schooling requires knowledge of the effects of many alternative subsidy schedules and of households' take-up decisions under these alternative programs. As noted earlier, a limitation of experiments is that they do not typically provide a reliable way of extrapolating to learn about effects of counterfactual policies. Although a small change in the subsidy schedule might be well approximated by a simple extrapolation of the experimental

[41] In a static quality-quantity fertility model, Willis (1973) shows that a decrease in the per-child price of quality, e.g., a subsidy to school attendance, will increase fertility as long as quality and quantity are complements in utility, which is what we find. We would expect the basic intuition of that model also to hold in the more complex model we estimate.

TABLE 19—THE EFFECTIVENESS AND COST OF ALTERNATIVE PROGRAMS

	Baseline[a]	Compulsory school attendance through age 15	Original subsidy	2× subsidy	0.5× subsidy	Restricted subsidy[b]	1.43× restricted subsidy
Mean completed schooling							
Girls	6.29	8.37	6.83	7.30	6.56	6.67	6.97
Boys	6.42	8.29	6.96	7.44	6.68	6.79	7.07
Percent completed grade 6 or more							
Girls	75.8	95.1	82.3	86.9	79.3	77.4	82.0
Boys	78.8	93.7	83.3	86.7	81.1	79.6	82.8
Percent completed grade 9 or more							
Girls	19.8	55.5	25.9	31.6	23.1	26.2	29.3
Boys	22.8	54.7	28.0	34.6	25.5	29.2	31.8
Cost per family	0	—	26,096	59,935	11,989	15,755	25,193
Mean number of children	4.24	4.21	4.28	4.32	4.27	4.25	4.27

	Bonus for completing 9th grade[c]	Junior secondary school in each village	Unconditional income transfer 5,000 pesos/yr	No child labor through age 15	Original subsidy and 25% wage increase
Mean completed schooling					
Girls	6.50	6.39	6.41	6.30	6.75
Boys	6.58	6.55	6.53	6.52	6.79
Percent completed grade 6 or more					
Girls	74.9	76.0	77.6	76.1	81.5
Boys	76.9	79.0	80.0	79.9	81.8
Percent completed grade 9 or more					
Girls	28.8	21.2	20.8	19.7	25.2
Boys	32.7	24.1	23.7	23.5	26.5
Cost per family	36,976	—	237,000	—	25,250
Mean number of children	4.20	4.24	4.24	4.25	4.29

[a] Predicted: control and treatment families.
[b] Subsidy for attending school in grades 6–9 only.
[c] The bonus is set at 30,000 pesos for girls and boys.

treatment effect, any extrapolation to a more radical change in the subsidy schedule would be ad hoc. For example, one might be interested in evaluating an unconditional income grant to families that removes the school attendance requirement.

Using the estimated behavioral model, we simulate the effects of a variety of counterfactual policy experiments that are alternative ways
T19 to increase school attainment. Table 19 reports the results based on simulations to mother's age 59, as in the previous two tables. The first column reports the predicted completed schooling and fertility for the baseline of no program. To establish an upper bound for the effect of alternative subsidy schemes on school completion levels, the second column reports the effect of a perfectly enforced school attendance requirement for all children between the ages of 6 and 15. Although maximum completed schooling by age 16 is 10 years, because failure rates are significant, mean completed schooling with compulsory attendance is only 8.37 years for girls and 8.29 for boys, an increase over the baseline of about 2 years for both.[42] Fertility declines Fn42
slightly with compulsory school attendance.

The next column shows the effect of the original subsidy scheme (as in the previous two tables), and the following two columns report

[42] Failure rates differ significantly among the types. For example, type 2s complete almost 9 years of schooling under the compulsory school attendance requirement, while type 3s complete fewer than 6.5 years.

tapraid1/z3y-aer/z3y-aer/z3y00506/z3y2008d06a | yodert | S=42 | 10/12/06 | 13:56 | Art: 20031016 | Input-dld(dld)

experiments that simply vary the subsidy amounts, first doubling and then halving them. Completed fertility increases slightly as subsidy levels rise, but is relatively invariant over the range of changes considered. Mean completed schooling increases at a linear rate with increments in subsidy amounts up to the original amount and then at a slightly diminishing rate. For example, for girls the increases in mean schooling between the baseline and one-half of the original subsidy and between one-half and the full original subsidy amount are both 0.27, while the increase from doubling the subsidy amount is 0.47 years. However, whether there are diminishing returns to the program depends, in addition, on how the total cost of the program increases with the subsidy levels. The next-to-the-last row of the table, which calculates the average cost of the subsidy program on a per-family basis, shows that doubling the subsidy amounts more than doubles the per-family cost. Based on these figures, the change in average schooling induced by a unit change in total costs is 41 percent higher at the one-half subsidy level than at the full subsidy level.

The next column restricts the subsidy to attendance in the sixth grade or higher, that is, the subsidy is zero for attending grades 3 through 5. Because the great majority of children complete at least the fifth grade, the subsidy to the earlier grades acts mainly as a direct income transfer program and, therefore, might have only a weak effect on schooling although a strong effect on the cost of the program. As seen, restricting the subsidy to higher grades reduces the per-family cost of the program considerably, from around 26,000 pesos to less than 16,000 pesos. Perhaps surprisingly, however, the fall in completed schooling is also not insignificant, with approximately 30 percent of the gain in mean schooling for girls and 33 percent for boys being lost.[43] Fn43

The reason that restricting the subsidy to attendance in grades 6 and higher diminishes nonnegligibly the gain in completed schooling levels, given that there is almost universal attendance through grade 5 without the subsidy, is due to the interdependence of parental decisions among children within the family. If there are multiple children of school age in the household, providing a subsidy to attendance for children at lower grades, at least in part because it increases family income, reduces the incentive for parents to have older children work. That this intra-household allocation mechanism is important can be seen from the fact that the reduction in the gain from the subsidy under the restriction falls with the number of children ever born. There is no reduction in one-child families, a 12.5-percent reduction in two-child families, a 16.7-percent reduction in three-child families, a 29-percent reduction in four-child families, and a 35-percent reduction in families with five or more children.

