The Intergenerational Transmission of Welfare Receipt: A Nonparametric Bounds Analysis
Author(s): John V. Pepper
Source: *The Review of Economics and Statistics*, Vol. 82, No. 3 (Aug., 2000), pp. 472-488
Published by: The MIT Press
Stable URL: http://www.jstor.org/stable/2646807
Accessed: 13/01/2010 14:31

THE INTERGENERATIONAL TRANSMISSION OF WELFARE RECEIPT: A NONPARAMETRIC BOUNDS ANALYSIS

John V. Pepper*

Abstract—Using a nonparametric bounding method and data from the Panel Study of Income Dynamics, I examine the effect that growing up in a household that receives Aid to Families with Dependent Children (AFDC) has on welfare participation as a young adult. In light of the ambiguities created by the selection problem, a number of alternative assumptions and estimates are presented. While the data alone cannot be conclusive, the results generally strengthen the evidence that being exposed to AFDC as a child increases both the probability and the expected duration of future welfare participation.

I. Introduction

FOR more than a century, scholars and political leaders in the United States have argued that charity and public aid causes a culture of dependency that is transmitted across generations (Katz, 1986). In 1888, for instance, the founder of the Indianapolis Charity Organization Society, Oscar McCulloch, professed that growing up in a household that received public welfare creates an "irresistible tendency" towards pauperism. Richard Dugdale's (1877) landmark study of the Jukes family in New York drew similar conclusions. Over a century later, President George Bush voiced related concerns when he suggested that welfare is "passed from generation to generation like a legacy" (1992).

Recently, researchers have found empirical evidence that supports this rhetoric. Growing up in a household that receives Aid to Families with Dependent Children (AFDC) appears to increase the probability that a child will receive AFDC as an adult. These findings have disturbing implications: apparently, participation in welfare today induces participation by future generations.

Given current efforts to reform the welfare system by both state and federal governments, these studies are particularly relevant. However, despite the existing literature, the intergenerational effects of growing up in an AFDC household remain uncertain. Although receiving AFDC today may induce future generations to participate, unobserved factors may also exist that jointly influence whether parents and children receive welfare. For instance, parents' human-capital characteristics, attitudes towards work and family, addictions, and emotional well-being may all affect both the parents' and child's propensity to receive AFDC. Thus, any observed relationships between the welfare participation behavior of parents and children could be spurious. A selection problem results from the fact that the data alone cannot reveal how a child growing up in an AFDC household would have behaved if the child were to have grown up in a non-AFDC household.

Using intergenerational data from the Panel Study of Income Dynamics, I reassess the effect that growing up in an AFDC household has on future welfare participation. Focusing primarily on black daughters who grew up in single-parent households, two contributions are made to the existing literature. First, this study investigates the intergenerational effects in light of the uncertainties created by the selection problem; second, by observing the behavior of parents and their daughters for multiple years, this study explores how the daughters' time spent receiving AFDC varies with the length of time that parents received welfare.

Although the distinction between participation and length of participation has not been entirely ignored in the literature,[1] researchers have focused on the effects that growing up in a household that received any AFDC has on the probability of future welfare participation. The length of receipt, however, has important policy implications. Clearly, political and academic attention focuses on dependence rather than occasional participation. Furthermore, the design of counteracting reforms may benefit from this distinction. As Antel (1992) suggests, "If only long-term exposure implies intergenerational transfer, then any counteracting policy need only be directed at young women from chronic welfare homes" (p. 473).

Thus, to better understand the intergenerational dynamics, I use five-year observation windows on both parents and daughters to examine the effects that growing up in a household that received AFDC for a specified duration during the early teenage years has on the length of receipt as a young adult. Even focusing on five-year intervals, however, does not allow one to draw inferences about lifetime relationships. Most AFDC spells last less than two years, and incomplete welfare histories may produce inconsistent estimates of lifetime intergenerational relationships (Gottschalk, 1992; Wolfe, Haveman, Ginther, & An, 1996). It seems likely, for instance, that the intergenerational effects

Received for publication February 3, 1997. Revision accepted for publication July 7, 1999.

* University of Virginia.

I wish to thank Arthur Goldberger, Robert Haveman, John Kennan, Charles Manski, Andrew Reschovsky, John Karl Scholz, Steven Ziliak, two anonymous referees, and workshop participants at Tufts University, the University of Kentucky, the University of Virginia, the University of Wisconsin, and Virginia Tech, for their many helpful comments. Also, I wish to thank the Wisconsin Alumni Research Foundation for financial support.

[1] Gottschalk (1992) and McLanahan (1988) estimate the relationship between the length of time parents receive AFDC and the probability that their children participate; Antel (1992) examines the relationship between parental participation and the expected length of time daughters receive AFDC. Pepper (1995) examines the association between the time parents participate and the expected duration of a daughter's receipt, finding that daughters exposed to AFDC as children spend on average more time receiving aid than those growing up in non-AFDC homes. Unlike the present analysis, however, Pepper does not attempt to identify the effects of growing up in an AFDC household, but instead merely examines the statistical association between the lengths of time parents and their children receive AFDC.

The Review of Economics and Statistics, August 2000, 82(3): 472–488

differ by when the daughter was exposed to parental receipt: infants and teenagers may be affected differently.

After the data are described in section II, section III formally defines the empirical questions and describes the selection problem. Then, sections IV through VI provide alternative estimates of the effect of growing up in a household that received AFDC on the distribution of time that a daughter receives aid. Each set of estimates relies on different a priori assumptions to address the selection problem.

Using a conventional parametric model coupled with the exogenous selection assumption that unobserved factors affecting parents and children are unrelated, section IV presents estimates that are consistent with the past literature: growing up in a household that receives any AFDC increases the expected duration of receipt. These point estimates, however, rely on strong and questionable assumptions.

In contrast, sections V and VI apply a nonparametric bounding method developed by Manski (1995) to examine what can be learned under weaker assumptions. Section V reveals that, in the absence of prior information about the selection problem, the data reveal little about the intergenerational effects of growing up in a household that received AFDC for a specified duration. Rather than taking the extreme positions of either making no assumptions (section V) or making the strong assumptions necessary for identification (section IV), section VI explores a middle ground. Here, four sets of predictions are made using assumptions motivated by economic theory or empirical convention.

Section VI begins by assuming that being exposed to AFDC as a teenager never decreases the time that a daughter receives AFDC as a young adult. Although this assumption does not generally enable identification, useful nonparametric bounds are derived. These estimated bounds not only rule out extreme contentions, but also reduce the range of possible values of the intergenerational effects by as much as 75%.

The identifying power of the traditional instrumental variable or exclusion restriction assumption is then explored. In particular, I assume that the local unemployment rate faced by parents is not related to the length of time that a daughter receives AFDC as an adult, but does affect the duration of her parents' receipt. Once again, I find that invoking this assumption is useful but that it does not necessarily enable identification. In fact, the estimates derived using this assumption do not even determine the sign of the effects. Still, under this relatively weak instrumental-variable assumption, the results suggest that, for some cohorts, the effects of being exposed to AFDC as a teenager on the length of receipt as a young adult are either negligible or substantially positive.

The final two sets of estimates identify the sign of the treatment effect. The first set of estimates rely on both the instrumental-variable assumption and the assumption that growing up in an AFDC household never decreases the time a daughter receives AFDC as a young adult, while the second set utilize an exogenous selection assumption. These estimates not only confirm the previous findings that growing up in an AFDC household increases the probability of future receipt, but they also suggest that being exposed to AFDC as a child increases the likelihood of becoming a chronic recipient.

Section VII compares the various estimates and finds several inconsistencies. In particular, the point estimates derived under the exogenous selection assumption fall outside the range of a number of the alternatives. Despite these discrepancies, the qualitative conclusions remain. Section VIII concludes by noting that, despite the ambiguity created by the selection problem, all of the alternative estimates presented in this paper are consistent with the inference that growing up in a household that receives AFDC increases the length of time a daughter will receive aid as an adult.

II. Data: The Panel Study of Income Dynamics

The primary data source used in this analysis is the 1989 Panel Study of Income Dynamics (PSID). In 1968, this longitudinal survey began interviewing a national sample of nearly 4,800 households that overrepresented the non-elderly, low-income minority population. Each year since, the heads of these 4,800 families and any other families formed by members or descendants of the original 1968 sample (split-off families) have been surveyed. Thus, the 1989 survey, which includes nearly 7,000 families and 27,000 individuals, contains socioeconomic data on young women and their families of origin (Morgan et al., 1992).

In the sample used for this analysis, I include 1,205 PSID daughters who were between 24 and 33 in 1989, and for whom information exists from 1984 through 1988 and ages 12 to 16.[2] Notice that for each respondent there are two five-year observation windows: one during the early teenage years, and one during the mid-1980s when these daughters were in their twenties and early thirties. Thus, these data can be used to draw inferences about the effects of being exposed to AFDC as a teenager on the duration of receipt as a young adult.

Nearly 1,300 observations are excluded due to attrition or nonresponse. In this analysis, I maintain the assumption that attrition is exogenous, or unrelated to the welfare participation behavior of parents and their children. There is some empirical support for this assumption. Specifically, Fitzgerald, Gottschalk, and Moffitt (1998) find that, after accounting for observed factors, attrition in the PSID does not appear to affect estimated intergenerational welfare participation relationships.[3] If, however, unobserved factors jointly

[2] This sample includes respondents from both the representative Survey Research Center subsample and the stratified Survey of Economic Opportunity subsample.

[3] To account for factors associated with attrition, Fitzgerald, Gottschalk, and Moffitt (1998) use both observed covariates and PSID sampling weights that provide a general adjustment for attrition based on observable measures. The results imply that attrition has little impact on estimated intergenerational welfare participation parameters. In particular, the bias

affect both attrition and intergenerational welfare participation, then these data may provide inconsistent estimates of the population relationships of interest.

Each observation contains information on the time that the respondent and the respondent's family of origin received AFDC. For the PSID, households are asked about the types and amounts of government assistance that they received during the prior year. From 1967 to 1982, AFDC receipt was measured on an annual basis, and, from 1983 onward, the annual survey began asking households to report information on a monthly basis. Thus, while the daughter's participation information is measured monthly, parental receipt is measured annually.

