

# Daddies, Devotion, and Dollars

## *How Do They Matter for Youth?*

By GARY PAINTER and DAVID I. LEVINE\*

**ABSTRACT.** Growing up in a family that lacks a biological father is correlated with a number of poor outcomes for youths. This study uses the National Educational Longitudinal Survey of 1988 to examine the extent to which differences in income or parental involvement can explain the effects of family structure on youth outcomes. We find that measurement error in income from single-parent homes has a large effect on the results because of the variability in income earned over a youth's teen years. Overall, we find that lower income explains most of the disadvantages of youths in single-parent homes, but neither gaps in income nor in parental involvement explain the disadvantages of families with stepfathers.

\*Gary Painter is an Associate Professor at the School of Policy, Planning, and Development, University of Southern California, Los Angeles, CA 90089-0626 <gpainter@usc.edu>. His interests include the importance of family background and neighborhoods on education outcomes, as well as the importance of race and housing policy in determining housing outcomes. He has recently published in the *Journal of Human Resources*, the *Journal of Urban Economics*, the *Journal of Housing Economics*, and *Industrial Relations*. David I. Levine is a Professor at the Haas School of Business, University of California, Berkeley, CA 94720-1900 <levine@haas.berkeley.edu>. Levine's research focuses on labor markets and workplaces, the determination of wages and inequality in the United States, and the causes and consequences of child poverty. He has recently published books entitled *The American Workplace: Skills, Pay and Employee Involvement* (Cambridge University Press, 1999) and *Working in the Twenty-First Century: Government Policies to Promote Opportunity, Learning and Productivity in the New Economy* (M. E. Sharpe, 1998). We appreciate useful comments from Timothy Biblarz. All programs used in this project are available from the first author. Data are available from the National Center for Education Statistics.

The American Journal of Economics and Sociology, Vol. 63, No. 4 (October, 2004).  
© 2004 American Journal of Economics and Sociology, Inc.

## I

**Introduction**

THE NONTRADITIONAL FAMILY is becoming the norm. For example, in 1988 approximately 35% of eighth graders lived in homes that were not headed by two biological parents (see Table 1), up from 25% in 1970 (*Statistical Abstract of the United States* 1996). What is even more striking is that less than 40% of African-American children have two biological parents in the household.

Children who do not grow up living with both biological parents are more likely to drop out of high school, have children out of wedlock, and be arrested. Such children are also less likely to attend college or hold a good job. While the mechanisms underlying these disadvantages remain subject to research, an understanding of them is critical in developing public policies that can help children with a higher risk of negative outcomes.

There are many theories that predict why, on average, children are disadvantaged in nontraditional families. For example, children in single-parent homes live in families with lower average incomes, are more likely to suffer school and/or residential dislocation, and suffer more sexual abuse (Acock and Demo 1994; Amato, Loomis, and Booth 1995; Biblarz, Raftery, and Bucur 1997; Coleman 1994; Forehand, Long, and Brody 1988; Garasky 1995; McLanahan 1985; McLanahan and Sandefeur 1994; Simons 1996). The quality of parenting may also suffer, where quality of parenting is defined as time spent with children and the level of parental involvement in their children's education and in their children's other activities (Acock and Demo 1994; Simons 1996; Coleman 1994; Downey 1995). One can infer from these studies that one reason nontraditional families fail is due to a reduction in "quality time" spent with their children.

This study uses the National Education Longitudinal Study of 1988 to examine the importance of two possible causal channels by which children in nontraditional families are disadvantaged: low parental income and lack of time or resources to devote to helping children by activities such as volunteering in schools and participating in extracurricular activities with them (parental involvement). The NELS is sponsored by the National Center for Education Statistics and carried

Table 1  
Family Structure by Ethnic Group

Family Structure of Children in the Eighth Grade	Ethnic Heritage				Total
	Asian	Latino	African	Caucasian	
Both biological parents	78.02%	62.43%	36.35%	68.55%	65.24%
Father, stepmother	0.99%	1.69%	0.72%	2.12%	1.84%
Mother, stepfather	3.56%	6.67%	8.48%	9.06%	8.30%
Single female	4.16%	13.67%	37.43%	11.63%	13.93%
Single male	2.38%	1.30%	0.93%	1.74%	1.65%
No biological parent	10.89%	14.24%	16.09%	6.90%	9.03%

Source: National Center for Education Statistics (1994). *National Education Longitudinal Study of 1988, Second Follow-Up: Student Component Data File User's Manual*, Washington DC: June.

out by the Bureau of the Census. The NELS is designed to provide trend data about critical transitions experienced by young people as they develop, attend school, and embark on their careers. The base year (1988) survey was a multifaceted study with questionnaires for students, teachers, parents, and schools. Due to the multiple questionnaires, these data are advantageous because they contain rich measures of the characteristics of youth, their schools, and their parents.

These features allow us to control for many aspects of family background that the hypotheses mentioned above suggest are important. Therefore, this research is subject to little of the biases due to omission of key variables that other studies have been forced to work around due to data constraints. We have multiple measures of several of the key constructs, which reduces problems of measurement error. We also examine multiple outcomes, which permits us to identify how income and parental involvement may affect education attainment differently than out-of-wedlock fertility.

Most past research has used statistical techniques that assume that predetermined variables such as race and maternal education do not affect youth, in part due to their effects on family structure (e.g., Astone and McLanahan 1991; Downey 1995). This strong assumption can hide the true effects of income and parental involvement. Similarly, most past research has used statistical techniques that ignore how income can partly proxy for unobserved aspects of the family. This strong assumption can overstate the effects of income as a channel for why family structure matters. We examine how relaxing these assumptions changes the results.

The study of how family structure affects youth outcomes is complicated by the fact that family structure may be correlated with poor outcomes for youth, but not be causally related. This would occur if the parents of children living in single-parent households were disadvantaged prior to becoming parents or single parents, so that their lack of marriage or continued marriage is not the cause of their children's disadvantage (Manski et al. 1992; Painter and Levine 2000). Understanding causality is critical for policy purposes, because, for example, promoting marriage and discouraging divorce may not help children if the observed disadvantages in single-parent households are not caused by family structure.

Research is not conclusive as to the importance of the noncausal channels. Cherlin et al. (1991) examined the lives of children at ages 7 (in the United Kingdom) or 7 to 11 (in the United States), and resurveyed the same children roughly four years later, after their parents divorced. They found the noncausal channel to be important in predicting behavior and achievement deficiencies in boys, but less important in girls. In contrast, in a companion paper (Painter and Levine 2000), we found that little of the disadvantage of divorce during a teen's high school years was due to predivorce disadvantages of the family. Morrison and Cherlin (1995) also found little effect of the noncausal channel.

Although we examine only two of the many causal channels, they have important policy implications. If the disadvantages are largely due to the lower incomes of single-parent households, reducing welfare, without strengthening other income support programs such as the earned income tax credit, may be very expensive for the next generation. Conversely, if the disadvantages are largely due to different levels of parental involvement, finding ways to promote parental involvement becomes a target of policy.

We find that both income and parental involvement play important roles in explaining the disadvantages of youth who grow up with a single mother. This study finds a much larger role for income than previous studies because we partially control for measurement error in income, which is likely to be a problem in cross-sectional data. Income itself eliminates most of the estimated negative impact of growing up in a female-headed household when predicting educational attainment and out-of-wedlock fertility. Parental involvement is important for explaining educational attainment, but not for out-of-wedlock teen fertility. On the other hand, neither income nor parental involvement can explain the observed disadvantage of youth who grow up in a family with a stepfather.

## II

### **Theory and Methods**

THE LITERATURE ON the importance of income as a determinant of youth outcomes is vast. Youth outcomes are usually either educational or behavioral. (In this study, we will focus on the likelihood that

someone will drop out of high school, attend college, or (for females) have a child out of wedlock.) Many studies find that the reduction in income associated with a marital disruption is an important contributing factor to the observed disadvantaged outcome in the youth. For example, Sandefur, McLanahan, and Wojtkiewicz (1992) find that the reduction in income accounts for 15% of the observed gap in high school graduation rates between children living with one versus two parents. Downey (1995) finds that most of the disadvantage of growing up in a stepfather household instead of an intact family is due to parental education, income, and race.

Mayer (1997) and others have noted that income can improve outcomes through many channels, including higher consumption levels of a family, improved access to enrichment activities such as tutoring, after-school classes, or camp, improved parental well-being, and raising youths' perceptions that the family can afford college. Conversely, poverty may limit the ability of the family to provide certain amenities, such as visiting museums. While Mayer (1997) cautions that income alone cannot determine consumption choices, a higher income can free up more time. A low income can also increase financial stress in the household. Financial stress, in turn, can change the parenting practices either by decreasing the time available to spend with children or by creating an atmosphere of inferior discipline.