In light of this finding, the next column reports an experiment in which the subsidy is again restricted to attendance in grades 6 through 9, but the subsidy schedule is set at a level (1.43 times the subsidy amounts at each grade) at which the cost per family is the same as the original subsidy without the restriction (column 3). The gain in mean completed schooling is predicted to be 0.14 years more for girls and 0.11 years more for boys than the original subsidy, an increase of about 25 percent over the original gain. There is also a difference, as compared to the original subsidy, in the distribution of completed schooling, with a very slight increase in the proportion of children completing fewer than six years of schooling and a significant increase in the proportion completing nine or more years. Ignoring distributional impacts, namely that families whose children do not attend the higher grade levels receive no income transfer under the restricted subsidy program, the restricted subsidy program would appear to be more efficient in producing higher completed schooling levels in this population.

An alternative subsidy scheme rewards grade completion rather than attendance. In the next column, we assess the impact of the ninth grade graduation bonus in the form of a payment of 30,000 pesos to families when a child graduates from junior secondary school.[44] Clearly, the Fn44

[43] Moreover, relative to the original subsidy, there is a substantially smaller gain in the fraction of children graduating from elementary school (sixth grade), although offset by a slight gain in the fraction of children graduating from junior secondary school (ninth grade). With the restricted subsidy, 77.6 percent of girls and 80.0 percent of boys graduate from elementary school, an increase of 1.6 percentage points for girls and 0.8 for boys as compared to increases of 6.5 and 4.5 percentage points with the original subsidy.

[44] Keane and Wolpin (2000) assess such bonus-type schemes in the U.S. context.

1 LINE LONG

tapraid1/z3y-aer/z3y-aer/z3y00506/z3y2008d06a | yodert | S=42 | 10/12/06 | 13:56 | Art: 20031016 | Input-dld(dld)

effect of such a bonus scheme requires, as the model assumes, that families be forward-looking.[45] The simulations show that the bonus increases the percentage of children completing junior secondary school significantly, by about 10 percentage points for both girls and boys, but has a relatively small impact on average schooling. In fact, the increase in average schooling is not as large as the effect of the original subsidy, even though the cost of the bonus program is about 40 percent higher. Interestingly, the proportion of children who complete at least sixth grade actually falls below the nonsubsidy level, suggesting that, in order to earn the bonus, families are substituting more schooling for some children and less for others. In fact, as evidence for this behavior, the within-family coefficient of variation in completed schooling increases by 25 percent with the bonus, as compared to the no-subsidy case. Thus, the effect of the bonus is largely to induce children who were already attending junior secondary school to complete ninth grade.

The additional interventions we consider all have relatively minor effects on schooling. In particular, enforcing a child labor law that prohibits children under the age of 16 from working has almost no effect on completed schooling for girls, and increases mean schooling for boys by 0.1 years. Relative to the subsidy scheme, which draws children from both the work and home alternatives, restricting work does not change the relative attractiveness of attending school and remaining at home. Building a junior secondary school in each village where it is absent, thereby setting to zero the cost of attending a secondary school due to its distance from the village, would raise mean schooling by 0.1 years for girls and by 0.13 years for boys. We also simulate the impact of a pure income transfer program, one that pays 5,000 pesos per year to families without any school attendance requirement. This amount is close to the maximum benefit that families may currently receive under the program in any year, and represents about a 50-percent increase in annual family income. As seen, mean schooling increases due to the subsidy because schooling is a normal good.[46] However, the increase in schooling is only about 20 percent as large as the original attendance-based subsidy. Moreover, its cost per family is an order of magnitude larger.

An important caveat to our evaluation of the long-term impact of counterfactual experiments is that our analysis is partial equilibrium. As school attendance rates rise due to the program, and children withdraw from the child labor market, one would expect child wage rates to rise and the increase in school attendance rates due to the subsidy to be somewhat mitigated. The fact that our forecast of the subsidy effect does not incorporate such a labor market equilibrium response, and is yet reasonably accurate, may imply that such adjustments are small. Nevertheless, to get some idea of the quantitative significance of potential equilibrium effects, we also performed a counterfactual experiment which combines the original subsidy program with a concomitant increase in child wage rates of 25 percent.[47] The results of that experiment are shown in the last column of Table 19. The degree of mitigation of the increase in mean schooling differs by sex. The increase above the baseline accounting for the wage increase is 85 percent of the partial equilibrium effect of the original subsidy for girls and 69 percent for boys. Of course, this example is at best illustrative because we do not know how elastic the market demand for child labor is. A complete analysis would require a general equilibrium model of the rural labor market, which we leave for future work.

VII. Conclusions

The value of empirically determining the underlying structure of economic relations in order to evaluate the impact of policy interventions that radically depart from past experience is well understood (Jacob Marschak, 1953;

Fn45 [45] The effect of a bonus program would also be sensitive to assumptions about the ability of families to smooth consumption intertemporally. Increased opportunities to smooth consumption through saving/borrowing would presumably increase the value of large lump-sum payments. Recall that, in the model, families cannot smooth consumption between periods at all.

Fn46 [46] As with all of the other subsidy schemes, fertility, though falling with the income subsidy, is essentially unchanged.

Fn47 [47] We assume that parental income would be unchanged, although there may be an effect on adult wages depending on the degree to which child and adult labor are substitutes.

1 LINE LONG

tapraid1/z3y-aer/z3y-aer/z3y00506/z3y2008d06a | yodert | S=42 | 10/12/06 | 13:56 | Art: 20031016 | Input-dld(dld)

Heckman, 2000). However, the credibility of applications of structural estimation in forecasting the impact of new policies is often a matter of dispute because of concerns about the validity of modeling assumptions. Within-sample goodness-of-fit tests provide useful but not necessarily compelling evidence of the validity of the model, because pre-testing of the model's structure on the estimation sample is a common practice.[48] To mitigate the effects of pre-test estimation, there have been some attempts to assess model validity using out-of-sample forecasts. Such applications are sparse, however, and are limited by the nature of the data.[49]

Fn48

Fn49

An alternative to using observational data to evaluate the impact of new policies is to design and implement randomized social experiments. However, social experiments also have some common drawbacks. For example, they are usually of short duration, making it difficult to assess long-term effects of social programs, are very costly, and do not typically provide much variation in the range of treatments.

