A Sibling Study of Stepchild Well-being

Eirik Evenhouse
Siobhan Reilly

ABSTRACT

Examining 33 indicators of well-being from the National Longitudinal Study of Adolescent Health, we conclude that stepchildren's inferior outcomes are not entirely explained by sample selection. Using sibling comparisons to control for unobserved family characteristics, we identify stepparent effects by comparing half-siblings in families in which one child

Eirik Evenhouse is a fellow at the Institute for Health Policy Studies, University of California San Francisco. Siobhan Reilly is a professor of economics at Mills College. *For assistance and advice the authors thank George Akerlof, Kathryn Anderson, J. S. Butler, Anne Case, Ken Chay, Donna Ginther, Kathleen Mullan Harris, David Lee, David Levine, Shelley Lundberg, Charles Mullin, Steve Raphael, Claudia Strow, Joyce Tabor, Anne Winkler, and two anonymous referees. They also thank participants at the 2000 meetings of the Midwestern Economics Association and the Population Association of America, the 2001 Add Health Users Workshop, and the 2002 meeting of the American Economic Association. The authors acknowledge the financial support of the Vanderbilt Institute for Public Policy Studies, and of the National Institute for Child Health and Human Development under grant RO3HD39933.*

This research is based on data from the Add Health project, a program project designed by J. Richard Udry (PI) and Peter Bearman, and funded by grant P01-HD31921 from the National Institute of Child Health and Human Development to the Carolina Population Center, University of North Carolina at Chapel Hill, with cooperative funding participation by the National Cancer Institute; the National Institute of Alcohol Abuse and Alcoholism; the National Institute on Deafness and Other Communication Disorders; the National Institute on Drug Abuse; the National Institute of General Medical Sciences; the National Institute of Mental Health; the National Institute of Nursing Research; the Office of AIDS Research, NIH; the Office of Behavior and Social Science Research, NIH; the Office of the Director, NIH; the Office of Research on Women's Health, NIH; the Office of Population Affairs, DHHS; the National Center for Health Statistics, Centers for Disease Control and Prevention, DHHS; the Office of Minority Health, Centers for Disease Control and Prevention, DHHS; the Office of Minority Health, Office of Public Health and Science, DHHS; the Office of the Assistant Secretary for Planning and Evaluation, DHHS; and the National Science Foundation.

Because of restrictions on the use of Add Health data, the authors may share only their programming code. The data themselves can be obtained directly from The National Longitudinal Study of Adolescent Health, by arrangement with the Add Health Project, Carolina Population Center, 123 West Franklin Street, Chapel Hill, NC 27516-2524 (email: addhealth@unc.edu).

[Submitted June 2001; accepted July 2002]
ISSN 022-166X

has both parents and the other a parent and stepparent. Most estimated effects retain their sign after differencing across siblings, and a third remain statistically significant. The estimates' sensitivity to the choice of indicator suggests that studies based on a single measure of child well-being may be misleading.

I. Introduction

Most studies of family structure and child well-being conclude that children from stepfamilies "have outcomes very similar to children who grow up with only one parent, and worse than children who are raised by both of their biological parents" (Case, Lin, and McLanahan 1999, p. 234). One explanation, patterned on the folk model of medieval fairy tales, and rationalized by evolutionary psychologists (Daly and Wilson 1988, for instance), is that stepchildren fare poorly because stepparents stint them. An alternative view, supported by the work of Gennetian (1999), Ginther and Pollak (2002), and Hofferth and Anderson (2000), is that the observed correlations between family structure and children's well-being arise not from discriminatory treatment (or other stepfamily dynamics) but from the unobserved process by which people self-select into stepfamilies. The evidence presented in this paper suggests that the truth lies somewhere in between.

Drawing on the rich array of measures of child well-being in the National Longitudinal Study of Adolescent Health (Add Health), this study seeks to reduce selection bias by using sibling comparisons to control for unobserved parental and household characteristics. The effect of living with a stepparent is identified by differences between half-siblings in families in which one child has both biological parents present and the other has their common parent and a stepparent. These within-family comparisons suggest that the poorer outcomes of stepchildren cannot be attributed solely to sample selection. Despite a radical reduction in sample size as we move from between-family to within-family estimates, a third of the stepparent effects remain statistically significant. Over four-fifths of the point estimates retain their sign after differencing, and as many increase as shrink. Furthermore, the estimates' sensitivity to the choice of indicator suggests that studies that focus on only one or two indicators of child well-being may be misleading.

II. Background and Significance

How family structure matters to children's well-being has become an area of active research (see the surveys in McLanahan 1998 or Ginther and Pollak 2002, for example). The attention is warranted. Only a minority of American children spend their entire childhood with both biological parents (Dawson 1991).[1] A third of children live apart from one biological parent, and 5 percent apart from both (Evenhouse and Reilly 2003). Thirty percent spend time in a stepfamily (Bumpass, Raley, and Sweet, 1995). One goal of recent welfare reforms is encouraging mar-

1. In the 1994 data used in this study, only 47 percent of seventeen-year-olds live with both parents.

riage, and many marriages that could result would be to stepfathers rather than to fathers.[2] The past decade has also seen a wave of legislative efforts to tighten state divorce laws. The most commonly offered motivation has been a desire to benefit children, suggesting a belief by legislators and advocates that children do better with both biological parents than with a single parent or a stepparent.

While most research on family structure has contrasted one- with two-adult families (McLanahan and Sandefur 1994, for example), a growing strand of the literature compares children who live with both parents to those who live with a parent and stepparent (usually a stepfather). These studies usually find that stepchildren have significantly worse outcomes. For example, they score six to eight percentage points lower on standardized high school entrance exams at age 14 (Dronkers 1994), complete six months less of schooling (Sandefur and Wells 1997), and are 8 percent less likely to finish high school (Garasky 1995). They are nearly three times as likely to be incarcerated by age 18 (McLanahan and Harper 1998), and five times as likely to leave home due to conflict (Mitchell 1994). In the extreme, young children are 40 times as likely to suffer abuse (Daly and Wilson 1985), and 70 times as likely to be killed by the man in the household (Daly and Wilson 1988). By many criteria, children with stepfathers fare no better than the children of single mothers (McLanahan 1998).

Stepparent effects are mismeasured, however, if selection bias arises from the omission of variables that affect children's well-being and that are also correlated with being in a stepfamily. Unobserved maternal characteristics are one possible example—the same factors that helped separate a woman from her child's father may also handicap her as a mother. Heritable paternal traits are another.

A common remedy for selection bias is to use instrumental variables. For instance, in analyzing the link between a male youth's odds of incarceration and his stepchild status, McLanahan and Harper (1998) use his state's 1975 divorce rate, an indicator that his mother's education exceeds his father's, and several demographic controls to create an instrument for family type. Noting the low correlation between the error terms of the equation predicting incarceration and of the equation predicting stepchild status, they conclude that selection bias is negligible.

A drawback of the instrumental variable approach is the difficulty of finding suitable instruments. For the question at hand, an ideal instrument is highly correlated with family structure but uncorrelated with the error term of the outcome equation being estimated, and such a variable is hard to find. The higher the correlation between the instrument and family structure, the less plausible the claim that it is uncorrelated with the disturbances. The weaker the correlation between instrument and family structure, the less useful the approach. The latter consideration is particularly relevant, as researchers typically can explain only a very small fraction of the variation in individuals' marital status or living arrangements (6 to 10 percent in Evenhouse and Reilly 1997, 2003, for example).