Although the monthly measures are likely to provide more information on the intensity of participation over time, an important caveat must be considered when using the monthly as opposed to annual participation indicators. Because the monthly measurement period differs from the annual survey period, a disproportionate number of monthly transitions are observed at the seams of the survey period (Hill, 1992). To the extent that these monthly data tend to be inappropriately reported in yearly increments, the outcome variable merely measures the years of participation rather than the intensity of participation during the five-year observation period.

The months that a respondent received AFDC are measured during the five-year period from 1984 through 1988. For more than 80% of the observations in the sample, this information is directly recorded. If a respondent is a head of household or a spouse in year t (that is, a split-off daughter), then the PSID reports the number of months she received AFDC in year $t - 1$.

However, for daughters living with their parents (a non-split-off daughter) and receiving more than one source of transfer income, the panel does not provide disaggregated details about the types or amounts of aid. Of the 1,205 respondents, 266, or 22%, lived with their families of origin for at least one of the five observed years. In particular, 101 daughters lived with their families for all five years, 23 for four years, 36 for three years, 52 for two years, and 54 for one year. Thus, approximately 14% of the monthly welfare receipt information for the 1,205 respondents is imputed.

Consistent with previous studies, I impute the time a non-split-off daughter receives AFDC using a number of variables in the PSID (An, Haveman, & Wolfe, 1993; Haveman & Wolfe, 1994; Pepper, 1995).[4] A non-split-off daughter is first classified as receiving AFDC in year t if she has a child, has taxable income below the poverty line, and is recorded as either having received AFDC income only or multiple sources of transfer income. Then, the number of months that she received AFDC is imputed. In particular, the average number of months (conditional on the number of years) that split-off daughters received AFDC is used to predict the number of months that non-split-off daughters participated. So, for example, split-off daughters that received aid in one year spent an average of seven months receiving AFDC. Thus, if a non-split-off daughter received AFDC for one of the five observed years, she is assumed to have spent seven months on AFDC. Alternatively, if a non-split-off daughter received AFDC for two, three, four or five years, she is assumed to have spent, respectively, 9, 10, 10.5, or 11.5 months per year receiving aid.

The years that each respondent's family received AFDC are measured over the five-year period when the daughter was aged twelve to sixteen. This annual measure may overstate the time a respondent was "exposed" to the welfare system (Blank, 1989). For example, a family that received AFDC for one month each year while the respondent was between twelve and sixteen is recorded as receiving aid for five years, while a family that received AFDC for five months in a single year is recorded as receiving aid for one year. The monthly data from split-off daughters, however, suggest that individuals who receive AFDC for many years also receive AFDC for many months per year.

In addition to measuring the time that a family received AFDC, the sample contains measures of the five-year-average annual family income and whether the daughter lived with a single parent during any of the five observed years. The family-income variable, which is inflated using the CPI to reflect 1988 prices, measures all forms of earned and unearned income. Variables also measure the number of children in the household when the respondent was fourteen, and the race of the respondent, which I code as 1 if the respondent is black and 0 if white. Other races are excluded from the sample, as they are not well represented in the PSID.[5] Finally, to account for labor market conditions faced by the parents, the sample includes a measure of the local unemployment rate when the daughter was twelve, which I aggregate into five mutually exclusive and exhaustive categories ranging from less than 2% to greater than 10%.

The empirical distribution in the sample by selected covariates is presented in table 1. In total, 286 (24%) of the respondents grew up in households that received AFDC, with the majority being black daughters who grew up in single-parent households. Because AFDC is primarily intended for and used by single-parent households, a relatively small fraction of respondents grew up in intact (married)

estimates resulting from attrition are neither substantial nor statistically different than zero at the 10%-significance level. Without weights, they again find that the estimated bias due to attrition is inconsequential, although in certain cases statistically different than zero. In this analysis, I do not use the PSID weights to adjust the estimates for attrition.

[4] An alternative method circumvents the issue by focusing on children who have established their own households. (See, for example, McLanahan (1988).) This approach has the obvious shortcoming of excluding the potentially important group of young mothers who live with their parents. The primary conclusions of this paper are not sensitive to the method used to measure the AFDC status of daughters.

[5] Hispanics are included in the sample. Beginning in the 1985 wave of the PSID, Hispanics' ethnicity was coded as a separate variable from race. Prior to 1985, Hispanics are coded as a separate race. For the data used in this analysis, I use the post-1985 surveys to determine the race of the respondent.

TABLE 1.—NUMBER OF RESPONDENTS BY RACE, NUMBER OF CHILDREN IN HOUSEHOLD WHEN AGED 14, AND THE MARITAL AND AFDC STATUS OF THE PARENTS WHEN AGED 12–16

Family Background Characteristics	Number of Daughters			Percentage of Daughters Whose Families Received AFDC
	AFDC	No AFDC	Total	
White				
Single				
Four or more children	14	19	33	42%
Three or lower children	5	59	64	8%
Married				
Four or more children	7	180	187	4%
Three or fewer children	5	305	310	2%
subtotal	31	563	594	5%
Black				
Single				
Four or more children	136	49	185	74%
Three or fewer children	65	60	125	52%
Married				
Four or more children	36	137	173	21%
Three or fewer children	18	110	128	14%
subtotal	255	356	611	42%
Total	286	919	1205	24%

Note: *AFDC* indicates that the family received some AFDC when the daughter was aged twelve to sixteen. *Single* indicates that the daughter lived with a single parent during at least one of the five observed years.
Married indicates that the daughter lived in an intact family all five years.

households that received AFDC. Furthermore, because the PSID overrepresents the low-income minority population, a relatively large fraction of the black daughters in the sample grew up in AFDC households. Of the 310 black daughters who grew up in single-parent households, 201 (65%) received some AFDC between the ages of twelve to sixteen. In contrast, only 19 (20%) of the white daughters' families in the same cohort received AFDC.[6]

For those cohorts where the sample contains small numbers of respondents who grew up in AFDC households, only limited inferences can be made about intergenerational welfare behavior. In particular, the estimated intergenerational effects are either imprecise or require strong parametric assumptions. Thus, the remaining discussion focuses primarily on the 310 black respondents who grew up in single-parent households. The one exception is in section IV, which uses all 1,205 observations to estimate a parametric model. In this model, as in much of the literature, both race and family structure are accounted for using single indicator variables. With these data, models with interactions between race, family structure, and welfare participation cannot be precisely estimated.

III. The Empirical Questions and the Selection Problem

Studies examining the intergenerational transmission of welfare receipt may draw distorted conclusions due to the selection problem. This identification problem results from the fact that one cannot observe what the welfare status of AFDC daughters would be if they had grown up in non-AFDC households. Likewise, we cannot observe the participation status of a non-AFDC daughter if she were to have been an AFDC daughter. Without imposing prior and unverifiable assumptions, the effect of growing up in a welfare household cannot be identified.

To formally address the selection problem, I begin by distinguishing among several welfare participation variables. In particular, $Y(t)$ is the outcome (the number of months on AFDC) of a daughter who grew up in an household that received AFDC during t years. In this analysis, I evaluate $Y(t)$ at t equal to zero years, one to two years, three to four years, and five years. Thus, each observation includes four outcome variables: $Y(0)$, $Y(1\text{–}2)$, $Y(3\text{–}4)$, and $Y(5)$.

The purpose of this study is to compare the distribution of $Y(t)$ across different values of t, given observed covariates X. Here, covariates are merely used to define subpopulations of interest, not to control for spurious effects.[7] It is important to recognize that this use of covariates to define subpopulations of interest differs from the conventional use of covariates to account for nonrandom treatment assignment (Manski & Nagin, 1998). Observed covariates are typically motivated as a means of controlling for the unobserved factors associated with a parent's decision to receive welfare. In that framework, researchers attempt to "correctly" choose a set of covariates or "control variables" such that the exogenous selection assumption applies. If certain variables are omitted, the estimated intergenerational effects will be biased. Inevitably, however, these analyses are controversial with much debate about whether or not the researcher omitted "important" explanatory variables. After all, the data cannot resolve the selection problem.

In contrast, I am interested in predicting the intergenerational effects of growing up in an AFDC household for different subpopulations. This problem is well defined regardless of how the subpopulations are specified. In this framework, there is no correct set of control variables. I focus on black daughters who grew up in low-income, single-parent households of varying sizes. For each of these subpopulations, this paper examines the implications of a number of alternative assumptions for identifying the intergenerational effects of growing up in a household that received welfare.

Traditionally, the effect of growing up in a household that received AFDC for t' as opposed to t periods is defined as the difference in the respective conditional means. That is,

$$T_1(t', t|X) = E[Y(t')|X] - E[Y(t)|X]. \quad (1)$$

[6] Comparable estimates from the U.S. Bureau of the Census (1990) suggest that, in 1990, approximately 46% of black children aged six to seventeen residing in single-parent households received means-tested cash assistance (such as AFDC), while the analogous figure for white children is 26%. In 1977, 37% of black and 13% of white single-parent households received public assistance (U.S. Bureau of the Census, 1978).

[7] See Gottshalk (1996) for a discussion of spurious relationships in intergenerational welfare participation analyses. See Manski and Nagin (1998) for a general discussion on the use of covariates.

However, to the extent that one is interested in welfare dependence or long-term receipt, this measure is inadequate. Suppose that the treatment effect defined in equation (1) is positive for all values of t' greater than t. Hence, the expected time a daughter receives AFDC from 1984 to 1988 increases with the length of time that her family received aid. This finding has many possible explanations. It may be that growing up in an AFDC household increases the probability of participation, but decreases the probability of participating for long periods. Alternatively, growing up in an AFDC household may have a negative effect on the probability of participation, but a positive effect on the probability of receiving AFDC for many years.

To better understand the dynamics of this intergenerational process, other features of the distribution of $Y(t)$ should be examined. Ideally, the entire distribution of outcomes would be identified for each value of t. However, because in the absence of strong assumptions there does not generally exist a comprehensive method of summarizing the conditional distribution of $Y(t)$, my analysis focuses on three measures. I examine the expected duration of receipt (equation (1)), the probability of receiving any AFDC (that is, $P[Y(t) > 0|X]$), and the probability of receiving AFDC for more than two years (that is, $P[Y(t) > 24 \text{ months}|X]$). Showing both a central tendency and the tails of the distribution, these three measures better characterize the intergenerational effects of growing up in a household that received AFDC.