Researchers have also examined the effect of parenting practices on youth. Astone and McLanahan (1991) note that parental aspirations and supervision are important predictors of youth high school graduation. In a sample of eighth graders, Downey (1995) finds that children whose parents attend the local PTA and whose parents know their children's friends received higher grades and higher test scores than children whose parents do not. McLanahan and Sandefur (1994) mention that one reason that children in stepfamilies seem to do nearly as poorly as do children in female-headed households is inferior parenting practices when compared to the intact family. They suggest that while the stepfather brings in additional income to the family, constructing a new, satisfying marital relationship commands a great deal of the mother's time and attention. This lack of time is also present with regard to female-headed households, due to financial demands and lack of the additional parent.

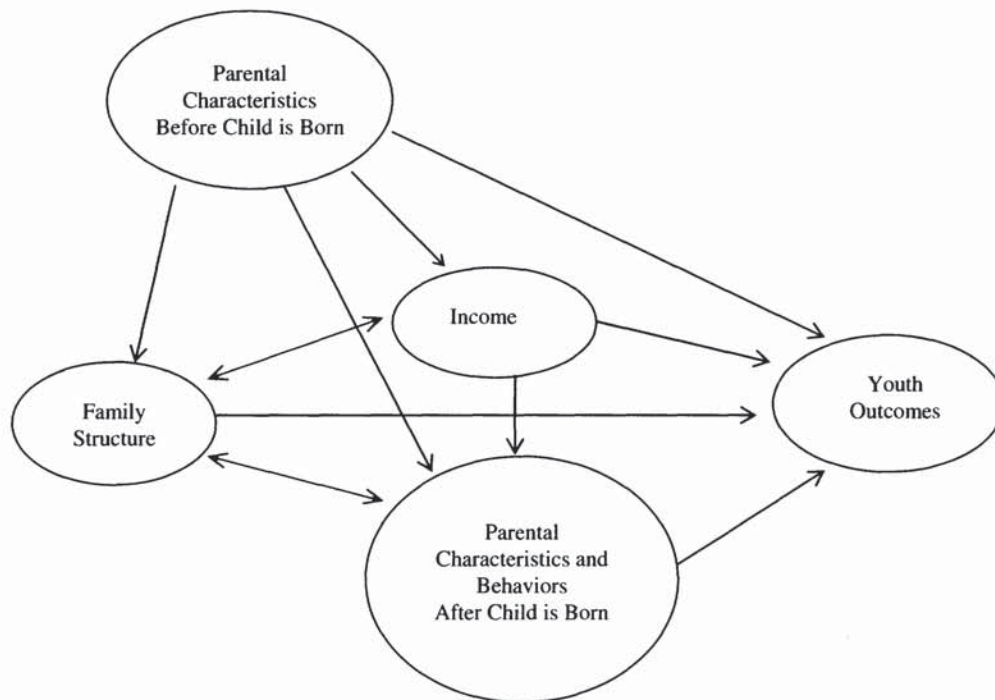
*The “Conventional Reduced Form”:* A general model of the determinants of youth outcomes posits that they depend on family structure, predetermined parental characteristics that predate the birth of the child, and contemporaneous characteristics of the family, parent, and youth ranging from parental occupation and employment status to whether the family has a library card:

Youth outcome

$$= \delta_0 + \delta_1 \text{ family structure} + \delta_2 \text{ predetermined characteristics} + \delta_3 \text{ income} + \delta_4 \text{ contemporaneous characteristics} + u_1. \quad (1)$$

As Figure 1 illustrates, the estimation of such a model is not straightforward. Income can affect other contemporaneous characteristics; for example, higher income can enable more trips to museums. Causation can run from contemporaneous characteristics to income; for example, occupation is a good measure of permanent income, even when controlling for a single year's income. To understand the true

Figure 1



role of income, we would like to be able to identify its effects on contemporaneous family characteristics. For some purposes, as noted below, we would also like to subtract how predetermined characteristics and omitted variables affect income. Further, if increasing income affects outcomes, then we want to increase income and ignore the parental characteristics that caused it. Given that no plausible instruments to distinguish the causation exist in the data set, we have to estimate a reduced form.

A number of researchers have estimated what Mayer (1997) referred to as the “conventional reduced form” to study the role of income or parental involvement in explaining how family structures affect youth outcomes (e.g., Astone and McLanahan 1991; Downey 1995). In these specifications, the authors typically divided variables into three groups: predetermined variables assumed to causally precede family structure; family structure; and a focal input such as income. They then ran:

$$\begin{aligned} &\text{Youth outcome} \\ &= A_1 \text{ predetermined characteristics} + B_1 \text{ family structure} \quad (2) \end{aligned}$$

and

$$\begin{aligned} &\text{Youth outcome} \\ &= A_2 \text{ predetermined characteristics} + B_2 \text{ family structure} \\ &\quad + \Gamma_2 \text{ focal input such as income.} \quad (3) \end{aligned}$$

These studies then focus on the decline in the estimated coefficient on family structure ( $B_1 - B_2$ ) as income is added to the regression. Implicitly, this specification assumes that the predetermined variables do not operate by changing income. These studies typically find that including income reduces the estimated negative effects of nontraditional family structures by about one-third.

*What If Predetermined Variables Operate Via Income?* The conventional reduced form may understate the true role of income because the predetermined variables such as race or maternal education affect youth in part by affecting income. Consider a world (similar to our own) in which low-educated mothers are more likely to be single parents and also to live in families that, on average, have below-average incomes.

When the regression does not include maternal education, the coefficient on income is biased up because it picks up some of the nonincome advantages associated with high maternal education, such as being raised in a home with more books. Moreover, the effect of income in reducing the coefficient on family structure will partly capture some of the true effect of higher maternal education. Thus, the effect of income in reducing the coefficient on family structure will be an upper bound.

Algebraically, if we regress

$$\text{Youth outcome} = B_1 \text{ family structure} \quad (2')$$

and

$$\begin{aligned} \text{Youth outcome} = B_2 \text{ family structure} + \Gamma_2 \\ \text{focal input such as income} \end{aligned} \quad (3')$$

the difference  $(B_1 - B_2)$  will be biased up and will be larger than  $(B_1 - B_2)$  estimated in the conventional reduced form (Equation (3)).

Conversely, the conventional reduced form includes maternal education in the regression without income (as in Equation (2)). In this formulation, maternal education picks up some of the effect of the omitted variable income. Thus, the estimated effects of family structure already have factored in the disadvantage due to lower income that was, in turn, correlated with low maternal education. That is, some fraction of the disadvantage of lower-educated parents is due to their lower income. Thus, when income is added to the regression, its effect is lower than the true effect. Thus the conventional reduced form's estimate of how income accounts for the disadvantage of some family structures  $(B_1 - B_2)$  is biased down. This result is modeled formally in Appendix 1.

*What If Income Proxies for Unmeasured Characteristics of the Family?* Conversely, income partly proxies for unmeasured characteristics of the parents and family that benefit the child (Mayer 1997). Consider the case where highly energetic parents earn more and also take their children to the library more. In this setting, the correlation between income and library use is not causal; instead, library use proxies for unobserved parental energy.

For concreteness, we can make the extreme assumption that all of

the contemporaneous characteristics that are correlated with income causally affect youth due to common omitted factors that lead to both income and to these characteristics. In this case, the correct method is to add income into the regression after all contemporaneous measures of the family's characteristics have already been entered. With accurate measures of income, this procedure leads to a lower bound on the role of income.

*Measurement Error on Income:* In fact, income is measured with error. Thus, the tests described above may understate the true effect of income. In addition, a single year's income is an imperfect measure of a family's permanent income, and it is permanent income that probably matters more for youth (Solon 1992; Zimmerman 1992). Transitory movements or measurement error are particularly large for single-parent households; in the NELS, such families have 44% more variability in income over the study period than do stepfamilies, and 56% more variability than intact families. In a standard setting, random measurement error will reduce the estimated coefficient of income on youth outcomes. Moreover, random measurement error will also reduce the effect of income in lowering the estimated effects of family structure in predicting youth outcomes.