An alternative to instrumental variables is a fixed-effects approach. Longitudinal data, for example, allow one to incorporate parent or child fixed effects. Cherlin *et*

2. By counting the income of resident fathers but disregarding much or all of the income of unrelated male cohabitors, welfare eligibility rules give low-income mothers an incentive to live with unrelated men rather than with the fathers of their children. See Evenhouse and Reilly (2003) or Moffitt, Reville, and Winkler (1998) for discussion of these rules.

al. (1991), studying the impact of divorce, find that controlling for a child's test scores at age seven reduces the significance and size of the estimated effect of divorce on test scores at age 11. In contrast, Painter and Levine (2000) find that controlling for parent and child characteristics measured when the child is in 8th grade does not significantly reduce the estimated effect of a subsequent divorce or, more pertinent for this study, a subsequent remarriage on the child's outcomes four years later. Such before-and-after comparisons, however, control only for omitted factors that are constant over the intervening period.

Sibling data permit the use of a family fixed effect. Geronimus and Korenman (1992), for example, studying the consequences of teen childbearing, compare sisters who had their first births at different ages and conclude that cross-sectional studies tend to overstate the consequences. Sandefur and Wells (1997) look at the educational attainment of stepchildren. They compare pairs of siblings in the National Longitudinal Survey of Youth (NLSY), relating each sibling's family structure at age 14 to his or her eventual educational attainment, and conclude that controlling for common family background lowers by half the educational disadvantage associated with being in a stepfamily at age 14, to roughly six months. A drawback of their approach is that important omitted variables may diverge between two siblings, because "family structure at age 14" is observed in a different calendar year for each sibling, and because the educational outcomes are observed years later.

This study, like Gennetian (1999), Ginther and Pollak (2002), and Hofferth and Anderson (2000), controls more fully for time-varying unobserved variables. All variables are measured at the same time. Pairwise sibling comparisons eliminate parental factors or household conditions, observed or not, that affect two siblings equally and that therefore cannot explain differences in their well-being. Stepparent effects are identifiable because, in some families containing half-siblings, one sibling has both biological parents present while another has a parent and a stepparent. Retrospective data are used to control for family structure transitions that were not experienced by both siblings. Such controls are vital because having had these experiences is strongly correlated with being the stepchild in a half-sibling pair.

Also like Gennetian (1999), Ginther and Pollak (2002), and Hofferth and Anderson (2000), we focus on half-siblings in blended families. Our data, from the National Longitudinal Study of Adolescent Health (Add Health), offer several advantages, however. Ginther and Pollak (2002) and Gennetian (1999) are limited by a dearth of measures in the NLSY-Child survey that are amenable to a sibling comparison approach, as the survey contains only a few questions asked of all children regardless of age.[3] Hofferth and Anderson (2000) use the Child Development Supplement (CDS) to the Panel Study of Income Dynamics (PSID), which has more measures that are comparable between siblings but many fewer half-sibling pairs. In contrast, Add Health contains more half-sibling pairs than either the PSID or NLSY, the respondents are in a narrower age range,[4] all respondents are asked the same questions, and the questions cover a wider range of

3. Both studies examine children's Peabody Individual Achievement Test (PIAT) scores in math and reading. In addition, Gennetian examines the cognitive subscore of the Home Observation for Measurement of the Environment (HOME) assessment, and Ginther and Pollak examine the NLSY's behavioral problem index (BPI).

4. Ninety-seven percent are teenagers in the 1994 wave.

topics. We exploit Add Health's richness by examining 33 different measures of child outcomes or parental investment.

III. Estimation strategy

The link between family structure and children's well-being can be modeled as operating through parental effort and productivity. A parent makes a constrained choice of the effort to invest in a given child. Parents are constrained by their resources (health, wealth, and time, for example). Effort may take many forms, but most are difficult to measure (expenditures of patience or thought, for example). Observable effort consists of reported investments of time or money in the child. Effort depends not only on a parent's own characteristics, but also on those of the child (Lundberg and Rose 2002, for example, find that a child's gender affects fathers' earnings and labor supply).

The productivity of a parent's effort depends on the characteristics of the parent (such as age or education), the child (such as age, sex, or physical attractiveness), and the child's environment (such as school quality or the local crime rate). The question motivating this research is whether a parent's effort or productivity depends on the parent's biological relationship to the child.

As indicators of child well-being, we use parental investment measures as well as child outcome measures. Parental investment measures are particularly germane to the differential-treatment hypothesis, as adults have more control over their parenting inputs than over child outcomes. They may also be more closely related to current variables than are child outcomes, making it less vital to control for differences in siblings' past experiences. The Add Health survey contains many more outcome than investment measures, however. Furthermore, adolescents have say, too, and it may be the child who resists participating in extracurricular programs or sharing activities with the stepparent.

Given the difficulty of observing important aspects of parenting behavior and the many theoretically plausible interactions among the observable characteristics of a family, we estimate two reduced-form models of child well-being. We first estimate a between-family model:

$$(1) \quad W_{ih} = \beta_0 + \beta_1 M_{ih} + \beta_2 F_{ih} + \beta_3 C_{ih} + \beta_4 E_{ih} + \varepsilon_{ih}$$

where i indexes the child and h the household; $\boldsymbol{W}$ is a measure of child well-being; $\boldsymbol{M}$, $\boldsymbol{F}$, and $\boldsymbol{C}_i$ are vectors of characteristics of the male adult, female adult, and child; $\boldsymbol{E}$ is a vector of environmental characteristics (such as school and neighborhood); and ε_{ih} is the error term. Estimating this model for each measure of well-being provides a benchmark for the size and significance of stepparent effects in the absence of a family fixed effect.

The error term is assumed to consist of a man-specific error, μ_{ih}, a woman-specific error, ω_{ih}, a child-specific error, γ_{ih}, an environment-specific error, η_{ih}, and a random error, v_{ih}:

$$(2) \quad \varepsilon_{ih} = \mu_{ih} + \omega_{ih} + \gamma_{ih} + \eta_{ih} + v_{ih}.$$

In estimating Equation 1, selection bias arises if any of the person- or environment-

specific error terms are correlated with family structure. If family structure is largely independent of children's characteristics, most of the bias will result from parent- or environment-specific errors.

We then estimate a sibling-difference, or within-family, model. By first-differencing between two siblings, the portion of the bias that is due to adult or environmental characteristics that are constant across the siblings is eliminated. This second model is given by:

$$(3) \quad \Delta_{ij} W_h = \beta_1 \Delta_{ij} M_{ijh} + \beta_2 \Delta_{ij} F_{ijh} + \beta_3 \Delta_{ij} C_{ijh} + \beta_4 \Delta_{ij} E_{ijh} + \Delta_{ij} E_{ijh}$$

with the subscript h indexing the household and ij denoting a comparison between sibling i and sibling j. The vectors of differenced variables for the man and woman, $\Delta \boldsymbol{M}_{ijh}$ and $\Delta \boldsymbol{F}_{ijh}$, contain zeros except in cases when two siblings do not have identical relationships to an adult (for instance, when they are half-siblings). These are the cases that identify the effect of living with a stepparent. This specification allows the estimation of distinct stepfather and stepmother effects. Comparing the within-family estimates from Equation 3 to the between-family estimates of Equation 1 gives an idea of the magnitude of selection bias present in the former.

Eliminating parental factors would leave only differences between siblings to predict differences in the dependent variable. Differences between siblings are potential sources of bias if they are omitted. For example, if eldest children have better outcomes on average, then it is important to control for birth order, as being an eldest child is highly correlated with being the stepchild in a half-sibling pair. The inclusion of sibling pairs from intact and single-parent families helps in this regard. It is also important to control for differences in the siblings' past experiences, as a sibling who is a stepchild may have had difficult experiences that a half-sibling has not.

Sibling differencing cannot remove all selection bias. Indeed, if explanatory variables are mismeasured, differencing may actually exacerbate the selection problem (Griliches 1979). Although family structure can be determined more reliably than many variables, there is nonetheless room for error.[5] In addition to measurement error, any unobserved differences between siblings that are correlated with family structure remain as sources of bias. A child's difficult personality, for example, may raise the odds that parents separate, as well as the child's odds of having poor outcomes. With personality largely unobserved, sibling differencing would lead one to overstate the causal importance of stepchild status. Similarly, if a stepchild's poor outcomes are due to unobservable genetic traits inherited from the absent parent, and those traits were a factor in the parents' separation, then the poor outcomes may be wrongly attributed to having a stepparent. Bias can go in the other direction, too. For example, efforts by a biological parent to compensate for the deficiencies of a stepparent may mask the stepparent effect.