To compare these other features of the distribution of $Y(t)$ across different values of t, equation (1) can be modified to read

$$T_2(t', t|X) = P[Y(t') > 0|X] - P[Y(t) > 0|X], \tag{2}$$

or,

$$T_3(t', t|X) = P[Y(t') > 24|X] - P[Y(t) > 24|X]. \tag{3}$$

$T_2(t', t|X)$ in equation (2) measures the effect of growing up in a household that received AFDC for t' as opposed to t periods on the probability of receiving AFDC, while $T_3(t', t|X)$ in equation (3) measures the effect on the probability of receiving AFDC for over two years. Notice that the probabilities in equation (2) and (3) must lie between [0, 1]. Thus, the effect of growing up in an AFDC household for t' versus t years on the probability of receiving AFDC for more than 0 or 24 months must fall in the range of $[-1, 1]$. Likewise, because $Y(t)$ lies between [0, 60] months, the effect on the expected duration of receipt (that is, $T_1(t', t|X)$) falls in the range of $[-60, 60]$ months.

The treatment effects defined in equation (1) through (3), however, cannot be identified by the data alone as the outcome $Y(t)$ is observed only if a respondent grew up in a household that received AFDC during t years. Recall that, for this analysis, I define four outcome variables: $Y(0)$, $Y(1\text{–}2)$, $Y(3\text{–}4)$, and $Y(5)$. Thus, for a daughter who grew up in a household that did not receive AFDC, $Y(0)$ is observed but $Y(5)$ is a latent variable. Similarly, for daughters growing up in a household that received AFDC during five years, $Y(5)$ is observed but $Y(0)$ is latent. In fact, for each respondent, one of these four variables is observed or selected, and the remaining three are unobserved or censored. This identification problem is highlighted using the Law of Iterated Expectations which shows that

$$\begin{aligned} E[Y(t)|X] = {} & E[Y(t)|X, Z = t] * P[Z = t|X] \\ & + E[Y(t)|X, Z \neq t] * P[Z \neq t|X], \end{aligned} \tag{4}$$

where Z is the actual number of years a respondent's family received AFDC. Because each observation in the sample contains information on $Y(Z)$, Z, and X, the sampling process identifies the selection probability $P[Z = t|X]$, the censoring probability $P[Z \neq t|X]$, and expectation of outcomes conditional on the outcome being observed $E[Y(t)|X, Z = t]$. Thus, the sampling process reveals every term in the right side of equation (4) except $E[Y(t)|X, \ Z \neq t]$, the expectation of outcomes conditional on the outcome variable being censored. In the absence of prior information restricting the distributions of $Y(t)$ and Z, observations with $Z \neq t$ reveal nothing about the latent outcome $Y(t)$. Thus, the sampling process does not identify the distribution of outcomes that would be observed if all households received AFDC for t periods.

IV. One Extreme: Estimates Assuming a Tobit Model with Exogenous Selection

To identify the intergenerational effects of growing up in an AFDC household, researchers rely on strong prior information restricting the distributions of observed and unobserved outcomes. Previous studies typically combine a parametric model with the assumption that parental receipt is exogenous. (See, for example, McLanahan (1988), Gottschalk (1992), An et al. (1993), and Pepper (1995).) Using these assumptions along with intergenerational data, researchers have found consistent evidence that growing up in an AFDC household increases the probability that a daughter receives AFDC as a young adult. This section presents similar estimates using a Tobit model with the exogenous selection assumption.

The model begins by assuming that the number of months a daughter receives AFDC if her family participated for t years equals

$$Y(t) = \begin{cases} 0 & \text{if } X'B + \theta_t + \epsilon \leq 0 \\ 60 & \text{if } X'B + \theta_t + \epsilon \geq 60 \\ X'B + \theta_t + \epsilon & \text{otherwise.} \end{cases} \tag{5}$$

Here, the vector B and the scalar θ_t are unknown parameters, where θ_t allows the intercept to differ with the time a family received AFDC, t. The coefficient vector B includes an intercept term, and θ_0 is set equal to 0. In addition to a constant, the vector of observed covariates X includes a variable that indicates if the daughter is black, and a number of variables that characterize the daughter's socioeconomic status when aged twelve to sixteen.[8] Factors that affect the length of participation yet are unobserved by the researcher are included in the random variable, ϵ. In the Tobit model, the distribution of ϵ given X is assumed to be normal with mean zero and unknown variance, σ^2. These assumptions imply that

$$E[Y(t)|X] = 60(1 - \Phi(u)) + (\Phi(u) - \Phi(l)) \times (X'B + \theta_t) + \sigma(\phi(l) - \phi(u)) \quad (6)$$

where,

$\Phi(\cdot)$ is the standard normal CDF,
$\phi(\cdot)$ is the standard normal PDF,
$u = (60 - (X'B + \theta_t))/\sigma$ and $l = (-(X'B + \theta_t))/\sigma$.

Although equation (6) provides strong parametric information about the conditional expectation of the latent variable $Y(t)$, the parameters B and θ_t cannot be identified without additional assumptions. After all, the data cannot reveal $Y(t)$ when $Z \neq t$. The most common identifying assumption used in the literature is that parental receipt is exogenous (Moffitt, 1992).[9] That is, unobserved factors affecting the welfare receipt of parents and their children are unrelated. Formally, the exogenous selection assumption implies that

$$E[Y(t)|X] = E[Y(t)|X, Z]. \quad (7)$$

Because $E[Y(t)|X, Z = t]$ is identified by the sampling process, equations (1) through (4) and (6) are also identified given exogenous selection.[10] In particular, equations (6) and (7) imply that

TABLE 2.—COEFFICIENT ESTIMATES AND THE STANDARD ERRORS GIVEN THE TOBIT MODEL WITH EXOGENOUS SELECTION $Y(t)$ = MONTHS (OUT OF SIXTY) THAT A DAUGHTER RECEIVED AFDC FROM 1984 TO 1988

Variables	Estimates	S.E.
Coefficients		
Constant	−82.1	11.4
$\theta_{\{1,2\}}$: Family received AFDC in one to two years	28.5	7.6
$\theta_{\{3,4\}}$: Family received AFDC in three to four years	28.2	9.6
$\theta_{\{5\}}$: Family received AFDC in five years	30.9	10.4
Average annual income (in thousands)	−0.44	0.18
Number of children, when age fourteen	3.3	1.4
Proportion of years living with one parent	7.1	7.2
Race (black = 1)	35.1	6.7
Sigma	59.9	3.8

Note: These estimates are for the entire population of daughters and thus are computed using all 1205 observations. See the text and table 1 for further description.
S.E. = standard error.

$$E[Y(t)|X, Z = t) = 60(1 - \Phi(u)) + (\Phi(u) - \Phi(l)) \times (X'B + \theta_t) + \sigma(\phi(l) - \phi(u)). \quad (8)$$

Maximum-likelihood estimation is used to derive consistent estimates of B, θ_t and σ. Recall that the stratified sampling scheme used to create the PSID overrepresents the non-elderly, low-income minority population. Because throughout the paper I condition the analysis on the age, race, and family income of the respondent, this stratification scheme is effectively exogenous. Thus, I report unweighted estimators.

The coefficient estimates, along with the asymptotic standard errors, are presented in table 2. These estimates are similar to those found in the previous literature. The expected time that a daughter receives AFDC decreases with family income and increases with the number of children in the household. The proportion of time spent in a single-parent household also appears to have a positive effect, although the estimate is statistically insignificant. Also, the expected time receiving AFDC is higher for black daughters than it is for whites.

Most importantly, the estimates of θ_t, which are provided for $t = 1$ to 2 years, 3 to 4 years, and 5 years, suggest that growing up in a household that receives any AFDC has a large positive effect on the length of time a daughter receives aid. Furthermore, this effect does not differ with the length of time a family received AFDC, t. Growing up in a household that received AFDC for one to two years appears to have the same positive effect on the expected duration of a daughter's receipt as growing up in a household that received AFDC for five years. A test of the null that $\theta_{\{1,2\}} = \theta_{\{3,4\}} = \theta_{\{5\}}$ supports this conclusion. Under the null hypothesis of no duration effects, the likelihood ratio statistic equals 0.07, substantially less than the conventional 10% critical value of 4.61.

Because the conditional expectation function in equation (8) is nonlinear in the parameters, the coefficient estimates of θ_t presented in table 2 do not directly translate into estimates of the intergenerational effects defined in equation (1) through (3). Estimated intergenerational effects given this Tobit model specification are presented in table 8 and discussed in section VII.

[8] Consistent with much of the past literature, the model specification does not allow for general interactions between the race of the respondent, the AFDC participation indicators, and the other covariates. Although some researchers (see Gottschalk (1992, 1996) and McLanahan (1988)) have specified separate models for black and white cohorts, the convention in the literature has been to account for racial differences with a single indicator variable. (See, for example, Antel (1992), An et al. (1993), Zimmerman and Levine (1993), and Pepper (1995).) These restricted models are supported by the data. In fact, using a classical hypothesis test, An et al. find no evidence of structural differences between the races. Likewise, I cannot reject the null hypothesis that coefficients on race interaction terms are all zero at the 1%-significance level. While this test might suggest that there are not important racial differences in the intergenerational transmission of welfare receipt, they also reflect the limitations of the data. Given the small number of white respondents who were exposed to AFDC as children, the data used in these intergenerational analyses are unable to precisely estimate more-general models. Estimates from a model with race interaction terms are available from the author.

[9] Alternatively, researchers have relied on exclusion restrictions (see section VI) coupled with assumptions on distribution of the unobserved factors affecting the welfare participation of parents and their daughters. (See, for example, Antel (1992), Levine and Zimmerman (1996), and Gottschalk (1996).) Using these assumptions, Antel and Gottschalk find that the intergenerational effects generally remain strong, while Levine and Zimmerman find that they tend to disappear. These empirical results, however, are sensitive to both the distributional (Arabmazar & Schmidt, 1982; Goldberger, 1983) and exclusion restriction assumptions.

[10] Nonparametric estimates of equation (7) are reported in section VI.