In this paper, we demonstrated a potential synergy between social experimentation and observational methods that can be exploited to overcome the limitations of each approach for use in policy analysis. In particular, if the data collected from the experiment are rich enough to estimate a behavioral model, then social experiments provide an opportunity for out-of-sample validation. Using data from the PROGRESA experiment, we estimated a behavioral model of family decisions about fertility and schooling without using post-program data on treated households. We validated the model by comparing its predictions about program impacts to those estimated directly from the experiment. The model produced reasonable forecasts of the effect of the program on school attendance rates of children. In our view, this evidence lent sufficient support to the model to use it to simulate the effects of a number of counterfactual policy experiments, which illustrate a menu of options that might be available to policymakers. We estimated the benefits and costs associated with alternative programs, such as doubling the subsidy at all grade levels, halving it, restricting it to higher grade levels, providing a graduation bonus, and providing pure income transfers without the school attendance requirement. The simulations showed that eliminating subsidies at lower grade levels and using the savings to increase the subsidy levels at higher grade levels would lead to a greater increase in average schooling completed. This last example shows more generally that, given a specific policy objective, the model can contribute to the design of an optimal program.

[48] Nonstructural estimation, when it can be used for the same purpose, suffers from the same limitations. Parametric assumptions are required to extrapolate outside the range of existing policy variation, and specification pre-testing is also widespread.

[49] For example, Keane and Wolpin (1997) forecast occupational choices for cohorts that were not used in estimating their model. Robin L. Lumsdane et al. (1992) estimate a model of retirement behavior using data prior to a regime change and then use the model to predict the effects of a new program giving incentives for workers in certain age ranges to retire early.

1 LINE LONG

APPENDIX A

Utility Function

$$u_t = \frac{1}{\lambda_{00}} C(t)^{\lambda_{00}} \left[1 + \exp(\lambda_{01} N(t) + \lambda_{02}\bar{S}(t) + \lambda_{03} \sum_{j=12}^{15} 1(t, t-j) n(t-j) \right] \tag{A1}$$

$$+ \sum_{j=1}^{3} \lambda_{1,j} \mathrm{I}(\text{type} = j) N(t) - \sum_{j=1}^{3} \lambda_{2,j} \mathrm{I}(\text{type} = j) N(t)^2 + \sum_{j=1}^{3} \lambda_{3,j} \mathrm{I}(\text{type} = j) \bar{S}(t)$$

$$+ \sum_{j=1}^{3} \lambda_{4,j} \mathrm{I}(\text{type} = j) N(t) \bar{S}(t) + \lambda_5 \sum_{j=1}^{3} n(t-j) - \lambda_6 \sum_{j=1}^{3} n(t-j)^2$$

$$+ \sum_{\mathrm{j}=1}^{3} \lambda_{7,\mathrm{j}} \mathrm{I}(\text{type} = j) \sum_{n} \mathrm{I}(S(t, \tau_n) \geq 6) + \sum_{j=1}^{3} \lambda_{8,j} \mathrm{I}(\text{type} = j) \sum_{n} \mathrm{I}(S(t, \tau_n) \geq 9)$$

$$+ \lambda_9 \mathrm{I}(p(t_m) = 1) + \lambda_{10} \mathrm{I}\left(\sum_{t=20}^{24} p(t) = 1 \right) + \lambda_{11} \mathrm{I}\left(\sum_{t=25}^{29} p(t) = 1 \right)$$

$$+ \lambda_{12} \mathrm{I}\left(\sum_{t=30}^{34} p(t) = 1 \right) + \lambda_{13} \mathrm{I}\left(\sum_{t=35}^{39} p(t) = 1 \right) + \lambda_{14} \mathrm{I}\left(\sum_{t=40}^{43} p(t) = 1 \right)$$

$$+ \lambda_{15} \mathrm{I}(p(t-1) = 1) \mathrm{I}(p(t) = 1) + \lambda_{16} \sum_{k=12}^{15} \mathrm{I}(S_b(t, \tau_n) \geq 6, s_b(t, \tau_n) = 1) | t - \tau_n = k))$$

$$+ \lambda_{17} \sum_{\mathrm{k}=12}^{15} \mathrm{I}(S_g(t, \tau_n) \geq 6, s_g(t, \tau_n) = 1) | t - \tau_n = k))$$

$$+ \sum_{\mathrm{j}=1}^{3} \lambda_{18,\mathrm{j}} \mathrm{I}(\text{type} = j) \sum_{k=12}^{15} \mathrm{I}(S(t, \tau_n) - (t - \tau_n - 6) = 0, s(t, \tau_n) = 1) | t - \tau_n = k))$$

$$+ \sum_{\mathrm{j}=1}^{3} \lambda_{19,\mathrm{j}} \mathrm{I}(\text{type} = j) \sum_{k=12}^{15} \mathrm{I}(S(t, \tau_n) - (t - \tau_n - 6) \geq 1, s(t, \tau_n) = 1) | t - \tau_n = k))$$

$$+ \lambda_{20} \sum_{\mathrm{k}=12}^{15} \mathrm{I}(S(t, \tau_n) - (t - \tau_n - 6) \geq 2, s(t, \tau_n) = 1) | t - \tau_n = k))$$

$$+ \lambda_{21} \sum_{k=12}^{15} \mathrm{I}(S(t, \tau_n) - (t - \tau_n - 6) \geq 3, s(t, \tau_n) = 1) | t - \tau_n = k))$$

$$+ \lambda_{22} \sum_{\mathrm{k}=12}^{15} \mathrm{I}(S_b(t, \tau_n) - (t - \tau_n - 6) = 0, s_b(t, \tau_n) = 1) | t - \tau_n = k))$$

$$+ \lambda_{23} \sum_{\mathrm{k}=12}^{15} \mathrm{I}(S_b(t, \tau_n) - (t - \tau_n - 6) \geq 1, s_b(t, \tau_n) = 1) | t - \tau_n = k))$$