A. Data

The data, from the first wave of the National Longitudinal Study of Adolescent Health (Add Health), are well suited to the purposes of this study. In 1994, Add

5. For instance, some families that we label "intact" may actually include an older stepchild who no longer lives at home and is therefore omitted from the Add Health survey.

Table 1A
Which parents do children live with?

	Both parents	Mother only	Mother and stepfather	Stepmother and father	Father only
Whole sample					
Number	8,961	4,409	2,978	571	503
Percentage	53.0	23.8	17.3	3.1	2.8
Sibling sample					
Number	1,964	927	745	186	96
Percentage	50.1	23.7	19.0	4.7	2.5

Note: Data from the first wave of the Add Health survey. Percentages incorporate weights. The 13.2 percent of the sample who are adopted, are foster children, or live with neither biological parent are excluded.

Health began following more than 20,000 adolescents. Among the 2,734 pairs of adolescent siblings are 442 pairs of half-siblings. Information on the adolescents comes from the children themselves, from their parents, from their network of school friends, and from school administrators.

Besides educational outcome measures, Add Health has information about the adolescents' sexual activity, their drug and alcohol use, the characteristics of their social networks, and their emotional health. It also contains proxies for parental investment, such as the number of activities shared by parent and child, the level of the child's extracurricular activities, whether the child attends a good school, whether the child attends private school, how often the parent is home when the child goes to bed, the fraction of evening meals eaten together, parent involvement in the child's schoolwork, the child's weekly allowance, and the hours the child spends watching television. In addition, the survey contains enough retrospective information to allow the construction of variables that capture aspects of a child's path into a stepfamily, such as "years spent living with a stepparent," "years spent with both parents," "years spent in a one-parent household," "child never lived with other biological parent," "child has experienced zero, one, or two or more residential moves," and "frequency of contact with absent biological parent."

Tables 1A and 1B summarize the family structures of Add Health children. Table 1A reports on the whole sample and, for the sake of comparison, the sibling sample. Table 1B reports on family types in the sibling sample, using the finer categories that surveying a child's sibling allows. Thus Table 1B differentiates among stepfamilies that we label as "pure" (one adult is the parent of all the children), "blended" (the adults also have a mutual child), and "Brady Bunch" (each parent brought children from a previous relationship and there may or may not be a mutual child). We define a "stepfamily" as one in which the biological parent reports living with a partner (married or unmarried) who is not the child's parent. Because we rely on half-siblings in blended families to identify stepparent effects, Table 1B also reports the distribution of family type within that subgroup. Table 1B shows nearly

Table 1B
Family types of children in the sibling sample

	Intact	Mother only	Pure stepfather	Blended stepfather	"Brady Bunch"	Father only	Pure stepmother	Blended stepmother
All siblings								
Number	1,804	830	592	296	190	82	72	52
Percentage	51.9	17.6	13.0	7.0	6.0	1.7	1.4	1.2
Half siblings								
Number	11	209	160	174	4	4	5	30
Percentage	1.9	38.7	22.6	29.7	0.8	0.0	0.7	5.5

Notes: Data from the first wave of the Add Health survey. Percentages incorporate weights. A "Brady Bunch" family is formed when parents each bring children from previous relationships. A blended stepfamily is formed when a parent and stepparent also have a mutual child.

two-thirds of half-siblings in mother-only or pure stepfather families.[6] Thus the identification of stepparent effects hinges on the 35.2 percent of half-siblings who are in blended stepfather or blended stepmother families.

IV. Empirical Results

The Add Health survey contains many more indicators of adolescent well-being than the 33 presented here, which are intended as a representative assortment. Some are subjective measures of the sort that economists usually avoid, but that we include because we think they contain valid information about adolescents' emotional and psychological well-being.[7]

For the sake of brevity, our discussion of results does not focus on individual indicators and their point estimates. The size or significance of a particular effect may be sensitive to the choice of regression technique, but our general conclusions are not. For ease of interpretation, all coefficients reported below are from OLS regressions.

As a precursor to regression analysis, Tables 2A and 2B present differences in unadjusted means, between families and between siblings, respectively. Table 2A reports, for all 33 indicators, the mean difference between children in stepfather, stepmother, or mother-only families and children in intact families.[8] Looking at the coefficients' signs, we see that, by 30 measures, children living with stepfathers fare worse than children living with both parents and, by ten measures, worse even than children in mother-only families. Similarly, children living with stepmothers do worse than those living with both parents by 27 measures and worse than children in mother-only families by 17 measures. Thirty of the stepfather effects are significantly different from zero, as are 26 of the stepmother effects.

When stepchildren are compared to the half-siblings they live with, their disadvantages look much less pronounced. Table 2B presents, for the same indicators, the difference between (a) the average differential between two siblings (full or half) who have the same relationship to the adult(s) in the family, and (b) the average differential between two half-siblings who do not (namely, when one of the adults is parent to one and stepparent to the other).[9] The figure of -0.289 for a child's

6. Eleven half-siblings are reportedly living in an intact family. While these apparent inconsistencies in family type may result from survey errors, they may also be cases of children who divide their time between two residences, an arrangement not captured well by Add Health's questions.

7. In medicine there is increasing recognition that individuals' subjective assessments of their physical health are good predictors of subsequent morbidity and mortality, as good as and often better than the objective assessments made by doctors (see Epstein 1990). Clinical assessments of mental health are often based on self-reported measures.

8. Recall that the classification by family type is child-based, with each child classified by his or her relationship to the household's adult(s). Thus two half-siblings in the same household may be classified differently. Sample size varies by indicator, and is typically around 16,800 observations, except for indicators concerning fathers or stepfathers specifically, when it is around 9,900 observations.

9. Each differential is a sibling comparison made by subtracting an indicator's value for the younger sibling from that for the elder. While it may seem more logical at first glance, a "stepchild minus biochild" rule for calculating sibling differences will work neither for full siblings nor for all half-siblings (for example, when only their common parent is present. The "older minus younger" rule can be applied to any sibling pair.)

Table 2A
Mean differences in child well-being across family types

Measures of child well-being	Stepfather vs. intact	Stepmother vs. intact	Mother-only vs. intact
Measures of parental investment			
Child attends a private school	**−0.031** (0.006)	**−0.032** (0.011)	**−0.039** (0.005)
Child saw dentist in past 12 months	**−0.094** (0.014)	**−0.134** (0.028)	**−0.127** (0.012)
Child wears braces on teeth	**−0.031** (0.008)	**−0.051** (0.013)	**−0.052** (0.007)
Number of child's extracurricular activities	**−0.264** (0.058)	**−0.484** (0.097)	**−0.340** (0.051)
Child shares few activities with (step)father	**0.146** (0.016)	−0.037 (0.023)	
Child shares few activities with (step)mother	−0.007 (0.012)	**0.189** (0.032)	−0.005 (0.011)
Outcome measures			
Education measures			
Child's self-reported four-subject GPA	**−0.262** (0.023)	**−0.267** (0.046)	**−0.276** (0.020)
Child ever held back a grade	**0.094** (0.012)	**0.112** (0.026)	**0.126** (0.011)
Child reports getting in trouble at school	**0.040** (0.005)	**0.047** (0.010)	**0.031** (0.005)
Child reports having been suspended from school	**0.153** (0.013)	**0.157** (0.027)	**0.183** (0.011)
Child reports having been expelled from school	**0.025** (0.005)	**0.043** (0.015)	**0.052** (0.006)
Measures of risky behavior			
Child reports having had sex	**0.171** (0.014)	**0.131** (0.028)	**0.163** (0.012)
Child reports drinking alcohol away from adults	**0.071** (0.014)	**0.069** (0.028)	**0.029** (0.012)
Child reports having used marijuana	**0.126** (0.013)	**0.178** (0.028)	**0.119** (0.012)
Child reports having used "hard" drugs	**0.219** (0.025)	**0.324** (0.058)	**0.167** (0.022)
Child's percentile rank for delinquent behavior	**4.89** (0.80)	**7.69** (1.54)	**4.47** (0.72)