V. Another Extreme: Estimates With No Prior Information

The Tobit model estimates suggest that growing up in an AFDC household increases the expected time that a daughter receives aid. These estimates, however, rely on strong unverifiable assumptions. In particular, this Tobit model restricts the functional relationships between the observed covariates X and the outcome variable $Y(t)$ (equation (5)), confines the conditional distribution of ϵ given X to be normal with mean zero and a homoskedastic variance, and assumes that parental receipt is exogenous (equation (7)). If any of these assumptions are incorrect (that is, if the parametric model is misspecified or if there exist unobserved factors that jointly determine the AFDC participation of parents and children), then these results should not support the conclusion that growing up in an AFDC household affects the distribution of time that a child receives AFDC.

Rather than relying on these strong identifying assumptions, a logical first step in examining intergenerational welfare participation is to ask what can be learned in the absence of any assumptions that are invoked to address the selection problem. Recall that, in the absence of prior information, equation (1) through (4) are not identified. A selection problem occurs because the sampling process fails to reveal $E[Y(t)|X, Z \neq t]$, the expectation of outcomes conditional on the outcome being unobserved or censored.

However, the data may still reveal information about the distribution of the latent variables, $Y(t)$. For example, because the unidentified conditional expectation $E[Y(t)|X, Z \neq t]$ lies between 0 and 60, bounds can be placed on the possible values of $E[Y(t)|X]$. A sharp upper bound is found by setting $E[Y(t)|X, Z \neq t] = 60$ in equation (4), while the lower bound sets this unobserved expectation equal to 0. Thus, the difference between these bounds equals $P[Z \neq t|X] * 60$. Without data, the range of possible values is [0, 60] months. Similarly, because for any constant r the unidentified conditional probability $P[Y(t) > r|X, Z \neq t]$ lies between 0 and 1, bounds can be placed around the possible values of $P[Y(t) > 0|X]$ and $P[Y(t) > 24|X]$. The width of these bounds equals the censoring probability, $P[Z \neq t|X]$.

These no-information bounds on $E[Y(t)|X]$ and on $P[Y(t) > r|X]$ are narrowest, and thus most informative, where censoring is the smallest. That is, if a large proportion of daughters grew up in households that received AFDC for t years (that is, $P[Z = t|X]$ is close to 1), then the width of the bound is relatively narrow. In the extreme, when all respondents grew up in households that received AFDC for t periods, the selection probability $P[Z = t|X]$ equals 1 and $E[Y(t)|X]$ is identified. However, given the same covariates X, the data cannot reveal much information about the distribution of $Y(t')$: if $P(Z = t|X)$ is close to 1, then $P(Z \neq t|X)$ must be close to 0.

In practice, the probability that a daughter grew up in a household that received AFDC during t years is usually small, and, thus, the worst-case bounds are generally uninformative. Table 3 presents estimated selection probabilities (that is, the probability that a daughter grew up in a household that received AFDC for t years) for black daughters who grew up in single-parent households.[11] This probability ranges from 0.12 when t equals 5 to 0.50 when t equals 0 years. In other words, only a small fraction of each of the four outcome variables—$Y(0)$, $Y(1\text{–}2)$, $Y(3\text{–}4)$, and $Y(5)$—are observed. The majority are censored. Thus, in the absence of prior information, little can be learned about the distribution of the outcome variables, $Y(t)$.[12]

TABLE 3.—ESTIMATED SELECTION PROBABILITIES: THE PROBABILITY A RESPONDENT GREW UP IN A HOUSEHOLD THAT RECEIVED AFDC FOR T YEARS

	Income ($'000)	One to Three Children	Four to Seven Children
t = 0 Years	6–10	0.40	0.21
	11–15	0.43	0.22
	16–20	0.47	0.23
	21–25	0.50	0.25
t = 1–2 Years	6–10	0.22	0.21
	11–15	0.23	0.22
	16–20	0.24	0.24
	21–25	0.24	0.26
t = 3–4 Years	6–10	0.16	0.27
	11–15	0.16	0.26
	16–20	0.15	0.24
	21–25	0.14	0.22
t = 5 Years	6–10	0.23	0.31
	11–15	0.18	0.30
	16–20	0.15	0.30
	21–25	0.12	0.28

Note: These estimates are for the subpopulation of black daughters who resided in non-intact households. See the text and table 1 for further description.

VI. The Middle Ground: Estimates with Alternative Assumptions

To derive useful inferences about the intergenerational transmission of welfare receipt, prior information must be utilized. However, the point estimates displayed in section IV remain suspect. In this section, I present estimates that rely on alternative, less-restrictive assumptions. After describing each assumption, nonparametric estimates of the treatment effects (equations (1) through (3)) are then presented. In total, four sets of assumptions and estimates are examined. The first restricts the ordering of outcomes, while the last three examine either the traditional instrumental vari-

[11] These selection probabilities are estimated using the nonparametric kernel estimator described in section VI.

[12] The treatment effects (equations (1) through (3)) can also be bounded. The upper bound equals the difference between the upper bound on $E[Y(t')|X]$ or $P[Y(t') > r|X]$ and the lower bound on $E[Y(t)|X]$ or $P[Y(t) > r|X)]$. The lower bound for $T_i(t', t|X)$ is determined similarly. Although these bounds do not sign the value of the treatment effect, they can reduce the range of possible values by as much as one-half. For example, when evaluating the difference in the expected duration of receipt (that is, $T_i(t', t|X)$ in equation (1)), the width of this bound equals $60 * (P[Z \neq t'|X] + P[Z \neq t|X])$, and must lie between 60 and 120. In the absence of data or prior information restricting the distributions of $Y(t)$ and $Y(t')$, the width of the range of possible values of this treatment effect equals 120. Given, however, that $P[Z \neq t|X]$ is relatively high for all t, these bounds will generally be uninformative. Furthermore, even the most-informative bounds cannot identify the sign of the treatment effect, because a bound with a width of 60 must include zero.

able or exogenous selection assumptions. Each set of estimates measures the effect of growing up in a household that received AFDC for t' years versus zero years ($t = 0$) on the expected duration of receipt, on the probability of receipt, and on the probability of receiving aid for over two years. Thus, in all cases, the effects described in equations (1) through (3) are measured relative to growing up in a household that did not receive AFDC.

A. *Ordered Outcomes Assumption*

A seemingly innocuous assumption is that the length of a daughter's participation, $Y(t)$, is nondecreasing with the length of time her family received AFDC when she was aged twelve to sixteen. This "ordered outcomes" assumption is implied in the current political discourse and has been asserted by various scholars for nearly a century. Arguably, growing up in a household that received AFDC lowers the cost of future receipt relative to other alternatives. Moffitt (1983) and Blank and Ruggles (1993), for example, have suggested that stigma and information costs deter welfare participation. These costs are likely to be lower for children exposed to parental welfare participation during the teenage years than for their counterparts growing up in non-AFDC households (Gottschalk, 1990; Antel, 1992; An et al., 1993). Similarly, Gottschalk, and Antel argue that the long-term welfare use of parents reduces a family's connections to the labor market and thus increases the (search) cost of working. With relatively higher cost of finding employment and lower cost of receiving government assistance, daughters exposed to AFDC as children are more likely to receive AFDC as adults.

This ordered-outcomes hypothesis implies that the effect of growing up in an AFDC household on the distribution of time a daughter receives aid must be nonnegative for all $t' > t$. However, that hypothesis and the resulting estimates are inconsistent with the belief that AFDC enables some families and children to overcome the burdens of poverty. Certainly, those uncomfortable with this assertion can make plausible assumptions that lead to different conclusions. For example, one could argue that AFDC enables parents to spend more time with their teenage children or that parents make choices that are best for their children.[13] Given these assumptions, growing up in an AFDC household may actually decrease the duration of future receipt for some daughters.

Still, for those willing to accept this assumption, improvements can be made to the worst-case bounds described in section V. Formally, what ordered outcomes imply is

$$t' > t \Rightarrow Y(t') \geq Y(t). \tag{9}$$

[13] There is some empirical evidence that AFDC receipt during infancy may lower the likelihood of future welfare receipt. Currie and Cole (1993), for example, find that AFDC receipt during the prenatal period has a positive effect on the birth weight of children. These findings, however, are not incompatible with this ordered-outcomes assumption, which is restricted to AFDC receipt during the early teenage years.

Observations where $Z \neq t$ may now be informative. For instance, ordered outcomes imply that daughters who grew up in households that received AFDC for more than t years ($Z > t$) would have at most spent the same amount of time receiving AFDC as young adults if their families were to have received aid for t years. Thus, the observed outcome $Y(z)$ provides an upper bound for the unobserved outcome $Y(t)$. So, when $Z > t$, we know that $E[Y(t)|X, Z] \leq E[Y(z)|X, Z] \leq 60$. Because the data reveal the distribution of $Y(z)$, the upper bound for the unidentified conditional expectation can be replaced by $E[Y(z)|X, Z > t]$. Similarly, when $Z < t$, the observed outcome $Y(z)$ provides a lower bound for the unobserved outcome $Y(t)$. Here, we know that $E[Y(t)|X, Z] \geq E[Y(z)|X, Z] \geq 0$; thus, the lower bound of the unobserved expectation for these observations equals $E[Y(z)|X, Z < t]$.

Given this ordered-outcomes assumption, the worst-case bounds for equation (4) can be replaced by the sharp bounds below (Manski, 1995, 1997).

$$\begin{aligned} &E[Y(z)|X, Z \leq t] * P[Z \leq t|X] \leq E[Y(t)|X] \\ &\quad \leq E[Y(z)|X, Z \geq t] * P[Z \geq t|X] + 60 * P[Z < t|X]. \end{aligned} \tag{10}$$

A similar reasoning can be used to place bounds on the $P[Y(t) > r|X]$. In the presence of censoring, these bounds may be uninformative or they may identify $E[Y(t)|X]$ or $P[Y(t) > r|X]$. For instance, the bounds presented in equation (10) are uninformative if $E[Y(t)|X, Z \leq t] = 0$ and $E[Y(t)|X, Z \geq t] = 60$. If, however, $E[Y(t)|X, Z < t] = 60$ and $E[Y(t)|X, Z > t] = 0$, then $E[Y(t)|X]$ is identified. In contrast, the no-information bounds identify $E[Y(t)|X]$ only if the censoring probability $P[Z \neq t|X]$ equals 0.