Fortunately, the data set has a measure of income four years after the data set began. We are able to use this measure to develop a measure of permanent income, using a standard instrumental variables technique. We use the income during a youth's 12<sup>th</sup>-grade year (1992) as an instrument for income during the youth's eighth-grade year (1988) to estimate the first-stage equation:<sup>1</sup>

$$\begin{aligned} \text{Income}_{88} = & \theta_0 + \theta_1 \text{Income}_{92} + \theta_2 \text{predetermined characteristics} \\ & + \theta_3 \text{contemporaneous characteristics} + u_2. \end{aligned} \quad (4)$$

We then substitute the predicted value  $\text{Income}_{88}^{\wedge}$  into Equation (1) for  $\text{Income}_{88}$  and reestimate Equation (1). This yields the equation:

$$\begin{aligned} \text{Youth outcome} \\ = & \delta_{0a} + \delta_{1a} \text{family structure} + \delta_{2a} \text{predetermined characteristics} \\ & + \delta_{3a} \text{Income}_{88}^{\wedge} + \delta_{4a} \text{contemporaneous characteristics} + u_{1a}. \end{aligned} \quad (1a)$$

The measurement error on  $Income_{88}$  and on  $Income_{92}$  are correlated; for example, neither pick up income from before 1988 and both are subject to bias from a family that systematically over- or underreports income. If the measurement error is correlated in this fashion, the coefficient from the instrumental variable estimate remains somewhat biased down.

*Parental Involvement:* We repeat the above procedures to study how parental involvement mediates the effects of family structure on youth outcomes. We do not create an instrumental variables approach for parental involvement. The data do not contain later measures of parental involvement, and other possible instruments are not likely to be uncorrelated with the error term in Equation (1). Therefore, we are not willing to make the strong assumptions that this method requires. Instead, we add multiple measures of parental involvement. To the extent that results are similar when new measures are added, we are reassured that better measures will not greatly change the results.

### III

#### Data

AS NOTED ABOVE, the National Education Longitudinal Study of 1988 (NELS) is sponsored by the National Center for Education Statistics and carried out by the Bureau of the Census.<sup>2</sup> Sampling was first conducted at the school level and then at the student level within schools. The data were drawn from a nationally representative sample of 1,000 schools (800 public schools and 200 private schools, including parochial institutions). Within this school sample, 25,000 eighth-grade students were selected at random. The three follow-ups revisited (most of) the same sample of students in 1990, 1992, and 1994; that is, when the respondents were typically in the 10<sup>th</sup> grade, in the 12<sup>th</sup> grade, and roughly two years after high school graduation. A randomized sample of approximately 14,000 students was interviewed in the 1994 survey. We restrict our sample ( $N = 9,260$ ) to the three family structures described below, and drop observations for which there was incomplete information available on the income of the family ( $N = 2,391$ ).<sup>3</sup>

*Family Structure:* Due to limited sample size, we focus on three family structures: intact families with both biological parents, families with a mother and a stepfather, and families with a mother alone.<sup>4</sup> We examined several other family structures—father and stepmother ( $N = 216$ ), father alone ( $N = 178$ ), mother and live-in companion ( $N = 123$ ), and no biological parents ( $N = 254$ )—but the sample sizes were either small or, in the case of children living without a biological parent, it was difficult to characterize the involvement of the parents, if any. These family structures are subsequently dropped from the main results. Respondents were also dropped if they lived with their parents less than 50% of the time ( $N = 40$ ), if a spouse died while the youth was in high school ( $N = 88$ ), or if the parent's and youth's surveys had conflicting reports on family structure ( $N = 27$ ).

Finally, we dropped those families that experienced transitions during the high school years of the youth: divorce from an intact family ( $N = 468$ ), (re)marriage by a single mother ( $N = 189$ ), and divorce from a stepfather ( $N = 146$ ). The inclusion of these families would preclude the instrumental variable technique that uses both eighth-grade and 12<sup>th</sup>-grade family income. The exclusion of these families would bias our results if the families that experienced transitions were systematically different than families that did not. Painter and Levine (2000) suggest in a study that uses the same sample of eighth graders that families which experience transitions are not systematically different from those that did not; therefore this exclusion should not affect the results.

Our measures of family structure accurately describe intact families and the current status of nonintact families, but they do not indicate how many times a particular woman may have been divorced or remarried. Garasky (1995) finds later transitions are more important than early ones, so the bias from not having data on these earlier transitions may be lessened.

*Socioeconomic Status and Family Background:* A missing ingredient in most analyses of the impact of family structure on the achievement of youths is adequate measures of family background and parental involvement in education. Studies have either used a socioeconomic status index provided by the data set (e.g., Lee et al. 1994),

created an *ad hoc* index of parent's characteristics (e.g., Herrnstein and Murray 1994), or used a limited set of family background measures intended to separate the effects of family structure on the achievement of youths from the effects of family background. This study employs much more detailed measures of family background and family involvement in education, which are intended to better isolate the effect of family structure on outcomes.

The measures of socioeconomic status are created from both the parent and student questionnaire. The set of variables include occupational status (using Duncan's index), parental education, and family income. Occupational status is converted into z-scores with mean 0 and standard deviation equal to 1. When mother's education is missing, it is set to a z-score of 0, and a categorical variable is included to note these important missing values. We are not able to do this in the case of father's education, because so many are missing values in female-headed homes. Thus, for father's education, we impute missing values based on the other predetermined characteristics of the family.<sup>5</sup>

Family income is comprised of income from all individuals and sources. Family income is presented in categories, and is top-coded at \$200,000 (less than 2% of the sample). It is assumed that income is at the midpoint of each of the categories, and for those households above \$200,000, an income of \$300,000 is assumed. To adjust family income for its size, family income is divided by the poverty line adjusted for family size. This is an improvement over most studies, which simply include some measure of family income in their estimated models. The log of this income:needs ratio (hereafter referred to as income:needs or income) is available for both the student's eighth-grade and 12<sup>th</sup>-grade years (or what would have been the student's 12<sup>th</sup>-grade year if he or she had previously dropped out of school).<sup>6</sup>

To create a rich list of family characteristics, we included a wide range of measures that prior research suggests are indicators of advantages or disadvantages for youth. From the student questionnaire, we use standard demographic characteristics: region, rural versus urban versus suburban, race categorical variables, and a female categorical variable. A second set of variables is indirectly related to parental

involvement in education, but is not exogenous to the outcome variable. These include whether a foreign language is spoken in the home, whether the mother or father is foreign born, the number of siblings, and whether the home has a library card, magazines, and many books.

From the parental questionnaire, indicators are obtained for whether the family was one of five religions, and any of four levels of religious observance. These variables may proxy for how closely a family is knit as well as proxy for the social capital (Coleman 1990) available to the children. Also, a categorical variable indicating if the mother was a teen when the youth was born is included. Unfortunately, the data set does not indicate whether the parents were married when the youth was born.

The final two variables measure parents' involvement in the youth's life and education. We refer to these variables as "parental involvement" throughout the paper. The first variable is equal to 1 if the parent belonged to a parent-teacher association or related organization, or volunteered at school. Finally, a categorical variable for whether the child had participated in clubs such as Boy or Girl Scouts, youth sports, or religious or community groups during elementary school is included to proxy for the quantity of time spent with the child outside of the home. While the data do not distinguish whether the parent directly participated in these groups, child participation would necessitate some parent role in most cases. (Recall that some currently female-headed households were married when the focal youth was in elementary school, the time period when the scouting question applied. Thus, those families may have higher levels of facilitating the participation of their children in extracurricular activities in the past than in the present.)

*Outcomes:* This study analyzes three outcomes that are observed when the youths were age 20. While the overall impact of family background and family structure is similar across the various outcome measures, there are subtle differences that may be important. The outcomes include permanently dropping out of high school (that is, dropouts who do not receive a GED), attending college if he or she has received a high school diploma, and having a child out of wedlock. The presentation below focuses primarily on the influence

on dropping out of high school, and notes any differences in the outcome measures in a robustness checks section.

Summary statistics for the analysis variables are presented in Table 2. The means for the outcome variables are taken from the estimation sample, while the means for the remainder of the variables are taken from the complete sample. For example, the estimation sample for having a child out of wedlock is made up of females, and the estimation sample for analyzing college attendance includes only high school graduates. Approximately 5% of the sample permanently dropped out of high school, while 75% of high school graduates had attended some college. Nine and a half percent of the females had a child out of wedlock.