Table 2A (*continued*)

Measures of child well-being	Stepfather vs. intact	Stepmother vs. intact	Mother-only vs. intact
Measures of child's social network			
Number of students naming child as a friend	**−0.531** (0.135)	**−0.629** (0.252)	**−0.838** (0.126)
Mean GPA of child's friends	**−0.150** (0.021)	**−0.188** (0.047)	**−0.181** (0.020)
Mean number of extracurricular activities of friends	**−0.303** (0.048)	**−0.282** (0.105)	**−0.277** (0.046)
High alcohol/tobacco/marijuana use by best friends	**0.038** (0.011)	**0.032** (0.022)	−0.005 (0.009)
Measures of relationship quality			
"My (step)mother is mostly warm and loving"	**−0.045** (0.009)	**−0.145** (0.025)	**−0.047** (0.008)
"My (step)father is mostly warm and loving"	**−0.147** (0.015)	−0.028 (0.022)	
"My (step)mother cares for me "	**−0.020** (0.009)	**−0.297** (0.029)	**−0.028** (0.008)
"My (step)father cares for me"	**−0.135** (0.012)	−0.009 (0.011)	
"I feel close to my (step)mother"	0.002 (0.008)	**−0.226** (0.028)	**−0.012** (0.007)
"I feel close to my (step)father"	**−0.200** (0.016)	**0.054** (0.018)	
"Have poor relationship with (step)mother"	**0.039** (0.009)	**0.121** (0.025)	**0.038** (0.008)
"Have poor relationship with (step)father"	**0.113** (0.015)	0.017 (0.020)	
"I badly want to leave home"	**0.054** (0.010)	**0.049** (0.020)	**0.054** (.009)
"Last physical fight was with a family member"	0.007 (0.007)	−0.002 (0.013)	0.005 (.006)
Measures of emotional health			
Child contemplated suicide seriously in past year	**0.036** (0.010)	**0.086** (0.024)	**0.017** (.008)
Child often feels depressed	**0.211** (0.042)	**0.143** (0.094)	**0.148** (.039)
Child has poor self-image	**0.047** (0.014)	0.045 (0.029)	**0.076** (.013)

Notes: Classification of family type is child-based, with child classified by relationship to household's adults. Sample size ranges from about 9,900 to 16,800. Standard errors (in parentheses) adjusted for within-family correlation. Bold font indicates statistical significance at the 10 percent level or higher. The appendix gives more detail about the indicators.

Table 2B
(Half-)sibling differences and stepparent effects

Differenced measures of child well-being	Stepfather vs. father	Stepmother vs. mother
Differenced measures of parental investment		
Child attends private school	0.009 (0.010)	0.000 (0.002)
Child saw dentist in past 12 months	−0.056 (0.069)	−0.057 (0.128)
Child wears braces on teeth	0.044 (0.056)	−0.071 (0.114)
Number of child's extracurricular activities	0.301 (0.364)	0.429 (0.628)
Child shares few activities with (step)father	0.124 (0.087)	−0.133 (0.176)
Child shares few activities with (step)mother	−0.033 (0.076)	**0.474** (0.174)
Differenced outcome measures		
Differenced education measures		
Child's self-reported four-subject GPA	**−0.254** (0.107)	−0.095 (0.224)
Child ever held back a grade	0.031 (0.064)	0.043 (0.088)
Child reports getting in trouble at school	**0.056** (0.026)	0.025 (0.063)
Child reports having been suspended from school	**0.134** (0.063)	**−0.198** (0.111)
Child reports having been expelled from school	−0.010 (0.014)	−0.049 (0.044)
Differenced measures of risky behavior		
Child reports having had sex	**0.279** (0.090)	**0.348** (0.164)
Child reports drinking alcohol away from adults	**0.147** (0.072)	**0.530** (0.123)
Child reports having used marijuana	**0.224** (0.071)	**0.230** (0.134)
Child reports having used any of five "hard" drugs	**0.324** (0.119)	0.115 (0.379)
Child's percentile rank for delinquent behavior	**9.41** (4.57)	2.72 (6.97)

Table 2B (*continued*)

Differenced measures of child well-being	Stepfather vs. father	Stepmother vs. mother
Differenced measures of child's social network		
Number of students naming child as a friend	−2.19 (1.36)	0.74 (2.82)
Mean GPA of child's friends	0.134 (0.115)	−0.225 (0.402)
Mean number of extracurricular activities of friends	0.680 (0.483)	0.118 (0.834)
High alcohol/tobacco/marijuana use by best friends	0.052 (0.046)	0.214 (0.154)
Differenced measures of relationship quality		
"My (step)mother is mostly warm and loving"	0.022 (0.042)	−0.145 (0.108)
"My (step)father is mostly warm and loving"	−0.091 (0.071)	−0.207 (0.139)
"My (step)mother cares for me "	0.069 (0.068)	**−0.346** (0.146)
"My (step)father cares for me"	**−0.162** (0.049)	0.064 (0.051)
"I feel close to my (step)mother"	**0.080** (0.049)	**−0.221** (0.127)
"I feel close to my (step)father"	**−0.296** (0.072)	0.118 (0.151)
"Have unsatisfactory relationship with (step)mother"	0.001 (0.052)	0.025 (0.054)
"Have unsatisfactory relationship with (step)father"	**0.129** (0.075)	0.151 (0.126)
"I badly want to leave home"	−0.101 (0.068)	−0.108 (0.121)
"Last physical fight was with a family member"	−0.007 (0.039)	−0.094 (0.068)
Differenced measures of emotional health		
Child contemplated suicide seriously within past year	**0.125** (0.065)	0.116 (0.085)
Child often feels depressed	−0.033 (0.247)	**−0.143** (0.081)
Child has poor self-image	**−0.199** (0.100)	**−0.496** (0.193)

Notes: Each variable measured as differential between two (half-)siblings. Table reports correlation between outcome differential and any difference between siblings' relationships to an adult in the family. Sample size varies slightly by measure, but is about 2,000, except for indicators concerning fathers or stepfathers specifically, where it is roughly 1,350. Standard errors (in parentheses) adjusted for within-family correlation (multiple pairings in some families). Bold font indicates statistical significance at the 10 percent level or higher. The appendix gives more detail about the indicators.

four-subject grade point average (GPA), for example, implies that the average GPA differential between two siblings is roughly three-tenths of a point smaller when the man is the stepfather, rather than the father, of the older sibling. For two-thirds of the measures, the sign on the difference is consistent with worse outcomes for the stepchild. The stepfather effect is statistically significant for only 14 of the 33 indicators, however, and the stepmother effect for only nine.

The picture given by unadjusted means is suggestive, but only a beginning. Families differ in many ways that matter to child well-being and that are correlated with family structure, and half-siblings may differ from each other in ways correlated with their stepchild status. Adding controls for family and child characteristics, we estimate for each indicator the between-family regression—Equation 1—and then the within-family regression—Equation 3. Table 3 lists these controls and their averages by family type. Most of the controls are ones that can differ between siblings, because in sibling comparisons, family-level characteristics such as the income-to-needs ratio will drop out.