Given sharp restrictions on $E[Y(t)|X]$ and $P[Y(t) > r|X]$, bounds can also be placed on the intergenerational effects defined in equation (1) through (3). Intuitively, given the ordered-outcomes assumption, the lower bound on the treatment effect equals 0, while the upper bound equals the difference between the upper bound for the distribution of $Y(t')$ and the lower bound for the distribution of $Y(t)$. These bounds are formally derived in appendix A.

Estimation: The bounds for equations (1) through (3) at different values of X and t are estimated using a nonparametric kernel estimator. These data are first divided into race and marital status cells. In particular, the analysis focuses on the 310 black respondents who grew up in single-parent households. The five-year average total family income and the number of children in the household when the daughter was fourteen are treated as continuous covariates. Then, within race and marital status cells, kernel estimates of the probabilities and expectations conditional on family size and income are computed.

A kernel estimate is a weighted average of the observed variables $Y(z)$ and Z conditional on covariates, where the weights are higher for observations that are "close" to the covariates of interest. I use the standard normal density as a kernel, or weighting function, while the measure of closeness, or bandwidth, equals \$9,000 for income and 1.0 for the number of children.[14] Thus, when estimating the bounds at a given income level L and number of children C, the weight given to the i^{th} observation (L_i, C_i) in averaging the outcome at (L, C) is proportional to $\phi[(L - L_i)/9000] * \phi[(C - C_i)]$, where $\phi(\cdot)$ is the standard normal density function. Notice that this kernel gives positive weight to all 310 observations.

Estimates are computed at every \$1,000 interval of income from \$6,000 to \$25,000, and every family size from one to seven children. For ease of presentation, the tables present an aggregated set of estimates by averaging the raw estimates over the respondents that fall within \$5,000 intervals and for households with one to three and four to seven children. Because these data come from a stratified sample, I use the 1968 PSID sampling weights to average over respondents who fall within the broad categories.

Bootstrapped confidence intervals are provided for each kernel estimate (Hardle, 1990, section 4.3). The first step in reporting these confidence intervals is to create a bootstrapped distribution of the upper and lower bounds. Suppose that, for a given set of discrete covariates, a sample has K (that is, 310) observations. From the empirical distribution of Y and Z, one randomly draws with replacement a sample of K pseudo observations, and nonparametrically estimates upper and lower bounds using these observations. This procedure is repeated 200 times to create a bootstrapped distribution of the estimates. Various quantiles of this distribution can then be reported. Below, I present the 0.05 quantile of the bootstrapped distribution of the lower bound and the 0.95 quantile of the bootstrapped distribution of the upper bound. This interval defines a conservative 90%-confidence interval for the bound estimate.

Results: Estimated bounds for the treatment effects (equations (1) through (3)) given the ordered-outcomes assumption are displayed in table 4. As described above, these estimates measure the effect of growing up in a household that received AFDC for t' years versus zero years ($t = 0$) on the expected duration of receipt, on the probability of receipt, and on the probability of receiving aid for over two years. Table 4 displays the upper-bound estimate and the respective 0.95 quantile. By assumption, the lower bound equals 0 in all cases. (See appendix A.)

Table 4 reveals that the bounds narrow as t' falls. For example, the upper-bound estimates for the effect on the expectation range from a low of around 27 months when $t' = \{1, 2\}$ to a high of nearly fifty months when $t' = 5$. Similarly, the upper-bound estimates for the effect on the probability of receipt and the probability of chronic receipt (that is, over 24 months) range from approximately 0.55 to 0.80 as t' moves from $\{1, 2\}$ to five years.

This result occurs because, as t' increases, observations with censored outcomes become less informative. Recall that the upper bound for the mean treatment effect (equation (1)) is derived by differencing the lower bound for $E[Y(0)|X]$ from the upper bound for $E[Y(t')|X]$. When evaluating the upper bound for $E[Y(t')|X]$ given ordered outcomes, only observations with $Z \geq t'$ are informative. We know that, when $Z > t'$, the observed outcome $Y(z)$ provides an upper bound for the unobserved outcome $Y(t')$. However, daughters who grew up in households that received AFDC for less than t' years, could have spent as many as sixty months receiving AFDC as young adults if their parents had received aid for t' years. Thus, observations with $Z < t'$ remain uninformative. As t' increases, the proportion of uninformative observations (that is, $P[Z < t'|X]$) rises, causing the upper bounds of the treatment effect to increase. In fact, for the case in which $t' = 5$, all of the observations with censored outcome variables are uninformative, and the upper bound given the ordered outcomes assumption equals the upper bound derived using no information at all.

Across all values of t', the results in table 4 are consistent with several interpretations. As suggested by the parametric estimates derived in section IV, it may be that growing up in a household that received AFDC during t' years substantially increases the expected length of receipt as well as the probability of future receipt and dependence (that is, over two years). Alternatively, since the lower bound equals zero in all cases, being exposed to AFDC as a child may have no effect on the distribution of time spent receiving welfare as an adult.

Despite this ambiguity, these bounds eliminate extreme inferences, and in certain cases narrow the range of possible values of the treatment effect by as much as 75%. In the absence of data and assumptions, the possible values for the effect on the expected duration range from $[-60, 60]$ months and the effect on the probabilities ranges from $[-1, 1]$. The ordered outcomes assumption cuts the range in half, to $[0, 60]$ months for the expected duration and to $[0, 1]$ for the probability of future receipt. Then, given data, the range of possible values is narrowed even further. For example, when evaluating the effect of growing up in a household that received AFDC during one to two years, the upper bound falls to around thirty months for the expectation and to around 0.55 for the probabilities of participation and long-term receipt.

B. *Instrumental-Variable Assumption*

Although often asserted, the ordered outcomes assumption has never been applied in empirical analyses. In contrast, researchers commonly rely on instrumental variables to help identify the effects defined in equations (1) through (3). For example, Antel (1992), Moffitt (1992), and

[14] Nonparametric estimation requires the researcher to choose both the kernel and bandwidth. The gaussian kernel is chosen for computational ease, and the bandwidths were subjectively selected using the cross-validation method as a guide. Although the results are derived by using bandwidths of 9,000 and 1 respectively, sensitivity tests reveal that the quantitative results are only slightly affected by reasonable changes in the bandwidths and that the qualitative conclusions remain unchanged. See Hardle (1990) for a more detailed description of nonparametric estimators.

TABLE 4.—ESTIMATED TREATMENT EFFECTS GIVEN THE ORDERED OUTCOMES ASSUMPTION

		$T(t', 0,) =$					
	Income ($'000)	$E(Y(t')\|X) - E(Y(0)\|X)$ (Expected Duration)		$P(Y(t') > 0\|X) - P(Y(0) > 0\|X)$ (Probability of Receipt)		$P(Y(t') > 24\|X) - P(Y(0) > 24\|X)$ (Probability of More than Two Years of Receipt)	
		UB	UB+	UB	UB+	UB	UB+
One to three children							
$t' = \{1, 2\}$	6–10	34.0	38.3	0.57	0.64	0.61	0.67
	11–15	34.6	38.6	0.58	0.65	0.61	0.66
	16–20	34.9	39.1	0.58	0.65	0.60	0.66
	21–25	35.2	39.3	0.58	0.65	0.60	0.65
$t' = \{3, 4\}$	6–10	44.0	47.2	0.72	0.77	0.76	0.83
	11–15	44.8	47.8	0.73	0.77	0.77	0.83
	16–20	45.3	48.0	0.73	0.78	0.77	0.83
	21–25	45.7	48.3	0.72	0.77	0.77	0.83
$t' = 5$	6–10	49.6	52.8	0.78	0.84	0.85	0.91
	11–15	50.2	53.1	0.79	0.84	0.85	0.90
	16–20	50.5	53.2	0.79	0.84	0.85	0.90
	21–25	50.6	53.3	0.78	0.84	0.85	0.89
Four to seven children							
$t' = \{1, 2\}$	6–10	27.9	31.8	0.53	0.60	0.50	0.55
	11–15	27.5	31.0	0.52	0.59	0.50	0.54
	16–20	27.4	30.6	0.52	0.58	0.49	0.54
	21–25	27.7	31.1	0.52	0.58	0.50	0.55
$t' = \{3, 4\}$	6–10	35.8	39.6	0.64	0.71	0.63	0.68
	11–15	36.6	40.0	0.66	0.72	0.64	0.69
	16–20	37.7	41.0	0.67	0.73	0.65	0.70
	21–25	39.3	42.7	0.69	0.74	0.68	0.73
$t' = 5$	6–10	46.4	49.8	0.77	0.83	0.78	0.84
	11–15	46.8	49.7	0.78	0.84	0.79	0.84
	16–20	47.4	50.2	0.79	0.84	0.80	0.85
	21–25	48.3	51.0	0.80	0.85	0.81	0.87

Note: UB is upper bound; UB+ is the 0.95 quantile of the bootstrapped distribution. These estimates are for the subpopulation of black daughters who resided in non-intact households. See the text and table 1 for further description.

Levine and Zimmerman (1996) all argue that the variation in local labor market conditions faced by parents does not affect the distribution of time that a child receives aid, $Y(t)$, but does affect the probability that parents participate, $P[Z|X]$. Similarly, I assume that the unemployment rate for the county of residence when the daughter was aged twelve is an instrumental variable, V.

For those willing to accept this assumption, improvements can be made to the no-assumption bounds. By the instrumental-variable assumption, the conditional expectation of the latent outcome variable $Y(t)$ does not vary across values of the instrument V. Formally, let $E[Y(t)|X, V] = E[Y(t)|X]$. This exclusion restriction allows one to replace the worst-case bounds for $E[Y(t)|X]$ with the intersection of these bounds across the values of the instrumental variable. (Manski, 1995; Manski & Pepper, 2000). That is,

$$\max_v \, [E[Y(t)|X, V, Z = t] * P[Z = t|X, V]] \leq E[Y(t)|X] \leq \min_v \, [E[Y(t)|X, V, Z = t] * P[Z = t|X, V] + 60 * P[Z \neq t|X, V]]. \quad (11)$$

Recall that the unemployment rate measure is aggregated into five mutually exclusive and exhaustive categories. Thus, the bound in equation (11) is found by taking the intersection of five no-assumption bounds. A similar reasoning results in bounds on the $P[Y(t) > r|X]$.