1 LINE LONG

$$+ \lambda_{24} \sum_{k=12}^{15} \mathrm{I}(S_b(t, \tau_n) - (t - \tau_n - 6) \geq 2, s_b(t, \tau_n) = 1) | t - \tau_n k))$$

$$+ \lambda_{25} \sum_{k=12}^{15} \mathrm{I}(S_b(t, \tau_n) - (t - \tau_n - 6) \geq 3, s_b(t, \tau_n) = 1) | t - \tau_n = k))$$

$$+ \lambda_{26} \sum_{k=12}^{15} \mathrm{I}(S(t, \tau_n) - (t - \tau_n - 6) \geq 1, 1(t, \tau_n) = 1) | t - \tau_n = k))$$

$$+ \lambda_{27} \sum_{k=12}^{15} \mathrm{I}(S(t, \tau_n) - (t - \tau_n - 6) \geq 2, 1(t, \tau_n) = 1) | t - \tau_n = k))$$

$$+ \lambda_{28} \sum_{k=12}^{15} \mathrm{I}(S(t, \tau_n) - (t - \tau_n - 6) \geq 3, 1(t, \tau_n) = 1) | t - \tau_n = k))$$

$$+ \lambda_{29} \sum_{k=12}^{15} \mathrm{I}(S_b(t, \tau_n) - (t - \tau_n - 6) \geq 1, 1_b(t, \tau_n) = 1) | t - \tau_n = k))$$

$$+ \lambda_{30} \sum_{k=12}^{15} \mathrm{I}(S_b(t, \tau_n) - (t - \tau_n - 6) \geq 2, 1_b(t, \tau_n) = 1) | t - \tau_n = k))$$

$$+ \lambda_{31} \sum_{k=12}^{15} \mathrm{I}(S_b(t, \tau_n) - (t - \tau_n - 6) \geq 3, 1_b(t, \tau_n) = 1) | t - \tau_n = k))$$

$$+ \lambda_{32} \mathrm{I}(s(t, \tau_n) = 1 | t - \tau_n = 12) n(t - 12) + \sum_{j=1}^{3} \lambda_{33,j} \mathrm{I}(\text{type} j) \mathrm{I}(1_b(t, \tau_n = 1 | t - \tau_n = 6)$$

$$+ \sum_{j=1}^{3} \lambda_{34,j} \mathrm{I}(\text{type} = j) \mathrm{I}(1_b(t, \tau_n = 1 | t - \tau_n = 7)$$

$$+ \sum_{j=1}^{3} \lambda_{35,j} \mathrm{I}(\text{type} = j) \mathrm{I}(1_g(t, \tau_n = 1 | t - \tau_n = 6)$$

$$+ \sum_{j=1}^{3} \lambda_{36,j} \mathrm{I}(\text{type} = j) \mathrm{I}(1_g(t, \tau_n = 1 | t - \tau_n = 7)$$

$$+ \sum_{j=1}^{3} \lambda_{37,j} \mathrm{I}(\text{type} = j) \sum_{k=8}^{11} \mathrm{I}(1_b(t, \tau_n = 1 | t - \tau_n = k)$$

$$+ \sum_{j=1}^{3} \lambda_{38,j} \mathrm{I}(\text{type} = j) \sum_{k=12}^{15} \mathrm{I}(1_b(t, \tau_n = 1 | t - \tau_n = k)$$

$$+ \sum_{j=1}^{3} \lambda_{39,j} \mathrm{I}(\text{type} = j) \sum_{k=8}^{11} \mathrm{I}(1_g(t, \tau_n = 1 | t - \tau_n = k)$$

1 LINE LONG

$$+ \sum_{j=1}^{3} \lambda_{40,j} \mathrm{I}(\text{type} = j) \sum_{k=12}^{15} \mathrm{I}(1_g(t, \tau_n = 1 | t - \tau_n = k)$$

$$+ \lambda_{41} \sum_{j=0}^{5} \mathrm{n}(\mathrm{t-j}) \sum_{k=14}^{15} \mathrm{I}(1_g(t, \tau_n = 1 | t - \tau_n = k)$$

$$+ \lambda_{42} \sum_{j=0}^{5} \mathrm{n}(\mathrm{t-j}) \sum_{k=12}^{15} \mathrm{I}(1_g(t, \tau_n) = 1 | t - \tau_n = k)$$

$$+ \lambda_{43} \mathrm{z_s} \sum_{n} \mathrm{I}(S(t, \tau_n) \geq 6, s(t, \tau_n = 1))$$

$$+ \sum_{j=1}^{3} \lambda_{44,j} \mathrm{I}(\text{type} = j) \sum_{n} \mathrm{I}(S(t, \tau_n) = 9, s(t, \tau_n) = 1)$$

$$+ \lambda_{45} \sum_{n} \mathrm{I}(S_b(t, \tau_n) = 9, s_b(t, \tau_n) = 1)$$

$$+ \sum_{j=1}^{3} \lambda_{46,j} \mathrm{I}(\text{type} = j)\varepsilon_p(t) + \varepsilon_{1_b}(t) \Bigg[\sum_{k=12}^{15} 1_b(t, \tau_n = 1 | t - \tau_n = k)$$

$$+ \lambda_{47} \sum_{k=6}^{11} 1_{\mathrm{b}}(\mathrm{t}, \tau_{\mathrm{n}} = 1 | \mathrm{t} - \tau_{\mathrm{n}} = \mathrm{k}) + \lambda_{48} \sum_{k=12}^{13} 1_{\mathrm{b}}(\mathrm{t}, \tau_{\mathrm{n}} = 1 | \mathrm{t} - \tau_{\mathrm{n}} = \mathrm{k}) \Bigg]$$

$$+ \varepsilon_{1_g}(\mathrm{t}) \Bigg[\sum_{k=12}^{15} 1_{\mathrm{g}}(\mathrm{t}, \tau_{\mathrm{n}} = 1 | \mathrm{t} - \tau_{\mathrm{n}} = \mathrm{k}) + \lambda_{49} \sum_{k=6}^{11} 1_{\mathrm{g}}(\mathrm{t}, \tau_{\mathrm{n}} = 1 | \mathrm{t} - \tau_{\mathrm{n}} = \mathrm{k})$$

$$+ \lambda_{50} \sum_{k=12}^{13} 1_{\mathrm{g}}(\mathrm{t}, \tau_{\mathrm{n}} = 1 | \mathrm{t} - \tau_{\mathrm{n}} = \mathrm{k}) \Bigg]$$