Two representative between-family regressions are presented in Table 4A, the first predicting a child's grade point average and the second whether a child has had sex yet. Table 4B presents the corresponding within-family regressions, in which the unit of observation is not an individual but a pair of (half-)siblings and the dependent variables and most explanatory variables are measured as differences between the two siblings. Age is treated nonparametrically in both types of regression.[10] The within-family regressions also use separate dummy variables for "older boy/younger girl" and "older girl/younger boy" comparisons (the default being same-sex comparisons).

As the two examples in Table 4A suggest, the estimated stepparent effects in a between-family regression are consistent with, if a bit smaller than, the simple differences in unadjusted means reported in Table 2A. Compared with children living with both biological parents, stepchildren tend to have lower grades (by 0.15 to 0.35 grade points, or by a fifth to a half of a standard deviation) and are more likely to have had sex (by 11 to 13 percentage points).

For the within-family model, in contrast, the differences in unadjusted means (Table 2B) are poorer predictors of the differences in regression-adjusted means. Looking at Table 4B, stepfather and stepmother effects on the probability that the child has had sex both become insignificant. The stepmother effect on the child's GPA also becomes insignificant. The stepfather effect on GPA, however, remains significant, and the point estimate rises slightly (from -0.254 in Table 2B to -0.288).

Taking all our indicators as a whole, the between-family stepparent effects are broadly consistent with the results of previous studies. Judging by the estimates summarized in Table 5A, children living with a parent and stepparent fare worse than children living with both parents by every single measure. Stepfather effects are statistically significant for 25 of the 33 indicators, and stepmother effects for 20. Moreover, stepchildren fare worse than children in mother-only families by 12 of the 28 measures computed for mother-only families.

Table 5B summarizes the estimates from the within-family model. These effects should be interpreted in the context of much reduced sample sizes, which are around

10. Between-family regressions use age dummies, and within-family regressions use a dummy for each possible age combination. Age is measured in years.

Table 3
Descriptive statistics for regression controls

Explanatory variable	Stepfather family	Stepmother family	Mother-only family	Intact family
Pre-teen years in two-parent family	3.7 (4.2)	7.3 (4.3)	4.5 (4.8)	
Pre-teen years in one-parent family	4.1 (3.9)	3.6 (3.9)	5.6 (4.6)	
Pre-teen years in stepfamily	4.2 (3.9)	1.0 (2.1)	1.9 (3.2)	
Exposed to two or more stepparents	0.37	0.03	0.15	
Female	0.52	0.42	0.53	0.50
Household income-to-needs ratio	2.7 (2.9)	3.1 (2.3)	2.0 (3.3)	3.4 (3.7)
(Step)mother works full-time	0.64	0.66	0.69	0.56
Alcoholic biological parent	0.25	0.18	0.21	0.06
Caucasian	0.66	0.75	0.47	0.71
African-American	0.24	0.14	0.43	0.14
Asian	0.03	0.08	0.04	0.08
Hispanic	0.16	0.11	0.18	0.17
Eldest child	0.63	0.50	0.54	0.44

Low birth weight (less than 5.5 lbs)	0.07	0.09	0.07	0.05
Retarded	0.01	0.01	0.02	0.01
Looks old for age	0.10	0.12	0.13	0.12
Looks young for age	0.14	0.10	0.12	0.09
Born in United States	0.94	0.91	0.93	0.92
12 years old	0.002	0.002	0.001	0.002
13 years old	0.035	0.025	0.033	0.037
14 years old	0.120	0.088	0.113	0.124
15 years old	0.150	0.158	0.149	0.143
16 years old	0.199	0.184	0.183	0.182
17 years old	0.199	0.203	0.197	0.196
18 years old	0.166	0.205	0.177	0.184
19 years old	0.109	0.114	0.122	0.121
20 or 21 years old	0.009	0.023	0.027	0.013
Number of observations	2,978	571	4,409	8,961

Notes: Data from 1st wave (1994) of Add Health survey.

Table 4A
Two examples of between-family regressions

Covariates	Outcome	
	Self-reported GPA	Has child had sex yet?
Child lives with stepfather	**−0.154** (0.034)	**0.127** (0.019)
Child lives with stepmother	**−0.351** (0.065)	**0.113** (0.038)
Child lives in mother-only family	**−0.084** (0.031)	**0.087** (0.018)
Number of pre-teen years in a one-adult family	**−0.008** (0.003)	0.000 (0.002)
Number of pre-teen years in a stepfamily	−0.003 (0.004)	0.001 (0.003)
Child has lived with two or more stepparents	0.025 (0.035)	0.002 (0.021)
Child is female	**0.213** (0.016)	**−0.017** (0.009)
Child is African-American	**−0.158** (0.020)	**0.155** (0.012)
Child is Asian	**0.118** (0.039)	**−0.040** (0.022)
Child is Hispanic	**−0.185** (0.026)	0.004 (0.014)
Child was born in United States	**−0.073** (0.035)	**0.085** (0.020)
Eldest child	**0.065** (0.016)	−0.005 (0.009)
Low birth weight	0.020 (0.039)	**−0.050** (0.019)
Child looks younger than actual age	**−0.108** (0.027)	−0.016 (0.014)
Child looks older than actual age	−0.032 (0.024)	**0.116** (0.014)
Income-to-needs ratio	**0.055** (0.004)	**−0.009** (0.002)
Income-to-needs ratio, squared ($\times 10^{-3}$)	**−0.695** (0.101)	**0.100** (0.032)
Child has an alcoholic biological parent	**−0.107** (0.025)	**0.067** (0.014)
Age dummies	Yes	Yes
Number of observations	16,365	16,343

Notes: Table reports OLS coefficients. Standard errors (in parentheses) adjusted for within-family correlation. Bold font denotes significance at the 10 percent level or higher.

1,350 for indicators that relate to fathers or stepfathers and roughly 2,000 for other indicators. Stepparent effects are identified by a much smaller number of half-sibling pairs. Recall that most half-siblings are not in a blended stepfamily, but in mother-only or pure stepfather families (Table 1B). Thus, stepfather effects are identified by about 90 half-sibling pairs, and stepmother effects by only fifteen.

Comparison of the within-family to the between-family estimates suggests that it would be wrong to interpret stepchildren's poor outcomes as largely a product of sample selection. Consider first the 28 measures that pertain specifically to (step)fa-

Table 4B
Two examples of regression using within-family differences

Differenced covariates	Outcome	
	Self-reported GPA	Has child had sex yet?
Child lives with stepfather	**−0.288** (0.171)	0.110 (0.126)
Child lives with stepmother	−0.005 (0.306)	−0.053 (0.228)
Number of pre-teen years in a one-adult family	0.001 (0.020)	−0.005 (0.014)
Number of pre-teen years in a stepfamily	0.014 (0.015)	**0.015** (0.009)
Child has lived with two or more stepparents	0.094 (0.151)	−0.089 (0.084)
Child was born in United States	−0.083 (0.303)	**−0.142** (0.085)
Eldest child	0.068 (0.052)	0.033 (0.030)
Low birth weight	0.023 (0.085)	0.005 (0.054)
Child looks younger than actual age	0.032 (0.067)	0.004 (0.041)
Child looks older than actual age	0.068 (0.063)	**0.100** (0.038)
Child is retarded	0.017 (0.360)	**0.351** (0.188)
Dummy indicating boy-to-girl comparison	**−0.201** (0.068)	**0.079** (0.045)
Dummy indicating girl-to-boy comparison	**0.284** (0.077)	−0.001 (0.045)
Dummy for each combination of ages	Yes	Yes
Number of observations	1,907	1,922

Notes: Except for last three explanators, each variable measured as differential between two siblings. Table reports OLS coefficients, with standard errors in parentheses. Standard errors adjusted for within-family correlation (multiple pairings in some families). Bold font denotes significance at the 10 percent level or higher.