#### IV

#### **Results**

SINCE THE OUTCOMES WE STUDY are binary, we estimate a probit specification. The tables present only the coefficients on family structure and the coefficients on the investigated causal channels. The coefficient estimates are converted into marginal changes in the probability of the outcome evaluated at the mean of the independent variables. We first present a detailed analysis of results for permanent dropouts from high school. We then present results on starting college and (for women) teen out-of-wedlock childbearing more briefly. Tables present the pseudo- $R^2$ , which is defined as  $1 - L_1/L_0$ , where  $L_1$  is the log likelihood of the estimated model and  $L_0$  is the log likelihood corresponding to the model with only a constant term.

The results of the equations predicting dropouts are presented in Table 3. Model 1A demonstrates that female-headed households have drop-out rates 6.3 percentage points higher and that stepfather families have rates 5.0 percentage point higher than intact families. Controlling for the predetermined factors (maternal and paternal education, race, and region of residence) cuts both the stepfather effect and the female-headed effect roughly by a third to a bit over 4 percentage points (Model 1B). With a mean drop-out rate in the population of 5% (Table 2), the effects of being raised in a nontraditional family remain large.

Table 2  
Summary Statistics  
(*N* = 9,260)

Analysis Variables	Mean	Std Dev.
<b>Family Structures</b>		
Female-headed family throughout the sample	0.146	
Stepparent family throughout the sample	0.083	
<b>Family Characteristics</b>		
Log (income/needs) in 8th grade	1.037	1.017
Log (income/needs) in 12th grade	1.022	1.151
Parent involved in educational system (e.g., PTA member)	0.570	
Parent involved in children's clubs (e.g., Boy or Girl Scouts)	0.889	
<b>Youth Outcomes</b>		
Permanent dropout (that is, no GED)	0.054	
College attender (among HS graduates)	0.747	
Had a child out of wedlock (among women)	0.100	
<b>Included as Predetermined Characteristics:</b>		
African American (omitted category is Caucasian)	0.094	
Asian	0.073	
Latino	0.107	
Female	0.520	
Native English speaker	0.836	
Father foreign born	0.154	
Mother foreign born	0.154	
Live in the south (omitted category is northeast)	0.347	
Live in the west	0.180	
Live in the central	0.296	
Live in urban area (omitted category is suburb)	0.242	
Live in rural area	0.321	
Oldest child	0.323	
Mother was a teen parent	0.096	

Table 2 *Continued*

Analysis Variables	Mean	Std Dev.
Father's education {z}	0.031	0.962
Mother's education {z}	0.087	0.989
<b>Included as Contemporaneous Characteristics (in Addition to Predetermined Characteristics):</b>		
Father's occupation {z}	0.055	0.904
Father unemployed	0.052	
Mother's occupation {z}	0.031	0.957
Mother unemployed	0.288	
Religious affiliation—Baptist (missing is other Protestant)	0.190	
Religious affiliation—Catholic	0.328	
Religious affiliation—other religion	0.116	
Religious affiliation—missing religion	0.030	
Religious affiliation—no religion	0.026	
Religiosity—very religious	0.435	
Religiosity—religious	0.152	
Religiosity—somewhat religious	0.154	
Number of siblings	2.152	1.498
More than 50 books in home	0.902	
Has at least one magazine subscription	0.778	
Family has a public library card	0.823	

*Note:* Variables marked {z} are z-scored to have mean 0 and s.d. 1 in the entire sample. Reported summary statistics differ from 0 and 1 due to exclusion of families with incomplete income reports (and those that divorced or remarried). Variables with no standard deviation indicated are dummy variables.

Model 2B presents the typical measure of the effects of income found in past research. Specifically, income is added to Model 1B, and we examine the decline in the coefficients on family structure. Youth in families with higher income in eighth grade had significantly lower odds of dropping out of high school. Consistent with past research, income lowers the estimated negative impact from residing

Table 3  
What are the Effects of Income and Parental Involvement on Permanently Dropping Out of High School Without a GED?

Probit Estimation ( <i>N</i> = 8,923)	Model 1: Baseline			Model 2: Adding Income			Model 3: Replacing Income with Permanent Income			Model 4: Adding Parental Involvement		
	DF/DX	Std. Error		DF/DX	Std. Error		DF/DX	Std. Error		DF/DX	Std. Error	
<b>A: No controls</b>												
Income:needs in 8 <sup>th</sup> grade				-0.018**	0.002		-0.034**	0.003		-0.027**	0.003	
Parent involved in educational system										-0.026**	0.004	
Parent involved in children's clubs										-0.054**	0.009	
Female headed	0.063**	0.008		0.034**	0.007		0.013**	0.006		0.013	0.006	
Stepfather	0.050**	0.008		0.046**	0.008		0.043**	0.008		0.042**	0.008	
Pseudo- <i>R</i> <sup>2</sup>	0.030			0.071			0.085			0.124		
<b>B: Predetermined characteristics</b>												
Income:needs in 8 <sup>th</sup> grade				-0.005**	0.002		-0.017**	0.003		-0.016**	0.003	



in a female-headed household about 16% (although the decline from 4.4 to 3.7 percentage points is not statistically significant). In contrast, income plays very little role in closing the gap in drop-out rates between stepfather and intact families. The complete results of Models 1B and 2B are provided in Appendix 2; further results are available from the authors.

*What If Predetermined Variables Operate via Income?* The conventional reduced form (comparing Models 1B and 2B) understates the true role of income to the extent that the predetermined variables operate by raising income. If all of the effects of race, maternal education, and other predetermined variables that are correlated with income affect youth outcomes via their effects on income, the correct test is to compare the family effects with no additional controls. In Table 3, adding income to the regression reduces the coefficients on family structure by more when it was entered into the regression without predetermined variables than when it was entered after the predetermined variables (Models 1B vs. 2B). Specifically, controlling for income alone reduces the effect of residing in a female-headed family from a 6.3 percentage points higher drop-out rate to 3.4 percentage points—a 46% reduction. The comparable impact of residing in a stepfather family falls by less (5.0 to 4.6 percentage points), and the difference is not statistically significant.

*What If Income Proxies for Other Characteristics of the Family?* Income is partly a proxy for measured and unmeasured characteristics of the parents and family that benefit the child. In an extreme case, all of contemporaneous (that is, eighth-grade) variables that are correlated with income might causally affect youth directly or due to common omitted factors. The appropriate lower bound of the influence of income is estimated by entering income after all contemporaneous measures of the family's and youth's characteristics.

After these contemporaneous characteristics have been added, the estimated effects of family structure on drop-out rates are much smaller (Model 1C). Now female-headed families have a 3.1 percentage point higher drop-out rate, and stepfather families have only a 2.1 percentage point higher drop-out rate (Model 1C). As expected, income has a smaller impact in lowering the impact of family structure if it is entered into the regression after an extensive set of con-

temporaneous measures of the family's characteristics than if it is entered after only the predetermined variables (Models 1C and 2C); this attenuation is to be expected as income itself is no longer a statistically significant predictor of dropping out. After controlling for income the single-parent effect fell very little (from 3.1 to 2.9 percentage points, difference n.s.) and the stepfather effect did not change at all. Thus, the lower bound estimate of the role of income is quite small.

*The Role of Measurement Error.* The tests of Hypotheses 1, 2, and 3 may understate the true effect of income because income measured in a single year is only a rough measure of permanent income experienced by and expected by a youth. The evidence in Model 3 of Table 3 demonstrates that the effect of income on drop-out status rises by 50% or more when correcting for measurement error. This result is obtained by comparing the coefficient on income in Models 2A, B, and C with the corresponding coefficient on income in Model 3. For example, the coefficient on income rises from  $-0.005$  to  $-0.017$  when using the instrumental variables technique (comparing Models 2B and 3B).

As expected, the impact of income in lowering the estimated effects of family structure on dropping out also rises substantially when correcting for measurement error. If we consider the baseline specification with predetermined controls, adding income reduces the effect of single-parent status by about 16%, and of stepfather status by very little (comparing Models 1B and 2B). When correcting for measurement error, income reduces the effect of single-parent status by two-thirds, but continues to have no influence on the coefficient on stepfather status (comparing Models 1B and 3B).

*The Role of Parental Involvement.* The effects of the parental involvement measures are different than income. As Painter and Levine (2000) show in the NELS, single-parent families have only about half as much income:needs as do intact families, while stepfamilies have only about 15% less income. At the same time, nontraditional families are nearly identical with respect to parental involvement measures, at levels 25% less than intact families. Consider adding the parental involvement measures to the conventional reduced form (Model 1B). The parental involvement measures have

a large influence on dropping out (Model 4B); specifically, youth whose parents were involved in the educational system (e.g., PTA members) had 1.3 percentage point lower drop-out rates, and youth involved in clubs had 2.3 percentage point lower drop-out rates. We also note that their contribution to pseudo- $R^2$  is quite similar to that of income adjusted for measurement error.