Budget Constraint

$$\text{(A2)} \qquad C(t) = y_p(t) + \sum_{n} y_o(t, \tau_n) h(t, \tau_n)$$

Parent Income Function

$$\text{(A3)} \qquad \log y_p(t) = \sum_{j=1}^{3} \gamma_{0,j}^{p} \mathrm{I}(\text{type} = j) + \gamma_1^p a_p(t) - \gamma_2^p a_p(t)^2 + \gamma_3^p z_c + \varepsilon_{y_p}(t)$$

Child Wage Function

$$\text{(A4)} \qquad \log y_0(t, \tau_n) = \sum_{j=1}^{3} \gamma_{0,j}^{o} \mathrm{I}(\text{type} = j) + \gamma_1^o \mathrm{I}(b(t, \tau_n) = 1) + \gamma_2^o z_c + \gamma_3^o (t - \tau_n)$$

$$+ \gamma_4^{\mathrm{o}} \mathrm{I}(14 \leq t - \tau_n \leq 15) + \gamma_5^o b(\tau_n) \mathrm{I}(14 \leq t - \tau_n \leq 15)$$

$$+ \gamma_6^{\mathrm{o}} \mathrm{I}(t - \tau_n = 15) + \gamma_7^o b(\tau_n) \mathrm{I}(t - \tau_n = 15) + \varepsilon_{y_o}(t)$$

School Failure Probability Function

$$\pi^c(t,\ \tau_n|s(t,\ \tau_n = 1)) = \Bigg[1 + \exp\Bigg(\sum_{j=1}^{3}\pi_{0,j}\mathrm{I}(\text{type} = j) + \pi_1 S(t, \tau_n) + \pi_2(t - \tau_n) + \pi_3 b(\tau_n) + \pi_4\mathrm{I}(8 \le t - \tau_n \le 15)\mathrm{I}(S(t, \tau_n = 0) + \sum_{j=1}^{3}\Bigg)\pi_{5,j}\mathrm{I}(\text{type} = j)\mathrm{I}(S(t, \tau_n \ge 7) + \pi_6 b(\tau_n)\mathrm{I}(S(t, \tau_n \ge 7))\Bigg]^{-1} \tag{A5}$$

Type Probability Function

1. Couples married seven years or less at 1997 survey date.

$$\Pr(\text{type} = j) = \exp(X\beta_j)/1 + \sum_{k=1}^{2}\exp(X\beta_k), \quad \text{for } 1, 2, \tag{A6}$$

where $X\beta_j = \beta^1_{0,j} + \beta^1_{1,j}t_m + \beta^1_{2,j}a_p(t_m) + \beta^1_{3,j}z_c + \beta^1_{4,j}z_s + \beta^1_{5,j}\mathrm{I}(\max(S^m, S^f) \ge 9)$.

2. Couples married eight years of more at 1997 survey date.

$$\Pr(\text{type} = j) = \exp(X\beta_j)/1 + \sum_{k=1}^{2}\exp(X\beta_k), \quad \text{for } j = 1, 2, \quad \text{where } X\beta_j = \beta^2_{0,j} + \beta^2_{1,j}t_m \tag{A7}$$

$$+ \beta^2_{2,j}a_p(t_{1997}) + \beta^2_{3,j}z_c + \beta^2_{4,j}z_s + \beta^2_{5,j}\mathrm{I}(\max(S^m, S^f) \ge 9) + \beta^2_{6,j}\sum_n \mathrm{I}(14 \le (t_{1997} - \tau_n) \le 15)$$

$$+ \beta^2_{7,j}\sum_n \mathrm{I}(12 \le (t_{1997} - \tau_n) \le 15) + \beta^2_{8,j}N(t_{1997}) + \beta^2_{9,j}\bar{S}(t_{1997})$$

$$+ \beta^2_{10,j}t_{1997} + \beta_{11,j}\sum_n \mathrm{I}(S(t_{1997}, \tau_n) \ge 7)$$

$$+ \beta_{12,j}\mathrm{I}\left(\sum_{j=0}^{5} n(t_{1997} - j) \ge 3 \ \vee \ \sum_{j=6}^{11} n(t_{1997} - j) \ge 3 \ \vee \ \sum_{j=12}^{15} n(t_{1997} - j) \ge 2\right)$$