thers. For all but two of the measures, the sign on the between-family estimate indicates a worse outcome for stepchildren than biological children. The within-family estimate has the same sign for all but six of the measures. Repeating the comparisons for stepmother effects, we see that the sign on the between-family estimate indicates a worse outcome for stepchildren than biological children in every case; the within-family estimate has the same sign for 15 of the 28 measures. Looking at the relative sizes of the point estimates before and after differencing, we find that stepparent effects tend to get slightly smaller. If we compute the ratio of each within-family

Table 5A
Stepparent effects measured between families

Measures of child well-being	Stepfather vs intact	Stepmother vs intact	Mother-only vs intact
Measures of parental investment			
Child attends private school	0.001 (0.010)	−0.013 (0.016)	0.004 (0.008)
Child saw dentist in past 12 months	**−0.052** (0.020)	**−0.170** (0.041)	−0.027 (0.019)
Child wears braces on teeth	0.003 (0.011)	**−0.029** (0.015)	0.001 (0.011)
Number of child's extracurricular activities	**−0.113** (0.086)	**−0.460** (0.154)	−0.116 (0.083)
Child shares few activities with (step)father	**0.120** (0.032)	−0.029 (0.039)	
Child shares few activities with (step)mother	**−0.043** (0.019)	**0.189** (0.044)	**−0.061** (0.017)
Outcome measures			
Education measures			
Child's self-reported four-subject GPA	**−0.154** (0.034)	**−0.351** (0.065)	**−0.084** (0.031)
Child ever held back a grade	**0.050** (0.018)	**0.090** (0.035)	**0.046** (0.017)
Child reports getting in trouble at school	**0.037** (0.008)	**0.051** (0.016)	**0.029** (0.009)
Child reports having been suspended from school	**0.083** (0.019)	**0.117** (0.039)	**0.073** (0.018)
Child reports having been expelled from school	0.002 (0.009)	**0.041** (0.025)	**0.016** (0.009)
Measures of risky behavior			
Child reports having had sex	**0.127** (0.019)	**0.113** (0.038)	**0.087** (0.018)
Child reports drinking alcohol away from adults	**0.072** (0.020)	0.050 (0.038)	**0.067** (0.018)
Child reports having used marijuana	**0.124** (0.020)	**0.172** (0.039)	**0.134** (0.018)
Child reports having used "hard" drugs	**0.234** (0.040)	**0.335** (0.079)	**0.233** (0.036)
Child's percentile rank for delinquent behavior	**3.40** (1.20)	**6.55** (2.28)	**3.50** (1.09)

Measures of child's social network			
Number of students naming child as a friend	−0.256 (0.198)	**−0.692** (0.352)	−0.290 (0.185)
Mean GPA of child's friends	**−0.077** (0.032)	**−0.147** (0.064)	**−0.049** (0.030)
Mean number of extracurricular activities of friends	**−0.199** (0.075)	−0.180 (0.145)	**−0.088** (0.068)
High alcohol/tobacco/marijuana use by best friends	**0.037** (0.017)	0.011 (0.031)	0.010 (0.015)
Measures of relationship quality			
"My (step)mother is mostly warm and loving"	−0.015 (0.014)	**−0.178** (0.039)	−0.013 (0.012)
"My (step)father is mostly warm and loving"	**−0.136** (0.031)	−0.037 (0.043)	
"My (step)mother cares for me"	−0.009 (0.015)	**−0.342** (0.039)	−0.016 (0.015)
"My (step)father cares for me"	**−0.134** (0.028)	−0.011 (0.031)	
"I feel close to my (step)mother"	−0.005 (0.014)	**−0.265** (0.041)	−0.011 (0.013)
"I feel close to my (step)father"	**−0.201** (0.032)	0.034 (0.037)	
"Have poor relationship with (step)mother"	0.020 (0.014)	**0.139** (0.039)	**0.022** (0.013)
"Have poor relationship with (step)father"	**0.099** (0.030)	0.010 (0.039)	
"I badly want to leave home"	0.030 (0.016)	0.053 (0.029)	0.021 (0.015)
"Last physical fight was with a family member"	**0.019** (0.011)	**0.039** (0.022)	0.014 (0.010)
Measures of emotional health			
Child contemplated suicide seriously in past year	**0.036** (0.015)	**0.132** (0.035)	0.012 (0.014)
Child often feels depressed	**0.175** (0.063)	0.043 (0.125)	**0.171** (0.059)
Child has poor self-image	**0.038** (0.022)	0.050 (0.042)	**0.056** (0.021)

Note: Table reports the difference in a variable's regression-adjusted mean between two types of families. Standard errors (in parentheses) adjusted for within-family correlation. Bold font indicates statistical significance at the 10 percent level or higher. The appendix gives more detail about the indicators.

Table 5B
Stepparent effects measured within families

Differenced measures of child well-being	Stepfather vs. father	Stepmother vs. mother
Differenced measures of parental investment		
Child attends a private school	0.014 (0.010)	0.005 (0.011)
Child saw dentist in past 12 months	−0.081 (0.120)	−0.186 (0.221)
Child wears braces on teeth	**0.171** (0.064)	0.124 (0.121)
Number of child's extracurricular activities	0.710 (0.585)	**1.58** (0.794)
Child shares few activities with (step)father	0.234 (0.193)	0.119 (0.280)
Child shares few activities with (step)mother	−0.045 (0.109)	**0.485** (0.193)
Differenced outcome measures		
Differenced education measures		
Child's self-reported four-subject GPA	**−0.288** (0.171)	0.005 (0.306)
Child ever held back a grade	−0.079 (0.108)	−0.155 (0.173)
Child reports getting in trouble at school	0.054 (0.058)	0.127 (0.102)
Child reports having been suspended from school	0.023 (0.123)	**−0.649** (0.169)
Child reports having been expelled from school	0.030 (0.039)	0.006 (0.077)
Differenced measures of risky behavior		
Child reports having had sex	0.110 (0.126)	−0.053 (0.228)
Child reports drinking alcohol away from adults	0.012 (0.123)	**0.424** (0.218)
Child reports having used marijuana	0.113 (0.117)	0.129 (0.196)
Child reports having used any of 5 "hard" drugs	−0.056 (0.205)	−0.356 (0.449)
Child's percentile rank for other delinquent behavior	8.08 (8.07)	8.97 (11.7)

Differenced measures of child's social network		
Number of students naming child as a friend	−0.206 (1.71)	1.65 (3.40)
Mean GPA of child's friends	0.312 (0.213)	−0.120 (0.401)
Mean number of extracurricular activities of friends	**−1.19** (0.56)	**−1.75** (0.557)
High alcohol/tobacco/marijuana use by best friends	−0.082 (0.116)	−0.067 (0.224)
Differenced measures of relationship quality		
"My (step)mother is mostly warm and loving"	**0.145** (0.083)	−0.305 (0.162)
"My (step)father is mostly warm and loving"	**−0.374** (0.178)	**−0.627** (0.251)
"My (step)mother cares for me "	0.153 (0.104)	0.140 (0.189)
"My (step)father cares for me"	−0.066 (0.095)	**0.144** (0.089)
"I feel close to my (step)mother"	**0.147** (0.077)	−0.031 (0.198)
"I feel close to my (step)father"	**−0.423** (0.190)	−0.373 (0.263)
"Have unsatisfactory relationship with (step)mother"	−0.045 (0.093)	−0.094 (0.130)
"Have unsatisfactory relationship with (step)father"	**0.290** (0.147)	0.370 (0.257)
"I badly want to leave home"	−0.077 (0.108)	0.019 (0.162)
"Last physical fight was with a family member"	**0.159** (0.083)	0.072 (0.106)
Differenced measures of emotional health		
Child contemplated suicide seriously within past year	0.100 (0.118)	0.130 (0.154)
Child often feels depressed	−0.076 (0.509)	0.173 (0.871)
Child has poor self-image	0.080 (0.147)	−0.300 (0.236)

Notes: Each variable measured as differential between two (half-)siblings. Table reports correlation between outcome differential and any difference between siblings' relationships to an adult in the family. Sample size varies slightly by measure, but is about 2,000, except for indicators concerning fathers or stepfathers specifically, where it is roughly 1,350. Standard errors (in parentheses) adjusted for within-family correlation (multiple pairs in some families). Bold font denotes statistical significance at the 10 percent level or higher. The appendix gives more detail about the indicators.

estimate to its corresponding between-family estimate, the median value of those ratios is 0.74. Finally, it is striking that, despite the smaller samples for the within-family estimates, almost a third of the stepparent effects remain significant.