At the same time, the addition of the parental involvement variables has little impact on the coefficients on family structure. The coefficient on *stepfather* is lowered slightly in Models 4A and 4B, but in Model 4C, the coefficient slightly increases. This may be due to the fact that these parental involvement variables are closely related to some of the other contemporaneous characteristics, such as having many books in the home or having a library card. In none of the models are the changes in the effects of family structure due to including parental involvement found to be statistically significant.

In results not shown, we also added income after adding parental involvement. The inclusion of income has no impact on the coefficient on parental involvement once other controls are included in the models. Similarly, the inclusion of parental involvement has little influence on the estimated effect of income (comparing Model 3 with Model 4). It appears that parental involvement and income act independently of each other.

As mentioned previously, we do not create an instrumental variables approach for parental involvement. Instead, we add additional measures that capture different dimensions of the construct "parental involvement." Based upon the literature (e.g., Acock and Demo 1994; Simons 1996; Coleman 1994; Downey 1995), we chose the number of friends and the number of friends' parents that the focal youth's parents knew, helping the student with homework, and having rules governing TV use and other behaviors as our supplementary measures of involvement. To the extent that results are similar when new measures are added, we are reassured that better measures will not greatly change the results. We find that additional measures of parental involvement have no impact on the coefficient on family structure (results not shown). This result suggests that measurement error for parental involvement—although surely present—is probably not driving the results.

*Robustness Checks*

*Other Outcome Variables: College Attendance and Out-of-Wedlock Childbearing.* The results for college attendance and out-of-wedlock childbearing differ only slightly from those for drop-out status. Family structure powerfully predicts college attendance for high school graduates. Children from persistently single-parent families are 8.6 percentage points less likely to attend college than are children from intact families, and children from stepfather families are 13.4 percentage points less likely. Controlling for predetermined variables (not shown), the gaps falls to 4.0 and 8.4 percentage points respectively.

We find that accounting for income (adjusted for measurement error) eliminates all of the negative impact of growing up in a female-headed household regardless of the controls included. At the same time, controlling for income has no impact on the disadvantage of growing up in a stepfather family. The contribution of the parental involvement variables was very similar to their contribution in Model 4. They appear to have an independent effect on college attendance distinct from the contribution of income, but they do little in explaining the negative impact of growing up with a stepfather.

Family structure also powerfully predicts childbearing out of wedlock for young women. Young women from persistently female-headed (stepfather) families are 14.3 (11.3) percentage points more likely to have a child out of wedlock than are those from intact families. Controlling for predetermined variables (not shown) lowers the gap to 6.4 and 7.2 percentage points.

Similar to the results on education attainment, income (adjusted for measurement error) fully explains the gap in out-of-wedlock childbearing rates between daughters in female-headed households and daughters in intact families. In contrast to the results on educational attainment, the parental involvement variables are not statistically significant.

*Not All Dollars May be Equal.* The conventional reduced form assumes that the income earned by a stepfather has the same influence on a youth as the income earned by a biological father. This assumption may not always hold. Most obviously, children with

stepfathers have spent some period with only one biological parent. Thus, they have typically spent a period with lower income. To the extent that the income experienced throughout childhood matters, income with a stepfather is an overestimate of permanent income of the youth's family. Second, stepfather families are more likely to divorce than intact families. For example, in the NELS sample, 10% of families with a stepfather when the child was in eighth grade divorced within the next six years, but only 5% of families with two biological parents did so. To the extent that youth consider this possibility, they will treat their expected future income as lower. Stepfathers are also more likely to have children from previous relationships, which reduces their ability to financially support the youth. Finally, some stepfathers have weaker emotional bonds with their stepchildren; thus, they may be less willing (or youth may expect them to be less willing) to pay for college. Any of these forces leads to the hypothesis that income would have a smaller influence in stepfather families than in intact families.

A simple test of the differences in the importance of income can be accomplished by interacting income with family structure. Case, Lin, and McLanahan (1999) proposed a similar test of the importance of family resources in different family structures on the food consumption of children. They find that after controlling for income, the food consumption of the child is statistically similar in stepfather families and the other families in which the biological mother is present.

We find inconsistent evidence of differential impacts of income by family structure. In the models without any controls, we find that income in stepfather families is two percentage points less important than income in other families, and that income in single-parent families is 40% more important. After controlling for predetermined characteristics of the family, the interaction term on residing in a stepfather family becomes insignificant. After controlling for contemporaneous characteristics of the family, both interaction terms are insignificant.

*Gender Differences.* As noted above, previous research suggests that fathers and stepfathers can, among other benefits, provide mentoring and work-related networks, and act as role models. Previous research has also suggested that these benefits may differ by the

gender of the child (Cherlin 1999; Lee et al. 1994; Cherlin et al. 1991; Evenhouse and Reilly 2000). It is likely that the role model and mentoring effects of fathers are strongest for sons and stepsons. Conversely, remarriage also brings increased risk of sexual abuse, especially for stepdaughters (Russell 1984; Gordon 1989). In order to test for the potential differential impact of residing in nontraditional families on male and female youth, we estimate separate equations for each gender and compare the estimated effects on family structure between the two equations.<sup>7</sup>

In this data set, many of the gender differences of the effects of family structures were of the expected direction, but they were also typically small (Table 4). We highlight several important exceptions. For example, as predicted by theory, being in a single-parent family predicted a slightly higher drop-out rate for young men than for young women (8.7% for young men vs. 8.3% for young women). But contrary to theory, having a stepfather also predicted higher drop-out rates for young men (8.1% vs. 6.8%). Neither of these gender differences in means are statistically significant. Including a broad set of controls does not change this conclusion.

For college entrance, single-parent status predicted lower college attendance for young women (74% for men, 72% for women), while stepparent status predicted no difference. Once a broad set of controls is included in the estimated model, there remains no difference in the single-parent coefficient, but the stepfather coefficient is statistically significantly lower for daughters at the 1% level.

In results not shown, we found conflicting evidence about the importance of a stepfather in reducing the incidence of female teens having a child out of wedlock. We found that being in a stepfather family throughout the high school years led to a higher incidence of teen births than residing in a female-headed household (difference n.s.). At the same time, having a female head that remarried during a teen's high school years sharply reduced the rate of teen births (Painter and Levine 2000).

Overall, the evidence appears to weakly support the notion that the presence of either fathers or stepfathers improves the outcomes of male children, but that the presence of a stepfather may reduce outcomes for females. It should be emphasized that, although the

Table 4  
Gender Differences

Outcome	Family Structure	Male		Female	
		Mean	Coef. W/Full Controls	Mean	Coef. W/Full Controls
Permanent Dropout <i>N</i> = 3,359	Persistently intact	0.031		0.028	
	Female headed	0.087	0.021* (0.009)	0.083	0.013* (0.007)
Males & 3,593 Females	Stepfather	0.081	0.009 (0.007)	0.068	0.010 (0.007)
	Persistently intact	0.845		0.795	
<i>N</i> = 3,162	Female headed	0.740	0.003 (0.024)	0.720	-0.039 (0.022)
	Stepfather	0.694	-0.034 (0.027)	0.697	-0.104* (0.029)

findings broadly support previous research, the estimate differences between male and female children were often not significant.

*School Fixed Effects.* The literature on neighborhood effects (e.g., Jencks and Peterson 1991) suggests that important unmeasured characteristics of the neighborhood may significantly influence a youth's outcomes either because of the influence of neighboring families on youth, youth interactions with neighboring youth, or some other factor. To the extent that nontraditional families cluster together in neighborhoods, these omitted neighborhood effects may bias the results on family structure estimated above.

To account for these omitted variables, we include school-level fixed effects. We found that while the fixed effects themselves are important, the impact on the coefficients on family structure was negligible.

## V

### **Conclusions and Policy Implications**

COMPARED WITH THOSE living with two biological parents, children growing up with a single mother have much higher rates of teen childbearing out of wedlock and of dropping out of high school, and much lower rates of entering college. For the outcomes analyzed in this study, the inclusion of income as a control lowers the coefficient on residing in female-headed household by more than two-thirds. The estimated effects of living in a female-headed household are higher if one assumes that predetermined variables such as maternal education and race act in part via their effects on family structure. Moreover, the effects of income in reducing the negative effects of a female-headed household are larger in absolute terms under these assumptions.