Appendix B

Table B1—Parameterizations and Parameter Estimates

I. Utility function

Variable	Estimate (s.e.)	Variable	Estimate (s.e.)	Variable	Estimate (s.e.)
CRRA parameter: λ_{00}	0.8715 (0.0190)	Number of boys age 12–15, with 6 years of school and currently attending school: λ_{16}	−78.78 (245.9)	Number of boys age 8–11 at home	
				Type 1: $\lambda_{37,1}$	582.9 (340.8)
				Type 2: $\lambda_{37,2}$	807.9 (4326)
				Type 3: $\lambda_{37,3}$	1043 (249.0)
Number of children		Number of girls age 12–15, with 6 years of school and currently attending school: λ_{17}	63.97 (274.1)	Number of boys age 12–15 at home	
Type 1: $\lambda_{1,1}$	2997 (1011)			Type 1: $\lambda_{38,1}$	1596 (370.8)
Type 2: $\lambda_{1,2}$	433.6 (1135)			Type 2: $\lambda_{38,2}$	1065 (555.4)
Type 3: $\lambda_{1,3}$	3818 (1512)			Type 3: $\lambda_{38,3}$	1085 (348.0)
Number of children squared		Number of children age 12–15 not behind in school and currently attending school		Number of girls age 8–11 at home	
Type 1: $\lambda_{2,1}$	1023 (54.92)	Type 1: $\lambda_{18,1}$	−385.0 (289.4)	Type 1: $\lambda_{39,1}$	333.0 (273.5)
Type 2: $\lambda_{2,2}$	1111 (67.13)	Type 2: $\lambda_{18,2}$	84.1 (188.1)	Type 2: $\lambda_{39,2}$	1179.4 (563.9)
Type 3: $\lambda_{2,3}$	1034 (264.7)	Type 3: $\lambda_{18,3}$	162.2 (316.5)	Type 3: $\lambda_{39,3}$	502.8 (221.2)
Consumption × number of children net of current birth: λ_{01}	−0.0014 (0.0202)	Number of children age 12–15, behind 1 year in school and currently attending school		Number of girls age 12–15 at home	
		Type 1: $\lambda_{19,1}$	−222.4 (251.8)	Type 1: $\lambda_{40,1}$	1815 (368.0)
		Type 2: $\lambda_{19,2}$	−266.3 (209.1)	Type 2: $\lambda_{40,2}$	969.3 (499.1)
		Type 3: $\lambda_{19,3}$	16.00 (244.2)	Type 3: $\lambda_{40,3}$	629.2 (352.1)
Consumption × average schooling: λ_{02}	0.0017 (0.0092)	Number of children age 12–15, behind 2 years in school and currently attending school: λ_{20}	−32.08 (170.6)	Number of girls age 14–15 at home × number of children age 0–5: λ_{41}	287.5 (570.7)
Number of children age 12–15 at home × consumption: λ_{03}	−0.1101 (0.0229)	Number of children age 12–15, behind 3+ years in school and currently attending school: λ_{21}	−10.80 (187.3)	Number of girls age 12–15 at home × number of children age 0–5: λ_{42}	3.27 (423.3)
Average schooling of all children		Number of boys age 12–15 currently attending school and behind × years in school		Number of children attending a secondary school × distance from a secondary school: λ_{43}	0.1298 (0.0076)
Type 1: $\lambda_{3,1}$	−12.16 (93.12)	0 years: λ_{22}	−156.03 (201.6)		
Type 2: $\lambda_{3,2}$	276.24 (87.68)	1 year: λ_{23}	32.16 (197.6)		
Type 3: $\lambda_{3,3}$	25.10 (81.52)	2 years: λ_{24}	−47.15 (194.5)		
		3+ years: λ_{25}	−138.7 (229.8)		

APPENDIX B1—*Continued.*

I. Utility function

Variable	Estimate (s.e.)
Stock of children × average schooling of children age 6–15	
Type 1: $\lambda_{4,1}$	121.4 (23.97)
Type 2: $\lambda_{4,2}$	134.64 (31.80)
Type 3: $\lambda_{4,3}$	113.81 (27.88)
Number of children age 0–2: λ_5	473.9 (7368)
Number of children age 0–2 squared: λ_6	−811.8 (1570)
Number of children with 6 or more years of schooling	
Type 1: $\lambda_{7,1}$	184.6 (41.66)
Type 2: $\lambda_{7,2}$	11.50 (35.82)
Type 3: $\lambda_{7,3}$	−54.21 (24.14)
Number of children with 9 or more years of schooling	
Type 1: $\lambda_{8,1}$	1.43 (48.41)
Type 2: $\lambda_{8,2}$	147.4 (43.05)
Type 3: $\lambda_{8,3}$	6.06 (44.92)
Pregnancy at: First yr. of marriage: λ_9	31834 (2194)
Age 20–24: λ_{19}	−1126 (1111)
25–29: λ_{11}	−3072 (1869)
30–34: λ_{12}	−24414 (3209)
35–39: λ_{13}	−27203 (3905)
40–43: λ_{14}	−59672 (8182)
Pregnancy × pregnancy in previous period: λ_{15}	−37001 (2128)

Variable	Estimate (s.e.)
Number of children currently at home and behind × years	
1 year: λ_{26}	475.3 (326.9)
2 years: λ_{27}	83.28 (350.8)
3 years: λ_{28}	284.2 (254.5)
Number of boys currently at home and behind × years in school	
1 year: λ_{29}	−641.1 (412.9)
2 years: λ_{30}	−86.90 (452.7)
3 years: λ_{31}	−264.7 (326.9)
Age 12 and in school: λ_{32}	186.85 (205.71)
Boy age 6 at home	
Type 1: $\lambda_{33,1}$	1205.91 (370.83)
Type 2: $\lambda_{33,2}$	1421.85 (1723.27)
Type 3: $\lambda_{33,3}$	1579.79 (361.94)
Boy age 7 at home	
Type 1: $\lambda_{34,1}$	697.49 (2235)
Type 2: $\lambda_{34,2}$	644.18 (3438466)
Type 3: $\lambda_{34,3}$	935.4 (251.4)
Girl age 6 at home	
Type 1: $\lambda_{35,1}$	582.9 (849.2)
Type 2: $\lambda_{35,2}$	1677 (493.7)
Type 3: $\lambda_{35,3}$	885.9 (266.5)
Girl age 7 at home	
Type 1: $\lambda_{36,1}$	356.1 (576.6)
Type 2: $\lambda_{36,2}$	1072.0 (1281)
Type 3: $\lambda_{36,3}$	680.2 (403.4)

Variable	Estimate (s.e.)
Number children attending grade 10	
Type 1: $\lambda_{44,1}$	1929 (4897)
Type 2: $\lambda_{44,2}$	1061 (249.9)
Type 3: $\lambda_{44,3}$	443.5 (922.3)
Number of boys attending grade 10: λ_{45}	673.8 (315.6)
Pregnancy shock	
Type 1: $\lambda_{46,1}$	1.0
Type 2: $\lambda_{46,2}$	0.907 (0.029)
Type 3: $\lambda_{46,3}$	0.9733 (0.1305)
Shock to preferences for boys age 12–15 at home × number of boys at home age	
6–11: λ_{47}	.3104 (2468)
12–13: λ_{48}	−.1192 (0.1380)
Shock to preferences for girls age 12–15 at home × number of girls at home age	
6–11: λ_{49}	.8793 (2469)
12–13: λ_{50}	−.2412 (0.1552)

tapraid1/z3y-aer/z3y-aer/z3y00506/z3y2008d06a | yodert | S=42 | 10/12/06 | 13:56 | Art: 20031016 | Input-DCT-msh