Notice that, in the presence of censoring, this instrumental-variable bound is uninformative if the no-assumption bounds do not vary with the instrumental variable V, while it identifies the conditional expectation of the latent variable if the intersection in equation (11) is a point. Thus, in this nonparametric framework, an instrumental variable is "weak" if it has little identifying power in that the bound in equation (11) barely improves on the no-assumptions bounds.[15]

These instrumental-variable restrictions can be used to narrow the no-assumption bounds for the treatment effects (equation (1) through (3). Again, the upper (lower) bound on $T_1(t', t|X)$ equals the difference between the upper (lower) bound on $E(Y(t')|X)$ and the lower (upper) bound on $E[Y(t)|X]$.

Results: Estimates derived using the local unemployment rate as an instrumental variable are displayed in table 5. The striking feature is that the estimated bounds do not indicate

[15] This differs from the classical linear-response model that is formally identified given a weak instrument. The problem with weak instruments in the classical literature is that the sampling distribution of the usual estimator may be poorly behaved.

TABLE 5.—ESTIMATED TREATMENT EFFECTS GIVEN THE INSTRUMENTAL-VARIABLE ASSUMPTION

		$T(t', 0) =$											
	Income	$E(Y(t')\|X) - E(Y(0)\|X)$ (Expected Duration)				$P(Y(t') > 0\|X) - P(Y(0) > 0\|X)$ (Probability of Receipt)				$P(Y(t') > 24\|X) - P(Y(0) > 24\|X)$ (Probability of More than Two Years of Receipt)			
	($'000)	LB−	LB	UB	UB+	LB−	LB	UB	UB+	LB−	LB	UB	UB+
One to three children													
$t' = \{1, 2\}$	6–10	−20.8	−7.1	41.9	45.9	−0.45	−0.24	0.53	0.66	−0.33	−0.03	0.72	0.79
	11–15	−19.6	−7.7	41.1	45.0	−0.44	−0.27	0.50	0.66	−0.31	−0.05	0.70	0.76
	16–20	−20.0	−10.6	40.7	44.0	−0.41	−0.31	0.50	0.67	−0.32	−0.10	0.67	0.72
	21–25	−20.7	−16.2	39.3	42.0	−0.37	−0.34	0.51	0.64	−0.32	−0.21	0.64	0.68
$t' = \{3, 4\}$	6–10	−25.3	−8.6	48.6	51.4	−0.51	−0.30	0.65	0.76	−0.38	−0.13	0.83	0.88
	11–15	−23.6	−10.4	48.7	51.1	−0.48	−0.35	0.64	0.76	−0.38	−0.16	0.82	0.85
	16–20	−24.2	−14.4	48.0	50.4	−0.49	−0.40	0.62	0.76	−0.39	−0.22	0.80	0.84
	21–25	−25.2	−20.9	46.3	49.3	−0.50	−0.48	0.62	0.72	−0.41	−0.32	0.77	0.81
$t' = 5$	6–10	−19.1	−2.3	43.0	46.2	−0.40	−0.19	0.59	0.68	−0.26	−0.01	0.74	0.81
	11–15	−20.2	−8.8	41.3	46.8	−0.45	−0.32	0.53	0.68	−0.32	−0.12	0.70	0.79
	16–20	−24.2	−15.1	40.5	47.0	−0.50	−0.42	0.50	0.69	−0.39	−0.23	0.67	0.79
	21–25	−27.2	−22.7	39.7	46.3	−0.55	−0.51	0.50	0.68	−0.44	−0.36	0.64	0.78
Four to seven children													
$t' = \{1, 2\}$	6–10	−36.4	−27.0	44.3	49.7	−0.63	−0.46	0.72	0.77	−0.61	−0.42	0.75	0.83
	11–15	−36.6	−27.9	43.7	47.6	−0.62	−0.46	0.71	0.74	−0.62	−0.44	0.73	0.80
	16–20	−36.7	−28.4	42.3	45.6	−0.60	−0.46	0.68	0.72	−0.61	−0.46	0.71	0.77
	21–25	−35.8	−27.8	39.6	43.9	−0.59	−0.46	0.63	0.68	−0.59	−0.45	0.65	0.74
$t' = \{3, 4\}$	6–10	−38.9	−33.3	44.5	47.1	−0.65	−0.55	0.75	0.79	−0.61	−0.51	0.77	0.81
	11–15	−40.4	−36.2	45.1	47.0	−0.66	−0.60	0.76	0.79	−0.65	−0.57	0.77	0.79
	16–20	−40.8	−37.7	44.4	47.1	−0.66	−0.61	0.73	0.79	−0.66	−0.61	0.75	0.79
	21–25	−39.9	−37.2	42.7	47.6	−0.67	−0.59	0.69	0.78	−0.66	−0.60	0.70	0.80
$t' = 5$	6–10	−37.8	−32.8	43.0	44.9	−0.61	−0.49	0.71	0.74	−0.64	−0.55	0.68	0.74
	11–15	−38.1	−34.8	42.3	44.5	−0.61	−0.50	0.72	0.75	−0.64	−0.58	0.66	0.73
	16–20	−38.0	−33.5	40.9	44.9	−0.60	−0.48	0.71	0.76	−0.62	−0.57	0.62	0.73
	21–25	−36.6	−30.8	38.8	45.4	−0.58	−0.42	0.69	0.76	−0.59	−0.51	0.58	0.74

Note: UB is upper bound; LB is lower bound; UB+ and LB− are the 0.95 and 0.05 quantiles of the respective bootstrapped distributions.
These estimates are for the subpopulation of black daughters who resided in non-intact households. See the text and table 1 for further description.

the sign of the intergenerational effect. For example, the estimated effect of a family receiving AFDC for five years on the probability that a child receives aid falls in the range [−0.49, 0.71] for daughters who grew up in low-income families with four to seven children. This result does not imply that there is a greater chance that the effect is positive. Rather, the effect of growing up in an AFDC household on the probability of future AFDC receipt lies somewhere within the estimated range: it may be negative or positive.

Still, this relatively weak assumption and these data reduce the range of possible values of the effect of growing up in an AFDC household. While the lower bound is always negative, in certain cases it is close to zero. For daughters who grew up in low-income households with one to three children, for instance, the estimated effect of receiving AFDC for five years on the expected duration of receipt falls in the range of [−2.3, 43.0] months, while the effect on the probability of receiving AFDC for over two years falls within [−0.01, 0.74]. Although in these cases the sign of the effect remains indeterminate, the estimated lower bound is close to zero. Thus, those comfortable with this instrumental-variable assumption learn that, for some cohorts, the effect of growing up in an AFDC household on the distribution of time that a daughter receives AFDC is either negligible or substantially positive.

C. *Instrumental-Variable and Ordered Outcomes (IO) Assumptions*

Those willing to make both the instrumental-variable and ordered outcomes assumptions—hereafter called the *IO assumption*—can narrow the bounds even further. Combining both assumptions allows one to replace the ordered outcomes bound in equation (10) with the intersection of those bounds across all values of the instrumental variable.

Results: Table 6 displays the estimated bounds using this joint assumption. Here, the bounds are much narrower than those derived by either using the instrumental-variable or the ordered outcomes assumptions. For instance, given this joint assumption, the width of estimated values for the effect of growing up in an AFDC household on the expected duration of receipt varies between seventeen months and thirty months. Thus, for those willing to accept the IO assumption, the range of debate about the intergenerational effects of growing up in an AFDC household on the expected duration of future welfare receipt is confined to a one- to two-year period, depending on the subpopulations of interest. In contrast, the bounds derived under either the ordered outcomes or instrumental-variable assumptions have widths ranging from nearly thirty months to over 75 months.

TABLE 6.—ESTIMATED TREATMENT EFFECTS GIVEN THE INSTRUMENTAL-VARIABLE AND ORDERED OUTCOMES ASSUMPTIONS

		$T(t', 0) =$											
		$E(Y(t')\|X) - E(Y(0)\|X)$ (Expected Duration)				$P(Y(t') > 0\|X) - P(Y(0) > 0\|X)$ (Probability of Receipt)				$P(Y(t') > 24\|X) - P(Y(0) > 24\|X)$ (Probability of More than Two Years of Receipt)			
	Income ($'000)	LB−	LB	UB	UB+	LB−	LB	UB	UB+	LB−	LB	UB	UB+
One to three children													
$t' = \{1, 2\}$	6–10	0.3	4.9	29.4	31.2	0.02	0.04	0.36	0.44	0.02	0.16	0.53	0.58
	11–15	1.0	4.3	30.1	31.7	0.03	0.07	0.37	0.47	0.02	0.15	0.53	0.56
	16–20	1.2	3.1	30.6	31.5	0.05	0.10	0.37	0.48	0.03	0.12	0.52	0.55
	21–25	2.2	2.9	30.4	31.3	0.09	0.18	0.39	0.48	0.06	0.11	0.49	0.53
$t' = \{3, 4\}$	6–10	4.0	8.6	38.7	40.6	0.06	0.09	0.53	0.61	0.06	0.21	0.67	0.72
	11–15	3.2	7.4	38.0	41.5	0.06	0.07	0.50	0.61	0.07	0.20	0.64	0.71
	16–20	3.5	5.5	37.9	41.8	0.09	0.12	0.47	0.63	0.08	0.16	0.62	0.69
	21–25	4.4	4.5	36.9	41.6	0.14	0.20	0.46	0.63	0.10	0.14	0.59	0.68
$t' = 5$	6–10	11.9	22.6	43.0	46.9	0.25	0.33	0.59	0.68	0.24	0.45	0.74	0.81
	11–15	8.0	16.2	41.3	46.6	0.18	0.21	0.53	0.68	0.19	0.35	0.70	0.79
	16–20	7.6	11.1	40.5	46.3	0.14	0.15	0.50	0.68	0.15	0.26	0.67	0.79
	21–25	6.6	7.1	39.7	46.3	0.16	0.21	0.50	0.68	0.15	0.18	0.64	0.77
Four to seven children													
$t' = \{1, 2\}$	6–10	0.0	2.8	23.7	24.9	0.00	0.04	0.43	0.47	0.00	0.05	0.39	0.44
	11–15	0.0	3.1	23.6	24.6	0.00	0.04	0.43	0.46	0.00	0.06	0.40	0.45
	16–20	0.0	3.7	23.0	24.9	0.00	0.05	0.42	0.46	0.00	0.07	0.38	0.44
	21–25	0.5	4.6	21.4	24.8	0.01	0.07	0.39	0.45	0.01	0.09	0.36	0.44
$t' = \{3, 4\}$	6–10	1.3	4.6	32.0	34.1	0.02	0.08	0.57	0.60	0.04	0.09	0.57	0.58
	11–15	0.4	4.9	33.2	35.4	0.03	0.08	0.59	0.62	0.03	0.10	0.58	0.59
	16–20	0.4	5.4	34.0	35.9	0.03	0.08	0.59	0.63	0.04	0.11	0.56	0.61
	21–25	1.4	6.4	33.2	36.8	0.07	0.13	0.57	0.64	0.06	0.12	0.51	0.62
$t' = 5$	6–10	8.3	12.2	43.0	44.4	0.19	0.32	0.71	0.75	0.15	0.20	0.68	0.75
	11–15	7.7	12.2	42.3	43.9	0.18	0.32	0.72	0.75	0.14	0.19	0.66	0.75
	16–20	8.6	13.6	40.9	44.1	0.20	0.35	0.71	0.75	0.15	0.21	0.62	0.73
	21–25	10.3	15.8	38.8	44.2	0.26	0.40	0.69	0.76	0.17	0.25	0.58	0.74

Note: UB is upper bound; LB is lower bound; UB+ and LB− are the 0.95 and 0.05 quantiles of the respective bootstrapped distributions.
These estimates are for the subpopulation of black daughters who resided in non-intact households. See the text and table 1 for further description.