Working in the other direction, if one includes contemporaneous controls for the family characteristics in eighth grade, the estimated effects of family structure are lower. For example, income closes the gap for permanent drop-out status by five percentage points under the assumption of predetermined variables operating through family structure, and by only two percentage points if one controls for a broad set of characteristics. Importantly, although the absolute effect

of income is smaller, income closes a similar proportion of the apparent effects of family structure.

We found a larger role for income than do some past studies because we are able to control partially for measurement error in income. This result suggests that a single measure of income in cross-sectional data sets is a poor predictor of permanent income. These findings are consistent with other research on the intergenerational transmission of income inequality (Solon 1992; Zimmerman 1992) and on how child poverty affects youth outcomes (Korenman, Miller, and Sjaastad 1994).

The results on income and stepfather families are simple to summarize: controlling for income never reduced the coefficient on residing in a stepfather family by much. When we stratified the sample by gender, we were unable to find consistent differential effects by gender that might help explain this puzzle. The most important influence on the effect of living with a stepfather was the inclusion of the contemporaneous controls such as the parents' occupations and the religiosity of the family, which lowered the estimated impact in half.

We further find that the estimated effects of parental involvement are almost independent of the effects of income. Parental involvement is important for predicting permanent drop-out status and attending college, but not for out-of-wedlock teen childbearing. At the same time, parental involvement has little role as a mediating variable that explains the disadvantages of female-headed or stepfather families.

As in all such analyses, measurement error and missing variables can continue to bias the results. As such, the importance of family structure on outcomes may still be overstated. The results here suggest that, even in the presence of unobservables, residing in a female-headed household does not independently affect educational outcomes. Therefore, controlling for more factors would only likely reduce the effect that was found in Table 3. At the same time, the reasons that residing in a stepfather family lead to negative outcomes are still unknown. Thus, future research is needed to determine the precise causal mechanism.

This analysis sheds some light on possible policies to address the

disadvantages of youth in nontraditional families. Because the effects of income and parental involvement are almost independent of each other, if policy-makers can identify cost-effective policies to raise parental involvement, they do not need to worry that they will need to also raise incomes sufficiently to enable the higher involvement. Lareau (1989) has noted that the reason that many parents are not involved is that they feel unqualified to help. This is predominately the case for parents with only a high school education. A focus of policy may be to develop programs that teach parents how to assist their children and become more involved in their lives. However, Lareau (1994) also cautions that one needs to be careful in widespread encouragement of parental involvement. Some forms of parental involvement in schooling, namely, criticism and anger from parents toward teachers, are likely to inhibit achievement of their students.

The findings in this study are consistent with a large body of research implying that raising the income of female-headed families will increase the educational attainment and lower the out-of-wedlock teen childbearing rates of their children. Raising the income level could be done by increasing child support vigilance or by providing some other type of income support. Mayer (1997) cautions that the type of income transferred may matter. She estimates that one dollar of child support income provides a greater benefit to children than does one dollar of labor income, while welfare income provides a much smaller benefit than labor income. As she notes, her correlations may be picking up omitted variables such as above-average paternal involvement in families with child support, and below-average paternal involvement in families on welfare. In our regressions, the effects of income do not appear to largely be proxying for parental involvement, but our measures are likely to underreport the involvement of the noncustodial father.

In short, nontraditional families predict less desirable outcomes for youth. It appears that the lower income of female-headed households accounts for a substantial portion of the disadvantage. Unfortunately, these results cast no light on why children living with stepfathers do so poorly. Further research should explore the youth's lower expectations that the stepfather will remain present in the long term (as

families with stepfathers have higher divorce rates than families with two biological parents), the scarring effects of lower income when the family is female-headed, higher conflict levels within the family, and other possible causal channels.

*Appendix 1: What if Predetermined Variables Operate via Income?*

This appendix presents a formal model of how the estimated effects of income on family structure are affected when predetermined variables such as race or parental education affect youth outcomes because they affect income.

For simplicity, suppose there are only three causal variables that affect the single outcome dropout ( $d$ ): racial status ( $r$ ), single-parent family status ( $s$ ), and income ( $y$ ). Then we can rewrite the conventional reduced form (Equation (3)) as:

$$d = d(s, y, b) = \delta_s \cdot s + \delta_r \cdot r + \delta_y \cdot y. \quad (i)$$

To simplify notation, we will make use of the coefficients from the following *noncausal* auxiliary regressions:

$$y = \gamma_s \cdot s + \gamma_r \cdot r \quad (ii)$$

and

$$r = \rho_s \cdot s + \rho_y \cdot y. \quad (iii)$$

Substituting (iii) into (ii) yields:

$$\begin{aligned} y &= \gamma_s \cdot s + \gamma_r [\rho_s \cdot s + \rho_y \cdot y] = [\gamma_s + \gamma_r \cdot \rho_s] \cdot s + \rho_y \cdot \gamma_r \cdot y \\ &= [(\gamma_s + \gamma_r \cdot \rho_s) / (1 - \gamma_r \cdot \rho_y)] \cdot s. \end{aligned} \quad (iv)$$

Substituting (iv) in (iii) yields:

$$\begin{aligned} r &= \rho_s \cdot s + \rho_y \cdot [(\gamma_s + \gamma_r \cdot \rho_s) / (1 - \gamma_r \cdot \rho_y)] \cdot s \\ &= \{\rho_s \cdot (1 - \gamma_r \cdot \rho_y) + \rho_y \cdot [\gamma_s + \gamma_r \cdot \rho_s] / (1 - \gamma_r \cdot \rho_y)\} \cdot s \\ &= (\rho_s + \rho_y \cdot \gamma_s) \cdot s / (1 - \gamma_r \cdot \rho_y). \end{aligned} \quad (v)$$

Now the conventional reduced form procedure is to first estimate the system including  $r$  but not  $y$ , as in our Equation (2)

$$d = \alpha_s \cdot s + \alpha_r \cdot r \quad (vi)$$

and then with all three variables as in our Equation (3). The change  $(\alpha_s - \delta_s)$  is the difference in the single-parent effect attributed to lower income.

Then to find  $\alpha_s$ , we plug (ii) into (i),

$$\begin{aligned} d &= \delta_s \cdot s + \delta_r \cdot r + \delta_y (\gamma_s \cdot s + \gamma_r \cdot r) \\ &= (\delta_s + \delta_y \cdot \gamma_s) \cdot s + (\delta_r + \delta_y \gamma_r) \cdot r \end{aligned}$$

so that

$$\alpha_s = \delta_s + \delta_y \cdot \gamma_s$$

and the difference

$$(\alpha_s - \delta_s) = \delta_y \cdot \gamma_s. \quad (\text{vii})$$

Alternatively, one can estimate the two equations without  $r$ :

$$d = \zeta_s \cdot s$$

and

$$d = \theta_s \cdot s + \theta_y \cdot y.$$

Equation (i) implies  $d = \delta_s \cdot s + \delta_y \cdot y + \delta_r \cdot r$ . We can substitute equations (iv) and (v) into Equation (i) to yield the following:

$$\begin{aligned} d &= \delta_s \cdot s + \delta_y \cdot [(\gamma_s + \gamma_r \cdot \rho_s)/(1 - \gamma_r \cdot \rho_y)] \cdot s \\ &\quad + \delta_r \cdot [(\rho_s + \rho_y \cdot \gamma_s)/(1 - \gamma_r \cdot \rho_y)] \cdot s \\ &= s \cdot (\delta_s + [\delta_y \cdot \gamma_s + \delta_y \cdot \gamma_r \cdot \rho_s + \delta_r \cdot \rho_s + \delta_r \cdot \rho_y \cdot \gamma_s]/(1 - \gamma_r \cdot \rho_y)) \end{aligned}$$

so  $\zeta_s = (\delta_s + [\delta_y \cdot \gamma_s + \delta_y \cdot \gamma_r \cdot \rho_s + \delta_r \cdot \rho_s + \delta_r \cdot \rho_y \cdot \gamma_s]/(1 - \gamma_r \cdot \rho_y))$ .