APPENDIX B1—*Continued.*

II. Parent earnings function, child earnings function, and failure probability function

II. Parent income function

Constant	
Type 1: $\gamma^p_{0,1}$	8.869 (0.0952)
Type 2: $\gamma^p_{0,2}$	9.0683 (0.0943)
Type 3: $\gamma^p_{0,3}$	8.7954 (0.1298)
Husband's age: γ^p_1	0.0203 (0.0049)
Husband's age squared: γ^p_2	0.0005 (0.0001)
Distance of village to nearest city: γ^p_3	−0.0018 (0.0001)

III. Child income function

Constant	
Type 1: $\gamma^o_{0,1}$	6.984 (1.448)
Type 2: $\gamma^o_{0,2}$	6.920 (1.474)
Type 3: $\gamma^o_{0,3}$	6.984 (1.475)
Child is a boy: γ^o_1	0.362 (0.1681)
Distance of village to nearest city: γ^o_2	−0.0001 (0.0005)
Child's age: γ^o_3	0.0281 (0.1150)
Child is age 14–15: γ^o_4	0.5209 (0.2262)
Child is a boy × child is age 14–15: γ^o_5	−0.0535 (0.2021)
Child is age 15: γ^o_6	0.1247 (0.2032)
Child is a boy × child is age 15: γ^o_7	−0.0805 (0.1953)

IV. School failure probability function

Constant	
Type 1: $\pi_{0,1}$	−2.214 (0.2773)
Type 2: $\pi_{0,2}$	−2.773 (0.2792)
Type 3: $\pi_{0,3}$	−1.251 (0.3550)
Highest grade completed: π_1	−0.1264 (0.0469)
Child age: π_2	0.0956 (0.0394)
Child is a boy: π_3	0.1194 (0.0868)
Child is age 8–15 and has zero years of schooling: π_4	1.623 (0.2586)
Child's grade >= 7	
Type 1: $\pi_{5,1}$	−0.1431 (1.487)
Type 2: $\pi_{5,2}$	0.7404 (0.2388)
Type 3: $\pi_{5,3}$	−0.6702 (1.158)
Child's grade >= 7 and child is a boy: π_6	−0.5201 (0.2836)

V. Type probabilities: Couples married 8 years or more

Variable	Std. error	Variable	Std. error
Constant		Number of kids 14–15	
Type 1: $\beta^2_{0,1}$	2.529 (1.753)	Type 1: $\beta^2_{6,1}$	−0.0041 (0.5219)
Type 2: $\beta^2_{0,2}$	1.765 (1.616)	Type 2: $\beta^2_{6,2}$	−0.1598 (0.5171)
Mother's age at marriage		Number of children 12–15	
Type 1: $\beta^2_{1,1}$	−0.0405 (0.0717)	Type 1: $\beta^2_{7,1}$	−0.1398 (0.3991)
Type 2: $\beta^2_{1,2}$	0.0424 (0.0591)	Type 2: $\beta^2_{7,2}$	−0.3188 (0.3993)
Father's age in 1997		Number of children as of 1997	
Type 1: $\beta^2_{2,1}$	0.0070 (0.0185)	Type 1: $\beta^2_{8,1}$	0.1030 (0.1577)
Type 2: $\beta^2_{2,2}$	−0.0086 (0.0179)	Type 2: $\beta^2_{8,2}$	−0.3497 (0.1356)
Distance of village to nearest city		Mother's age in 1997	
Type 1: $\beta^2_{3,1}$	0.0004 (0.0028)	Type 1: $\beta^2_{10,1}$	−0.0290 (0.0458)
Type 2: $\beta^2_{3,2}$	0.0018 (0.0027)	Type 2: $\beta^2_{10,2}$	0.0094 (0.0414)

VI. Type probabilities: Couples married 7 years or less

Variable	Std. error
Constant	
Type 1: $\beta^1_{0,1}$	2.0681 (10.26)
Type 2: $\beta^1_{0,2}$	1.124 (9.307)
Mother's age at marriage	
Type 1: $\beta^1_{1,1}$	−0.0163 (0.365)
Type 2: $\beta^1_{1,2}$	0.0619 (0.332)
Father's age at marriage	
Type 1: $\beta^1_{2,1}$	0.0142 (0.224)
Type 2: $\beta^1_{2,2}$	0.0016 (0.206)
Distance to nearest city	
Type 1: $\beta^1_{3,1}$	0.0020 (0.00845)
Type 2: $\beta^1_{3,2}$	−0.0014 (0.00819)

Variance-covariance matrix $f(\varepsilon) = N(0, \Omega)$

Variable	Std. error
Variance of pregnancy shock	3.274E9 (1.5114E8)
Variance of boy's age 12–15 leisure shock	1.584E6 (5.3911E5)
Variance of girl's age 12–15 leisure shock	1.015E6 (3.5834E5)
Variance of parent's income shock	0.3744 (7.1505E-3)

APPENDIX B1—*Continued.*

V. Type probabilities: Couples married 8 years or more				VI. Type probabilities: Couples married 7 years or less		Variance-covariance matrix $f(\varepsilon) = N(0, \Omega)$	
Variable	Std. error	Variable	Std. error	Variable	Std. error	Variable	Std. error
Distance to secondary school		Number of kids with 7 or more years of education		Distance to secondary school		Variance of child's wage shock	0.5051 (0.1345)
Type 1: $\beta^{2}_{4,1}$	0.00002 (0.0001)	Type 1: $\beta^{2}_{11,1}$	−0.8088 (0.4171)	Type 1: $\beta^{1}_{4,1}$	0.0016 (0.0011)		
Type 2: $\beta^{2}_{4,2}$	−5.684E-5 (9.785E-5)	Type 2: $\beta^{2}_{11,2}$	1.273 (0.3396)	Type 2: $\beta^{1}_{4,2}$	−0.0001 (0.0010)		
Maximum of mother's and father's education >= 9		Family has either 3 or more children age 0–5, or 3 or more age 6–11, or two or more age 12–15		Maximum of mother's and father's education >= 9		Child wage measurement error std. deviation	0.5402 (0.0536)
Type 1: $\beta^{2}_{5,1}$	−0.1347 (1.2250)	Type 1: $\beta^{2}_{12,1}$	0.1212 (0.4641)	Type 1: $\beta^{1}_{5,1}$	−0.3498 (2.618)		
Type 2: $\beta^{2}_{5,2}$	2.0033 (0.8521)	Type 2: $\beta^{2}_{12,2}$	−0.0594 (0.4238)	Type 2: $\beta^{1}_{5,2}$	1.0670 (2.356)		
Average schooling of children as of 1997							
Type 1: $\beta^{2}_{9,1}$	−0.0246 (0.1447)						
Type 2: $\beta^{2}_{9,2}$	0.0713 (0.1321)						