In addition to being more precise, these estimates are positive and are often substantial. Growing up in a household that receives AFDC for at least three years appears to substantially increase the time that a daughter receives aid. For example, for daughters who grew up in low-income families with one to three children, the estimated effect of being exposed to AFDC for five years on the expected duration of receipt is at least 22.6 months, reflecting a 426% increase over growing up in a non-AFDC household. For this same cohort, the probability of receiving AFDC increases by at least 0.33 (127%), and the probability of becoming a chronic recipient (that is, over two years) increases by at least 0.45 (600%). Thus, those willing to accept the IO assumption learn that the effect of growing up in a "chronic welfare household" on the duration of future welfare receipt is positive and substantial.

D. *Exogenous Selection*

Most analyses of intergenerational welfare participation assume that unobserved factors affecting the AFDC receipt of parents and their children are unrelated (Moffitt, 1992). In section IV, this exogenous selection assumption (equation (7)) was combined with a parametric model to identify the effects of growing up in an AFDC household. Here, I examine the implications of assuming only that the duration of parental receipt is exogenous. Recall that the exogenous selection assumption in equation (7) states that $E[Y(t)|X] = E[Y(t)|X, Z = t]$. Thus, without any additional restrictions (such as, equation (6)), the exogenous selection assumption identifies the intergenerational effect of growing up in an AFDC household.

Results: Nonparametric estimates derived given the exogenous selection assumption are displayed in table 7. Unlike the Tobit model results (table 2), these estimates do not always decline as the number of children falls and as family income increases. In certain cases, these nonparametric estimates suggest that the effect for daughters growing up households with four to seven children may be less than that for those in households with one to three children. For instance, for those who grew up in single-parent, low-income households with less than four children, being exposed to AFDC for three to four years increases the expected time that a daughter receives AFDC by 21.3 months, from 4.8 months to 26.0 months. For their counterparts growing up in large households, receiving AFDC for three to four years during childhood increases the expected duration of future receipt by 13.9 months, from 6.5 months to 20.4 months.

TABLE 7.—ESTIMATED TREATMENT EFFECTS UNDER THE EXOGENOUS SELECTION ASSUMPTION

	Income ($'000)	$T(t', 0) =$ $E(Y(t')\|X) - E(Y(0)\|X)$ (Expected Duration) (0.05)	Estimate	(0.95)	$P(Y(t') > 0\|X) - P(Y(0) > 0\|X)$ (Probability of Receipt) (0.05)	Estimate	(0.95)	$P(Y(t') > 24\|X) - P(Y(0) > 24\|X)$ (Probability of More than Two Years of Receipt) (0.05)	Estimate	(0.95)
One to three children										
$t' = \{1, 2\}$	6–10	2.6	10.5	19.3	−0.11	0.08	0.25	0.12	0.25	0.39
	11–15	2.0	10.0	17.2	−0.06	0.11	0.27	0.08	0.22	0.35
	16–20	1.8	9.2	16.4	−0.04	0.13	0.30	0.05	0.18	0.31
	21–25	1.0	8.4	15.8	−0.02	0.16	0.33	0.01	0.15	0.29
$t' = \{3, 4\}$	6–10	11.7	21.3	29.9	0.16	0.35	0.53	0.21	0.41	0.60
	11–15	11.1	21.0	28.9	0.20	0.37	0.52	0.18	0.40	0.58
	16–20	9.9	19.7	29.4	0.18	0.35	0.50	0.16	0.37	0.54
	21–25	7.4	17.2	27.7	0.10	0.30	0.47	0.11	0.32	0.49
$t' = 5$	6–10	9.2	**17.0**	26.8	0.03	**0.23**	0.43	0.20	**0.37**	0.55
	11–15	6.5	**13.9**	22.5	0.00	**0.19**	0.36	0.14	**0.30**	0.48
	16–20	3.2	**10.5**	18.7	−0.04	**0.13**	0.30	0.07	**0.22**	0.41
	21–25	−0.5	**6.7**	14.9	−0.12	**0.05**	0.24	0.01	**0.14**	0.32
Four to seven children										
$t' = \{1, 2\}$	6–10	7.3	16.7	25.8	−0.03	0.18	0.39	0.18	0.35	0.48
	11–15	4.6	12.8	21.0	−0.08	0.13	0.31	0.14	0.29	0.43
	16–20	3.2	9.9	17.8	−0.09	0.09	0.27	0.12	0.25	0.39
	21–25	1.0	7.2	14.4	−0.14	0.05	0.22	0.09	0.21	0.33
$t' = \{3, 4\}$	6–10	6.0	**13.9**	19.9	0.02	0.23	0.40	0.18	**0.34**	0.48
	11–15	5.8	13.0	19.7	0.02	0.22	0.39	0.17	0.33	0.45
	16–20	5.4	12.1	19.4	0.04	0.21	0.37	0.16	0.31	0.43
	21–25	4.7	11.2	18.8	0.01	0.19	0.35	0.14	0.29	0.41
$t' = 5$	6–10	7.9	14.2	22.4	0.00	**0.18**	0.35	0.15	**0.28**	0.42
	11–15	8.4	14.6	22.0	0.03	**0.19**	0.37	0.17	**0.28**	0.41
	16–20	8.4	15.3	23.3	0.06	0.21	0.39	0.17	**0.29**	0.43
	21–25	9.1	16.2	23.4	0.09	0.23	0.41	0.18	0.30	0.44

Note: (0.95) and (0.05) represent the 0.95 and the 0.05 quantiles of the bootstrapped distribution of the estimates. Bolded estimates violate the ordered outcomes assumption, and shaded estimates violate the instrumental-variable and ordered outcomes bounds. These estimates are for the subpopulation of black daughters who resided in non-intact households. See the text and table 1 for further description.

Regardless of family size, these nonparametric estimates imply that growing up in a household that received any AFDC has a large positive effect on the expected amount of time that a daughter receives aid. In addition, growing up in an AFDC household increases both the probability of participating and the probability of participating for over two years. Similar to past studies, the results in table 7 suggest that, relative to growing up in a non-AFDC household, being exposed to AFDC during the early teenage years increases the probability of future receipt by as much as 100% (An et al., 1993; Gottschalk, 1992; McLanahan, 1988). For example, for daughters growing up in small, low-income households, the probability of receiving AFDC increases 0.23, from 0.26 if the parents did not receive aid to 0.49 if the parents receive aid for all five years. The effect on the probability of receiving AFDC for over two years is even more extreme, resulting in an estimated increase of between 100% to nearly 700%. Again, for daughters growing up in small, low-income households, the probability of receiving AFDC for at least two years increases 0.37, from 0.05 if she grows up in a non-AFDC household to 0.42 if she is exposed to AFDC for all five years.

Seemingly, growing up in an AFDC household substantially increases the probability of long-term participation, especially for daughters who receive aid as young adults. For daughters who grew up in non-AFDC homes, the probability of receiving AFDC for over two years conditional on having received aid ranges from 0.20 to 0.50. In contrast, for respondents who both received aid as young adults and were exposed to AFDC as children, this conditional probability ranges from 0.64 to 0.90.

These estimates rely on the strong and perhaps implausible exogenous selection assumption. For those who accept the assumption, the estimates provide disturbing insights into the effects of growing up in an AFDC household. Apparently, receiving any AFDC today induces both increased welfare participation and dependence by the next generation. For those who reject this exogenous selection hypothesis, the estimates still suggest that daughters who grew up in AFDC households are more likely to both receive and spend long periods collecting welfare than their counterparts who grew up in non-AFDC homes. These results, however, may simply by an artifact of unobserved factors that jointly cause both parents and daughters to receive aid.

VII. Consistency of the Estimates

The alternative models and estimates presented in sections IV through VI reflect the identifying power of the various assumptions. Another related function of the bounds is to test hypotheses restricting the joint distribution of $Y(t)$ and Z. A model should be rejected if, for any value of X, the estimates lie significantly outside of the no-assumptions bounds, while, if the estimates lie within these bounds, the model cannot be rejected. Here, because the worst-case bounds are essentially uninformative, all but the most extreme estimates pass this test.

Still, by comparing estimates derived under various assumptions, we can test the joint hypothesis that these assumptions are valid. In particular, point estimates made under the exogenous selection assumption can be compared to the estimates made under alternative assumptions. Separately, the exogenous selection (equation (7)) and ordered outcomes (equation (9)) assumptions cannot be empirically refuted. Given the selection problem, censored data cannot reveal if outcomes are ordered or if parental receipt is exogenous.[16] However, if the estimates derived under these different information sets do not overlap, then at least one of the maintained assumptions may be invalid.

A number of the nonparametric estimates derived under exogenous selection lie outside the estimated bounds derived using either the ordered outcomes or joint IO assumption. In table 7, the "exogenous selection estimates" in bold print violate the ordered outcomes assumption, while the shaded figures are inconsistent with the estimates derived given the joint IO assumption. In total, nineteen of the 72 cases violate the ordered outcomes assumption, and sixteen lie below the bounds derived under the IO assumption. The magnitudes of these inconsistencies, however, are generally small, and the 90%-confidence intervals overlap in all cases.