While these results suggest that selection bias does not fully explain stepchildren's poor outcomes, they do not constitute evidence that stepparents mistreat stepchildren. Of the stepparent effects that persist after differencing, the majority are in the "Relationship quality" category. Poor stepchild-stepparent relationships need not imply differential treatment by stepparents. They could also arise from, say, teenaged stepchildren's own hostility toward stepparents, or notions of loyalty to an absent biological parent.

V. Discussion

In this paper we examine 33 indices of child well-being for evidence of selection bias in the measurement of stepparent effects. Consistent with the results of most studies, we find negative effects for virtually every measure when children are compared across families. When sibling comparisons are used to control for unobserved family factors, the signs of two-thirds of the point estimates are unchanged, and a third of the estimates remain statistically significant. Comparison of the within-family estimates to the between-family estimates suggests that many more of the within-family effects, particularly stepfather effects, would be significant if the sample were larger.

These results stand in contrast to those of Gennetian (1999), Ginther and Pollak (2002), or Hofferth and Anderson (2000), who find little evidence of stepparent effects when they compare half-siblings. Their findings may reflect both the small number of outcomes that they examine and their smaller samples of blended families.

Our results suggest that being a stepchild does lead to unfavorable outcomes. Sample selection may indeed bias the measurement of stepparent effects based on between-family estimates, but we conclude that stepchildren's unfavorable outcomes are not entirely attributable to sample selection. The estimates' sensitivity to the choice of indicator also suggests that studies based on a single indicator of child well-being may be misleading.

Given the proportion of American children who spend time in stepfamilies, stepfamily functioning deserves more investigation. Better studies might explain findings that children derive little or no benefit from having stepfathers. A clearer understanding of the disadvantages faced by stepchildren might shift policymakers' focus from promoting legal marriage to helping parents stay together. For example, perhaps welfare programs, and Transitional Aid to Needy Families (TANF) in particular, should be redesigned to stop privileging unrelated men over fathers. Perhaps "wedfare" and other pro-marriage measures targeted at single mothers should be given less priority, and child-support enforcement more. Perhaps family courts should not view a stepparent as an asset in making custody decisions. Almost certainly, more resources should be devoted to helping stepfamilies function better.

Half-sibling comparisons are not a panacea for measurement problems, and may mismeasure stepparent effects for a number of reasons. Bias is introduced by unobserved differences between siblings that are correlated with family structure. Thus

stepparent effects may be overstated as a result of unobserved heritable paternal traits, for instance, or inadequate controls for a child's prior history. Even if unobserved differences between siblings are not correlated with family structure, they may still reduce the sibling-difference model's explanatory power, as Duncan and Raudenbusch (1998) note.

On balance, however, we view the stepparent effects in this study as conservative estimates. One reason is that the families that permit us to identify the effects—families with half-siblings and an adult who is parent to one child and stepparent to the other—may themselves be a select group. Hofferth and Anderson (2000) suggest that the stepparents who go on to have children with a partner who already has children from a previous relationship are those with better parenting or relationship skills. If this is the case, one might expect more equal treatment of half-siblings.

Our data support the idea that stepparent effects measured in blended families may be smaller than in pure stepfamilies. Looking at all children who live with both biological parents, we find few significant differences between those in intact families and those in blended stepfamilies. By half of our 33 measures, children from intact families fare a bit better, and by the other half, they fare worse. Moreover, only a handful of the differences are statistically significant. By contrast, if we compare stepchildren in blended families with stepchildren in pure stepfamilies, we find that the former do better by 28 of the 33 measures, and for eight of those 28 measures the difference is statistically significant. Stepchildren in blended families do worse than other stepchildren by only five measures, and the difference is significant for only one of those.

Another reason half-sibling comparisons might understate stepparent effects is that a biological parent may try to compensate children for the shortcomings of a stepparent. If this were the case, one might expect children in a stepfather family to report better relationships with their mother, or to participate in more activities with her, than children in an intact family. The signs of the five mother-specific measures in Table 5B are all consistent with this hypothesis, and two of the estimates are statistically significant.

Half-sibling comparisons may also understate stepparent effects if overt parental discrimination between two children living under the same roof violates social norms. If some stepparents respond by stinting neither child, or both, the measured effects will be smaller. Differential treatment may be more apparent in data from countries where discriminatory treatment of siblings is more permissible, or in data from poor countries where parents face much harsher constraints in the intrahousehold allocation of food, education, or health care (see, for instance, Case, Lin, and McLanahan's 1999 study using South African data).

To the extent that differential treatment involves illicit behavior like abuse or neglect, it is bound to be underestimated. There are reasons to suspect that stepchildren are very disproportionately the victims of abuse (Daly and Wilson's 1988 report that children are 70 times as likely to be murdered by a stepfather as by a father is a much more pronounced stepparent effect than any in our data). To the best of our knowledge, however, no large-scale surveys, including Add Health, ask questions about parental neglect or abuse. Reasons for this omission include the fear that parents would forbid participation in the survey, and concern over legal requirements to report known or suspected child abuse. We inject here a plea that,

given the serious impact of neglect and abuse on children's well-being and life outcomes, greater efforts be made to collect such data in the context of large surveys, if only by asking retrospective questions of respondents once they reach adulthood.

Finally, there are three aspects of Add Health's design that are likely to lead one to underestimate negative stepparent effects. First, the survey's school-based design means that there are no dropouts in the initial sample. If, as research suggests, stepchildren are more likely to drop out of high school than children who live with both parents (Garasky 1995, Bogess 1998), then the children with the worst relationships with their stepparents may be the most likely to drop out. Furthermore, youths in juvenile detention are not surveyed, but are disproportionately from stepfamilies (McLanahan and Harper 1998) and can be presumed to have poorer outcomes in general. Second, the 10 percent of respondents who live with neither biological parent[11] are excluded from our analysis because too much family history information is missing. Stepchildren are more likely than other children to leave home because of conflict (Mitchell 1994), and it is reasonable to suppose that the homeleavers, omitted from our sample, had worse-than-average relations with their stepparents. Third, when biological parents separate, which parent houses the child is not random. Teenaged children often have a voice in custody determination, and their preferences are affected by the behavior of parents' new partners. Even if children had no voice in the matter, stepparents' sentiments toward their spouses' children would influence custody arrangements. Hence, children observed in stepparent households are likely to be living with the most successful among potential stepparents.

To sum up, our results cast doubt on the argument that stepchildren's poor outcomes mostly reflect sample selection bias. Whether their poor outcomes are attributable to differential treatment by stepparents, however, remains an open question.

Appendix 1

Definitions of measures of child well-being in AddHealth data

Number of child's extracurricular activities

Number of activities child participates in (from 30-item list). Topcoded at ten, which corresponds to the 99th percentile.

Child shares few activities with (step)mother

Dummy variable coded "Yes" if child participates in two or fewer of ten possible activities with parent during preceding month (true of 23 percent of children).

Child shares few activities with (step)father

Dummy variable coded "Yes" if child participates in one or none of ten possible activities with parent in preceding month (true of 25 percent of children in two-adult families).

11. Even among adolescents 17 and younger, the proportion is 8 percent.

Child's self-reported four-subject GPA

Grade point average can range from 1.0 to 4.0. Child reports letter grades for four subjects (mathematics, science, history or social studies, and language arts).