Now estimating the system with  $s$  and  $y$  but not  $r$ :

$$d = \theta_s \cdot s + \theta_y \cdot y.$$

As before, we substitute Equation (iii) into Equation (i):

$$\begin{aligned} d &= \delta_s \cdot s + \delta_y \cdot y + \delta_r \cdot (\rho_s \cdot s + \rho_y \cdot y) \\ &= (\delta_s + \delta_r \cdot \rho_s) \cdot s + (\delta_y + \delta_r \cdot \rho_y) \cdot y \end{aligned}$$

so that

$$\theta_s = \delta_s + \delta_r \cdot \rho_s.$$

The difference  $(\zeta_s - \theta_s)$  is the difference in the single-parent coefficient attributed to lower income in the absence of predetermined characteristics.

$$\begin{aligned}
 \zeta_s - \theta_s &= [\delta_y \cdot \gamma_s + \delta_y \cdot \gamma_r \cdot \rho_s + \delta_r \cdot \rho_s + \delta_r \cdot \gamma_r \cdot \gamma_s] / (1 - \gamma_r \cdot \rho_y) - \delta_r \cdot \rho_s \\
 &= [\delta_y \cdot \gamma_s + \delta_y \cdot \gamma_r \cdot \rho_s + \delta_r \cdot \rho_s + \delta_r \cdot \gamma_r \cdot \gamma_s - \delta_r \cdot \rho_s + \delta_r \cdot \rho_s \cdot \gamma_r \cdot \rho_y] / \\
 &\quad (1 - \gamma_r \cdot \rho_y) \\
 &= [\delta_y \cdot \gamma_s + \delta_y \cdot \gamma_r \cdot \rho_s + \delta_r \cdot \gamma_r \cdot \gamma_s + \delta_r \cdot \rho_s \cdot \gamma_r \cdot \rho_y] / (1 - \gamma_r \cdot \rho_y).
 \end{aligned}
 \tag{viii}$$

We can rewrite the difference

$$(\alpha_s - \delta_s) \text{ as } [\delta_y \cdot \gamma_s - \delta_y \cdot \gamma_s \cdot \gamma_r \cdot \rho_y] / (1 - \gamma_r \cdot \rho_y). \tag{ix}$$

Then we can notice the difference between the estimated coefficient on family structure in the specification without predetermined variables and its estimate in the conventional reduced form. Both Equations (viii) and (ix) have common terms, so that the difference in the importance of income in lowering out the effect of single-parent when entered before predetermined variables  $(\zeta_s - \theta_s)$  and when it is entered after  $(\alpha_s - \delta_s)$  is

$$\begin{aligned}
 (\zeta_s - \theta_s) - (\alpha_s - \delta_s) &= [\delta_y \cdot \gamma_s + \delta_y \cdot \gamma_r \cdot \rho_s + \delta_r \cdot \gamma_r \cdot \gamma_s + \delta_r \cdot \rho_s \cdot \gamma_r \cdot \rho_y] / (1 - \gamma_r \cdot \rho_y) \\
 &\quad - [\delta_y \cdot \gamma_s - \delta_y \cdot \gamma_s \cdot \gamma_r \cdot \rho_y] / (1 - \gamma_r \cdot \rho_y) \\
 &= [\delta_y \cdot \gamma_r \cdot \rho_s + \delta_r \cdot \rho_y \cdot \gamma_s + \delta_r \cdot \rho_s \cdot \gamma_r \cdot \rho_y + \delta_y \cdot \gamma_s \cdot \gamma_r \cdot \rho_y] / (1 - \gamma_r \cdot \rho_y) \\
 &= [\delta_r \cdot \rho_y \cdot (\gamma_s + \rho_s \cdot \gamma_r) + \delta_y \cdot \gamma_r \cdot (\rho_s + \gamma_s \cdot \rho_y)] / (1 - \gamma_r \cdot \rho_y).
 \end{aligned}$$

To simplify notation, note that  $(\gamma_s + \rho_s \cdot \gamma_r)$  is just the coefficient  $C_{\gamma s}$  we would estimate from the noncausal equation  $Y = C_{\gamma s} \cdot s$ , and  $(\rho_s + \gamma_s \cdot \rho_y)$  is the coefficient  $C_{\rho s}$  from the equation  $R = C_{\rho s} \cdot s$ . With this notation we can substitute and find:

$$(\zeta_s - \theta_s) - (\alpha_s - \delta_s) = [\delta_r \cdot \rho_y \cdot C_{\gamma s} + \delta_y \cdot \gamma_r \cdot C_{\rho s}] / (1 - \gamma_r \cdot \rho_y) > 0.$$

In words, the importance of income in lowering out the apparent effect of single-parent status is higher when we do not control for predetermined variables (that is,  $[\zeta_s - \theta_s]$ ) than when we do  $(\alpha_s - \delta_s)$ .

The gap between the two estimated effects of income is particularly large when:

- 1)  $\delta_r$  is large, so predetermined variables are strong predictors of outcomes.
- 2)  $C\gamma_s$  is large, so income is highly correlated with single-parent status (whether the correlation is due to both the direct effect  $[\gamma_s]$  and via its correlation with the predetermined variable  $[\gamma_r]$ ).
- 3)  $\delta_y$  is large, so family structure matters a lot for income.
- 4)  $C\rho_s$  is large, so predetermined variables correlate closely with family structure both directly ( $\rho_s$ ) and indirectly via income (the  $\gamma_s \cdot \gamma_r$  term's meaning here is unclear to me).
- 5) Finally, when income is strongly affected by the predetermined variables ( $\rho_y$  is high), then  $1/(1 - \gamma_r \cdot \rho_y)$  is high and  $\delta_r \cdot \rho_y$  is high.

Thus, when income is strongly affected by the predetermined variables, the omission of predetermined variables matters more.

*Channels (3) and (4):* For policy purposes, if income matters a lot for outcomes ( $\delta_y$  is large), then we do not want to include predetermined variables that eliminate the possibility of this effect. This bias argues for dropping predetermined variables from the reduced form.

*Channels (1) and (2):* In the other direction, if predetermined variables matter a lot for the outcome ( $\delta_r$  is large), then NOT including predetermined variables will permit their effect to "load onto" our estimate (as long as income correlates with single-parent status and the predetermined variables). This bias argues for including predetermined variables.

If the two estimated effects of family structure conditional on income (that is,  $\theta_s$  and  $\delta_s$ ) are similar, then we know  $\delta_s - \theta_s = \delta_s - (\delta_s + \delta_r \cdot \rho_s) = \delta_r \cdot \rho_s$  must be small. That is, the direct effect of predetermined variables on outcomes (other than operating via income) must be small.

## Appendix 2

Probit Estimation ( <i>N</i> = 8,923)	Model 1B		Model 2B	
	DF/DX	Std. Error	DF/DX	Std. Error
<b>Dependent Variable = Dropping Out of High School</b>				
	<i>Pseudo-R</i> <sup>2</sup> = 0.168		<i>Pseudo-R</i> <sup>2</sup> = 0.172	
Income:needs in 8 <sup>th</sup> grade			−0.005**	0.002
Female headed	0.044**	0.006	0.037**	0.006
Stepfather	0.039**	0.007	0.038**	0.007
African American (omitted category is Caucasian)	−0.010*	0.004	−0.011*	0.004
Asian	−0.012	0.007	−0.013	0.007
Latino	0.014*	0.007	0.012	0.007
Female	0.004	0.003	0.004	0.003
Father foreign born	0.001	0.007	0.001	0.007
Mother foreign born	−0.016**	0.004	−0.016 **	0.004
Live in the south (omitted category is northeast)	0.016**	0.005	0.015**	0.005
Live in the west	0.016**	0.007	0.015*	0.007
Live in the central	0.010	0.006	0.010	0.006
Live in urban area (omitted category is suburb)	0.006	0.005	0.005	0.005
Live in rural area	−0.005	0.003	−0.006	0.003
Oldest child	−0.009**	0.003	−0.008**	0.003
Mother was a teen parent	0.010*	0.006	0.009	0.006
Father's education {z}	−0.026**	0.003	−0.024**	0.003
Mother's education {z}	−0.013**	0.002	−0.012**	0.002

*Notes:* Omitted family type is persistently intact. Standard errors are corrected for the presence of heteroskedasticity (White/Huber estimator) and for correlated residuals within schools.

\* and \*\* = test that coefficient is statistically significantly different from 0 at the *p* < 0.05 and 0.01 levels.

### Notes

1. In regressions that do not include either predetermined characteristics or contemporaneous characteristics, those characteristics are dropped from the first-stage Equation (4).
2. The parent questionnaire was filled out by the mother in 85% of the cases, and therefore may reflect the mother's characteristics rather than the father's. The school questionnaire was filled out by an administrator. Less than 2% of the sample was lost to attrition.
3. The main results of the study are unchanged when the missing values are included. The disadvantage of living in a stepparent family is less, but the impact of parental involvement and income is the same.
4. Additional stratification was explored, but did not significantly change the implications of the analysis. These include the various reasons for being a single parent such as being divorced, widowed, and having never been married.
5. A total of 1,245 values were imputed. For families with stepfathers, it is likely that most of the reports of father's education refer to the stepfather, not the biological father. Results are invariant to the inclusion or exclusion of the father's education variable.
6. The square of family income was also entered to capture nonlinear effects; its inclusion did not affect the results on family structure.
7. The analogous test would be to interact sex with family structure, which has been accomplished by numerous researchers (e.g., Lee et al. 1994).