REFERENCES

Attanasio, Orazio, Costas Meghir, and Ana Santiago. 2001. "Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate Progresa." Unpublished.

Becker, Gary S., and H. Gregg Lewis. 1973. "On the Interaction between the Quantity and Quality of Children." *Journal of Political Economy*, 81(2): S279–88.

Behrman, Jere R. 1997. "Intrahousehold Distribution and the Family." In *Handbook of Population and Family Economics.* Vol. 1A, ed. Mark R. Rosenzweig, and Oded Stark, 125–87. Amsterdam: Elsevier Science, North-Holland.

Behrman, Jere R., Robert A. Pollak, and Paul Taubman. 1986. "Do Parents Favor Boys?" *International Economic Review*, 27(1): 33–54.

Behrman, Jere R., Piyali Sengupta, and Petra Todd. 2005. "Progressing through PROGRESA: An Impact Assessment of a School Subsidy Experiment in Rural Mexico." *Economic Development and Cultural Change*, 54(1): 237–75.

Behrman, Jere R., and Petra Todd. 2000. "Randomness in the Experimental Samples of PROGRESA (Education, Health, and Nutrition Program)." Washington, DC: International Food Policy Research Institute.

Eckstein, Zvi, and Kenneth I. Wolpin. 1999. "Why Youths Drop Out of High School: The Impact of Preferences, Opportunities, and Abilities." *Econometrica*, 67(6):1295–1339.

Gomez de Leon, José, and Susan Parker. 2000. "The Impact of Anti-Poverty Programs on Children's Time Use: The Case of Progresa in Mexico." Washington, DC: International Food Policy Research Institute.

Heckman, James J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica*, 47(1): 153–61.

Heckman, James 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,* ed. Charles F. Manski and Daniel McFadden, 179–95. Cambridge, MA: MIT Press.

Heckman, James J. 2000. "Causal Parameters and Policy Analysis in Economics: A Twentieth Century Retrospective." *Quarterly Journal of Economics*, 115(1): 45–97.

Holtz, V. Joseph, Jacob A. Klerman, and Robert J. Willis. 1997. "The Economics of Fertility in Developed Countries." In *Handbook of Population and Family Economics.* Vol. IA, ed. Mark R. Rosenzweig and Oded Stark, 275–347. Amsterdam: Elsevier Science, North-Holland.

Holtz, V. Joseph, and Robert A. Miller. 1993. "Conditional Choice Probability and the Estimation of Dynamic Models." *Review of Economic Studies*, 60(3): 497–529.

Keane, Michael P., and Kenneth I. Wolpin. 1994. "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence." *Review of Economics and Statistics*, 76(4): 648–72.

Keane, Michael P., and Kenneth I. Wolpin. 1997. "The Career Decisions of Young Men." *Journal of Political Economy*, 105(3): 473–522.

Keane, Michael P., and Kenneth I. Wolpin. 2000. "Eliminating Race Differences in School Attainment and Labor Market Success." *Journal of Labor Economics*, 18(4): 614–52.

Keane, Michael P., and Kenneth I. Wolpin. 2001. "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment." *International Economic Review*, 42(4): 1051–1103.

Lise, Jeremy, Shannon Seitz, and Jeffrey Smith. 2003. "Equilibrium Policy Experiments and the Evaluation of Social Programs." Queen's University Working Paper 1012.

Lumsdaine, Robin L., James H. Stock, and David A. Wise. 1994. "Pension Plan Provisions and Retirement: Men and Women, Medicare, and Models." In *Studies in the Economics of Aging,* ed. David A. Wise, 183–212. Chicago: University of Chicago Press.

Marschak, Jacob. 1953. "Economic Measurements for Policy and Prediction." In *Studies in Econometric Method,* ed. William Hood and Tjalling Koopmans, 1–26. New York: John Wiley & Sons, Inc.

McFadden, Daniel, and Antti P. Talvitie and Associates. 1997. "Validation of Disaggregate Travel Demand Models: Some Tests." In *Urban Demand Forecasting Project, Final Report: Volume V: Demand Model Estimation and Validation.* Part II, Chapter 3. Berkeley: Institute of Transportation Studies, University of California, Berkeley.

McFadden, Daniel. 1989. "A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration." *Econometrica*, 57(5): 995–1026.

Parker, Susan, and Emmanuel Skoufias. 2000. "The Impact of PROGRESA on Work, Leisure and Time Allocation: Final Report." Washington, DC: International Food Policy Research Institute.

Schultz, T. Paul. 2000. "School Subsidies for the Poor: Evaluating a Mexican Strategy for Reducing Poverty." Unpublished.

Willis, Robert J. 1973. "A New Approach to the Economic Theory of Fertility Behavior." *Journal of Political Economy*, 81(2): S14–64.

Wise, David A. 1985. "Behavioral Model versus Experimentation: The Effects of Housing Subsidies on Rent." In *Methods of Operations Research 50*, ed. Peter Brucker and R. Pauly, 441–89. Konigstein: Verlag Anton Hain.

Wolpin, Kenneth I. 1984. "An Estimable Dynamic Stochastic Model of Fertility and Child Mortality." *Journal of Political Economy*, 92(5): 852–74.