Child getting in trouble at school

Dummy variable coded "Yes" if child admits to two or more of four possible types of trouble at school (true of 10 percent of children).

Child reports having used "hard" drugs

"Hard" does not include marijuana.

Child's percentile rank for delinquent behavior

Children are asked about 15 types of delinquent behavior (other than use of tobacco, alcohol, or illegal drugs). Each type is scored zero to three (three corresponds to most frequent) and the scores summed. The raw scores are converted into percentiles.

Number of children naming child as friend

Number of fellow students who include child in their own list of friends.

High alcohol/tobacco/marijuana use by best friends

Child is asked how many of three best friends (a) smoke one or more cigarettes per day, (b) drink alcohol at least once a month, or (c) use marijuana at least once a month. Dummy variable coded "Yes" if child's answers sum to six or more (true of 16 percent of children).

"My (step)mother is mostly warm and loving"

Dummy variable coded "Yes" if child marks four or five on a scale of one (strong disagreement) to five (strong agreement) (true of 84 percent of children).

"My (step)father is mostly warm and loving"

Dummy variable coded "Yes" if child marks four or five on a scale of one (strong disagreement) to five (strong agreement) (true of 83 percent of children in two-adult families).

"My (step)mother cares for me"

Dummy variable coded "Yes" if child marks five on a scale of one (strong disagreement) to five (strong agreement) (true of 82 percent of children).

"My (step)father cares for me"

Dummy variable coded "Yes" if child marks four or five on a scale of one (strong disagreement) to five (strong agreement) (true of 82 percent of children in two-adult families).

"I feel close to my (step)mother"

Dummy variable coded "Yes" if child marks four or five on a scale of one (strong disagreement) to five (strong agreement) (true of 83 percent of children).

"I feel close to my (step)father"

Dummy variable coded "Yes" if child marks four or five on a scale of one (strong disagreement) to five (strong agreement) (true of 80 percent of children in two-adult families).

"Have a poor relationship with (step)mother"

Dummy variable coded "Yes" if child marks one, two, or three on a scale of one (very dissatisfied) to five (very satisfied) (true of 16 percent of children).

"Have a poor relationship with (step)father"

Dummy variable coded "Yes" if child marks one, two, or three on a scale of one (very dissatisfied) to five (very satisfied) (true of 18 percent of children in two-adult families).

"I badly want to leave home"

Dummy variable coded "Yes" if child marks four or five on a scale of one (doesn't want to leave home) to five (badly wants to leave) (true of 16 percent of children).

"Child often feels depressed"

Dummy variable coded "Yes" if child marks three, four, or five on a scale of one (rarely feels depressed) to five (always feels depressed) (true of 15 percent of children).

"Child has poor self-image"

Dummy variable coded "Yes" if child is below the 20^{th} percentile for positive self-image. Children are asked about 11 indicators of positive self-image. Each indicator is scored one (strongly disagree) to five (strongly agree) and the scores summed. The raw scores are converted into percentiles.

References

Bogess, Scott. 1998. "Family Structure, Economic Status, and Educational Attainment." *Journal of Population Economics* 11(2):205–22.

Bumpass, Larry, R. Kelly Raley, and James A. Sweet. 1995. "The Changing Character of Stepfamilies: Implications of Cohabitation and Nonmarital Childbearing." *Demography* 32(3):425–36.

Case, Anne, I-Fen Lin, and Sara McLanahan. 1999. "How Hungry Is the Selfish Gene?" 2000 *Economic Journal* 110(466): 781-804.

———. 1999. "Household Resource Allocation in Stepfamilies: Darwin Reflects on the Plight of Cinderella." *American Economic Review* 89(2):234–38.

Cherlin, Andrew J., Frank F. Furstenberg, Jr., P. Lindsay Chase-Lansdale, Kathleen E. Kiernan, Philip K. Robins, Donna Ruane Morrison, and Julien Teitler. 1991. "Longitudinal Studies of Effects of Divorce on Children in Great Britain and the United States." *Science* 252(5011):1386–9.

Daly, Martin, and Margo Wilson. 1985. "Child Abuse and Other Dangers of Not Living with Both Parents." *Ethology and Sociobiology* 6:197–210.

———. 1988. *Homicide*. New York: A. de Gruyter.

———. 1996. "Violence Against Stepchildren." *Current Directions in Psychological Science* 5(3):77–81.

Dawson, Deborah A. 1991. "Family Structure and Children's Health and Well-Being: Data from the 1988 National Health Interview Survey on Child Health." *Journal of Marriage and the Family* 53(3):573–84.

Duncan, Greg J., and Stephen W. Raudenbush. 1998. "Neighborhoods and Adolescent Development: How Can We Determine the Links?" Working Paper 12-3, Joint Center for Poverty Research. Chicago: Northwestern University.

Epstein, A. M. 1990. "The Outcomes Movement—Will It Get Us Where We Want to Go?" *New England Journal of Medicine* 323(4):266–70.

Evenhouse, Eirik, and Siobhan Reilly. 1997. "Divorce Laws and Divorce Rates: Evidence From Panel Data." Mills College. Unpublished.

———. 2003. "Father Taxes and Father Absence." Institute for Health Policy Studies. San Francisco: University of California. Unpublished.

Garasky, Steven. 1995. "The Effects of Family Structure on Educational Attainment." *American Journal of Economics and Sociology* 54(1):89–105.

Gennetian, Lisa. 1999. "One or Two Parents? Half or Step Siblings? The Effect of Family Composition on Cognitive Outcomes for Young Children." Manpower Demonstration Research Corporation. Unpublished.

Geronimus, Arline T., and Sanders Korenman. 1992. "The Socioeconomic Consequences of Teen Childbearing Reconsidered." *Quarterly Journal of Economics* 107(4):1187–214.

Ginther, Donna A., and Robert A. Pollak. 2002. "Does Family Structure Affect Children's Educational Outcomes?" University of Kansas, Department of Economics. Unpublished.

Griliches, Zvi. 1979. "Sibling Models and Data in Economics: Beginning of a Survey." *Journal of Political Economy* 87(5):S37–S64.

Hofferth, Sandra L., and K. G. Anderson. 2000. "Biological and Stepfather Investment in Children." Institute for Social Research. Ann Arbor: University of Michigan. Unpublished.

Lundberg, Shelley, and Elaina Rose. 2002. "The Effects of Sons and Daughters on Men's Labor Supply and Wages." *Review of Economics and Statistics* 84(2):251–68.

McLanahan, Sara. 1998. "Father Absence and the Welfare of Children." In *Coping with Divorce, Single Parenting, and Remarriage: A Risk and Resiliency Perspective*, ed. M. Hetherington, 117–45. Mahwah, N.J.: L. Erlbaum Associates, Inc.

McLanahan, Sara, and Cynthia Harper. 1998. "Father Absence and Youth Incarceration." Princeton University. Unpublished.
McLanahan, Sara, and Gary Sandefur. 1994. *Growing up with a Single Parent: What Hurts, What Helps.* Cambridge, Mass.: Harvard University Press
Mitchell, Barbara. 1994. "Family Structure and Leaving the Nest: A Social Resource Perspective." *Sociological Perspectives* 37(4):651–71.
Moffitt, Robert A., Robert T. Reville, and Anne E. Winkler. 1998. "Beyond Single Mothers: Cohabitation, Marriage, and the U.S. Welfare System." *Demography* 35(3):259–78.
Painter, Gary, and David L. Levine. 2000. "Family Structure and Youths' Outcomes: Which Correlations Are Causal?" *Journal of Human Resources* 15(3):524–9.
Sandefur, Gary, and Thomas Wells. 1997. "Using Siblings to Investigate the Effects of Family Structure on Educational Attainment." Discussion Paper 1144–97, Institute for Research on Poverty. Madison: University of Wisconsin.