### References

- Acock, A., and D. Demo. (1994). *Family Diversity and Well-Being*. Thousand Oaks, CA: Sage.
- Amato, P., and B. Keith. (1991). "Parental Divorce and Adult Well-Being: A Meta-Analysis." *Journal of Marriage and Family* 53: 43–58.
- Amato, P., L. Loomis, and A. Booth. (1995). "Parental Divorce, Marital Conflict, and Offspring Well-Being During Early Adulthood." *Social Forces* 73 (3): 895–915.
- Astone N., and S. McLanahan. (1991). "Family Structure, Parental Practices, and High School Completion." *American Sociological Review* 56: 309–320.
- Biblarz, T., A. Raftery, and A. Bucur. (1997). "Family Structure and Social Mobility." *Social Forces* 75 (4): 1319–1339.
- Bumpass, L. (1984). "Children and Marital Disruption: A Replication and Update." *Demography* 21: 71–82.
- Case, A., I-Fen Lin, and S. McLanahan. (1999). "Household Resource Allocation in Stepfamilies: Darwin Reflects on the Plight of Cinderella." *American Economic Review* 89 (2): 234–238.

- Cherlin, A. (1999). *Public and Private Families: An Introduction*. Boston, MA: McGraw-Hill.
- Cherlin, A. J., F. F. Furstenberg, P. Chase-Lansdale, P. Lindsay, K. E. Kiernan, P. K. Robins, D. R. Morrison, and J. O. Teitler. (1991). "Longitudinal Studies of Effects of Divorce on Children in Great Britain and the United States." *Science* 252 (5011): 1386–1389.
- Coleman, J. (1990). *Foundation of Social Theory*. Cambridge: Harvard University Press.
- . (1994). "Family Involvement in Education," in C. Fagnano and B. Werber, eds., *School, Family and Community Interaction: A View from the Firing Lines*. Boulder, CO: Westview Press.
- Downey, D. (1995). "Understanding Academic Achievement Among Children in Stepphouseholds: The Role of Parental Resources, Sex of Stepparent, and Sex of Child." *Social Forces* 73: 875–894.
- Entwisle, D., and K. Alexander. (1995). "A Parent's Economic Shadow: Family Structure Versus Family Resources as Influences on Early School Achievement." *Journal of Marriage and Family* 57: 399–409.
- Evenhouse, E., and S. Reilly. (2000). *Is the Wicked Stepmother Just a Fairytale?* Vanderbilt University, Department of Economics Working Paper No. 00-W49.
- Forehand, R., N. Long, and G. Brody. (1988). "Divorce and Marital Conflict: Relationship to Adolescent Competence and Adjustment in Early Adolescence," in E. Mavis Hetherington and J. D. Arasteh, eds., *Impact of Divorce, Single Parenting, and Stepparenting on Children*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Garasky, S. (1995). "The Effects of Family Structure on Educational Attainment: Do the Effects Vary by the Age of the Child?" *American Journal of Economics and Sociology* 54 (1): 89–105.
- Goldhaber, D., and D. Brewer. (1997). "Why Don't Schools and Teacher Seem to Matter? Assessing the Impact of Unobservables on Educational Productivity." *Journal of Human Resources* 32 (3): 505–523.
- Gordon, M. (1989). "The Family Environment of Sexual Abuse: A Comparison of Natal and Stepfather Abuse." *Child Abuse and Neglect* 13 (1): 121–130.
- Herrnstein, R., and C. Murray. (1994) *The Bell Curve*. New York: Free Press.
- Hoffman, S. D., E. M. Foster, and F. F. Furstenberg. (1993). "Reevaluating the Costs of Teenage Childbearing: Response to Geronimus and Korenman." *Demography* 30: 291–296.
- Hsiao, C. (1986). *Analysis of Panel Data*. Cambridge: Cambridge University Press.
- Jencks, C., and P. E. Peterson, eds. (1991). *The Urban Underclass*. Washington DC: Brookings Institute.
- Korenman, S., and S. Kinship. (1995). "A Reanalysis of *The Bell Curve*," NBER w.p. 5230.

- Korenman, S., J. E. Miller, and J. E. Sjaastad. (1994). *Long-Term Poverty and Child Development in the United States: Results from the NLSY*. Institute for Research on Poverty, University of Wisconsin–Madison. Discussion Paper 1044–94.
- Lareau, A. (1989). *Home Advantage: Social Class and Parental Intervention in Elementary Education*. New York: Falmer.
- . (1994). "Parental Involvement in Schooling: A Dissenting View," in C. Fagnano and B. Werber, eds., *School, Family and Community Interaction: A View from the Firing Lines*. Boulder, CO: Westview Press.
- Lee, V., D. Burkam, H. Zimiles, and B. Ladewski. (1994). "Family Structure and Its Effect on Behavioral and Emotional Problems in Young Adolescents." *Journal of Research on Adolescence* 4 (3): 405–437.
- Levine, D., and G. Painter. (1999). "The NELS Curve: Replicating *The Bell Curve* with the National Education Longitudinal Survey." *Industrial Relations* 38 (3): 364–401.
- Manski, C. (1993). "Identification Problems in the Social Sciences." *Sociological Methodology* 23: 1–56.
- Manski C., G. D. Sandefur, S. McLanahan, and D. Powers. 1992. "Alternative Estimates of the Effect of Family Structure During Adolescence on High School Graduation." *Journal of the American Statistical Association* 87 (417): 25–37.
- Mayer, S. (1997). *What Money Can't Buy: Family Income and Children's Life Chances*. Cambridge: Harvard University Press.
- McLanahan, S. (1985). "Family Structure and the Reproduction of Poverty." *American Journal of Sociology* 90 (4): 873–901.
- McLanahan, S., and G. Sandefur. (1994). *Growing Up With a Single Parent*. Cambridge: Harvard University Press.
- McLoyd, V. (1990). "The Impact of Economic Hardship on Black Families and Children: Psychological Distress, Parenting, and Socioemotional Development." *Child Development* 61: 311–346.
- McLoyd, V., and L. Wilson. (1991). "The Strain of Living Poor: Parenting, Social Support, and Child Mental Health," in *Children in Poverty: Child Development and Public Policy*. New York: Cambridge University Press.
- Milken, L. (1994). "Families, Communities, and Schools: A Perspective from the Foundations," in C. Fagnano and B. Werber, eds., *School, Family and Community Interaction: A View from the Firing Lines*. Boulder, CO: Westview Press.
- Morrison, D., and A. Cherlin. (1995). "The Divorce Process and Young Children's Well-Being: A Prospective Analysis." *Journal of Marriage and the Family* 57 (3): 800–812.
- National Center for Education Statistics. (1994). *National Education Longitudinal Study of 1988, Second Follow-Up: Student Component Data File User's Manual*. Washington DC.
- Painter G., and D. Levine. (2000). "Family Structure and Youths' Outcomes:

- Which Correlations are Causal?" *Journal of Human Resources* 35 (3): 524-549.
- Russell, D. (1984). "The Prevalence and Seriousness of Incestuous Abuse: Stepfathers vs. Biological Fathers." *Child Abuse and Neglect* 8 (1): 15-22.
- Sandefur, G., S. McLanahan, and F. Wojtkiewicz. (1992). "The Effects of Parental Marital Status During Adolescence on High School Graduation." *Social Forces* 71 (1): 103-121.
- Simons, R. L. (1996). *Understanding Differences Between Divorced and Intact Families: Stress, Interaction, and Child Outcome*. Thousand Oaks, CA: Sage.
- Solon, Gary. (1992). "Inter-Generational Income Mobility in the United States." *American Economic Review* 82 (3): 393-408.
- Statistical Abstract of the United States*. (1996). U.S. Bureau of the Census.
- Thomson, E., T. Hanson, and S. McLanahan. (1994). "Family Structure and Child Well-Being: Economic Resources vs. Parental Behavior." *Social Forces* 73 (1): 221-242.
- Wallerstein, J. S. (1991). "The Long-Term Effects of Divorce on Children: A Review." *Journal of American Academy of Child Adolescence: Psychiatry* 30 (3): 349-350.
- Zimmerman, David J. (1992). "Regression Toward Mediocrity in Economic Stature." *American Economic Review* 82 (3): 409-429.