Table 8 displays the analogous estimates from the Tobit model presented in section IV. Here, the expected duration of welfare receipt for a particular subsample defined by X is estimated using the sample analog to equation (8), evaluated using the characteristics of the subpopulation of interest. Because this paper focuses on black daughters who grew up in single-parent households of various sizes and incomes, we are interested in learning

$$\begin{aligned} E[Y(t)|X = {} & \text{(black, non-intact household, average} \\ & \text{income, number of siblings)}, Z = t] = 60 \\ & * (1 - \Phi(u)) + (\Phi(u) - \Phi(l))(X'B + \theta_t) \\ & + \sigma(\phi(l) - \phi(u)). \end{aligned} \tag{8'}$$

Consistent estimates of the conditional expectation in equation (8′) are obtained by replacing the unknown parameters B and θ_t with the estimates presented in table 2. Given the parametric specification, consistent estimates for other subpopulations (such as white daughters who grew up in intact households) can be computed similarly.

Many of the Tobit model estimates in table 8 are also inconsistent with the bounds. Twenty-four of the 72 estimates violate the ordered outcomes assumption, while 28 lie below the bounds derived using the joint IO assumption. These violations are often substantial, although in all cases the 90%-confidence intervals overlap. For example, the Tobit model estimate of the effect of growing up in a household that received AFDC for five years on the expected duration of receipt equals 8.7 months for daughters who grew up in low-income households with fewer than three children. The corresponding lower bound derived using the IO assumption is 22.6 months. Here, the 0.95 quantile of the parametric model estimate of 13.8 months lies slightly above the 0.05 quantile of the bootstrapped distribution of the lower bound. Thus, these results alone do not prove that the Tobit model or the alternative nonparametric models are incorrect. Still, these substantial inconsistencies suggest that at least one of the assumptions—ordered outcomes, instrumental variable, parametric model, and exogenous selection—may be invalid.

The parametric specification can be further evaluated by comparing the Tobit and nonparametric model estimates found under the exogenous selection assumption. Nonparametric models provide a less restrictive alternative. Maintaining the assumption of exogenous selection, the nonparametric model assumes only that the conditional distributions of $Y(t)$ and Z given X vary continuously with X. Thus, to the extent that two estimates differ, the parametric model may be misspecified.

The qualitative conclusions found given the exogenous selection assumption remain consistent across the two models. In particular, the nonparametric estimates displayed in table 7 and the Tobit model estimates in table 8 are all positive and substantial, suggesting that growing up in an AFDC household increases the expected time a daughter will receive AFDC as an adult.

However, tables 7 and 8 reveal substantial differences between both the slopes and the magnitudes of these two sets of estimates. As described in section VI, the nonparametric results do not monotonically decline with income and increase with family size. Furthermore, in many instances, the Tobit model estimates appear to be biased downward. For the estimated effects on the expected duration of receipt and on the probability of chronic receipt, the nonparametric results exceed the Tobit model analogs in all but a few cases. In many cases, these differences are substantial.

If the exogenous selection assumption is incorrect, these inconsistencies carry no implications for the validity of the parametric form. However, under the hypothesis that the exogenous selection assumption is correct, these systematic differences suggest that the basic Tobit model may be misspecified.

In fact, there appear to be interactions in the conditional distribution that are not appropriately accounted for by the parametric specification of the Tobit model evaluated in section IV. Consistent with much of the past literature, the model uses a single indicator variable to account for the race of the respondent. If modified to allow for interactions between the race of the respondent and all other variables,

[16] The instrumental variable assumption can be empirically refuted. If, abstracting from sampling variation, the intersection of the bounds is empty (that is, if the lower bound exceeds the upper bound), the exclusion restriction may be rejected.

TABLE 8.—ESTIMATED TREATMENT EFFECTS UNDER THE TOBIT MODEL ASSUMPTIONS

		$T(t', 0) =$								
	Income	$E(Y(t')\|X) - E(Y(0)\|X)$ (Expected Duration)			$P(Y(t') > 0\|X) - P(Y(0) > 0\|X)$ (Probability of Receipt)			$P(Y(t') > 24\|X) - P(Y(0) > 24\|X)$ (Probability of More than Two Years of Receipt)		
	($'000)	(0.05)	Estimate	(0.95)	(0.05)	Estimate	(0.95)	(0.05)	Estimate	(0.95)
One to three children										
$t' = \{1, 2\}$	6–10	4.3	7.9	11.5	0.10	0.18	0.26	0.08	0.14	0.21
	11–15	4.1	7.7	11.2	0.10	0.18	0.25	0.08	0.14	0.20
	16–20	4.0	7.5	10.9	0.10	0.17	0.25	0.07	0.14	0.20
	21–25	3.8	7.2	10.6	0.09	0.17	0.25	0.07	0.13	0.19
$t' = \{3, 4\}$	6–10	3.2	**7.8**	12.4	0.08	**0.18**	0.27	0.06	**0.14**	0.23
	11–15	3.1	**7.6**	12.1	0.08	**0.17**	0.27	0.06	**0.14**	0.22
	16–20	2.9	**7.4**	11.8	0.07	**0.17**	0.27	0.05	**0.13**	0.22
	21–25	2.8	**7.1**	11.5	0.07	**0.17**	0.27	0.05	**0.13**	0.21
$t' = 5$	6–10	3.6	8.7	13.8	0.09	0.19	0.30	0.07	0.16	0.25
	11–15	3.4	8.4	13.4	0.08	0.19	0.30	0.06	0.15	0.24
	16–20	3.3	8.2	13.1	0.08	0.19	0.30	0.06	0.15	0.24
	21–25	3.1	7.9	12.8	0.08	0.19	0.29	0.06	0.15	0.23
Four to seven children										
$t' = \{1, 2\}$	6–10	5.2	9.1	13.1	0.11	0.19	0.27	0.09	0.17	0.24
	11–15	5.0	9.0	12.9	0.11	0.19	0.27	0.09	0.16	0.23
	16–20	4.9	8.8	12.7	0.11	0.19	0.26	0.09	0.16	0.23
	21–25	4.7	8.6	12.4	0.10	0.18	0.26	0.09	0.16	0.22
$t' = \{3, 4\}$	6–10	4.0	**9.0**	14.1	0.08	**0.19**	0.29	0.07	**0.16**	0.26
	11–15	3.8	**8.8**	13.9	0.08	**0.18**	0.29	0.07	**0.16**	0.25
	16–20	3.7	**8.7**	13.6	0.08	**0.18**	0.29	0.07	**0.16**	0.25
	21–25	3.6	**8.4**	13.3	0.08	**0.18**	0.29	0.07	**0.15**	0.24
$t' = 5$	6–10	4.4	10.0	15.6	0.09	0.20	0.31	0.08	0.18	0.28
	11–15	4.3	9.8	15.3	0.09	0.20	0.31	0.08	0.18	0.28
	16–20	4.1	9.6	15.0	0.09	0.20	0.31	0.08	0.17	0.27
	21–25	4.0	9.4	14.7	0.09	0.20	0.31	0.07	0.17	0.27

Note: (0.95) and (0.05) represent the 0.95 and the 0.05 quantiles of the asymptotic distribution of the estimates. Bolded estimates violate the ordered outcomes assumption, and shaded estimates violate the instrumental-variable and ordered outcomes bounds. These estimates are for the subpopulation of black daughters who resided in non-intact households. See the text and table 1 for further description.

the Tobit model results are much more consistent with the nonparametric estimates displayed in table 7. Apparently, the discrepancies between these models reflect the way in which race and the other covariates enter the parametric specification.

Still, the restricted specification evaluated in section IV is arguably a valid and interesting model to compare to the nonparametric alternative. The more-general parametric specification is neither the norm in the literature nor is it supported by the data.[17] Only ex post, after evaluating the nonparametric estimates are these sharp and systematic differences clearly revealed. Furthermore, parsimonious specifications lead to powerful tests between the parametric and nonparametric models. The nonparametric estimates are completely flexible, effectively allowing for interactions between all of the covariates. As the parametric specification becomes increasingly general, the estimates will converge to the nonparametric alternatives.

VIII. Conclusions

The long-standing debate regarding the intergenerational effects of growing up in a household that received welfare can be informed by empirical analysis. In fact, by combining parametric latent variable models with other identifying assumptions, a number of recent studies have drawn strong conclusions about these intergenerational effects. Being exposed to AFDC as a child increases the probability of receipt as an adult.

Invariably, however, these analyses are controversial with much debate about the validity of the maintained assumptions. For those who believe in the identifying assumptions, the corresponding results follow. For those who do not believe in the assumptions, the results are uninformative. The inferences that can be drawn depend upon the assumptions that one is willing to impose. The analysis in section VII, for instance, highlights substantial inconsistencies between the estimates derived using a conventional parametric model and those that rely on the less-restrictive, nonparametric alternatives. Apparently, either the exogenous selection assumption is invalid or the simple parametric form used in section IV is misspecified, failing to properly account for differences between the races.

Explicitly recognizing the ambiguity created by the selection problem, this analysis presents alternative estimates that rely on a number of weaker and arguably more-credible assumptions. These assumptions and the resulting estimates do not exhaust all of the possibilities. Rather, the assumptions used are motivated by the prevailing

[17] Under the parametric model, as discussed in section IV (see footnote 8), I accept the null hypothesis that the coefficients on the race interaction variables are all zero. With only 31 white respondents in the sample growing up in AFDC households, these interaction coefficient estimates are too imprecise to draw precise inferences.

TABLE 9.—SUMMARY TABLE
Alternative Estimates of the Effect of Growing Up in a Household that Received AFDC for Three to Four Years on the Time a Daughter Receives AFDC for Black Daughters who Grew Up in Low-Income, Single-Parent Households with Four to Seven Children

Section	Assumptions	Estimates of the Effects on the Expected Duration	Probability of Receipt	Probability of More than Two Years
IV.)	Parametric model with exogenous selection	9.0	0.19	0.16
	A.) $Y(t) = \text{Max}(0, (\text{Min}(XB + \theta_t + \epsilon, 60))$			
	B.) $\epsilon \mid X \sim N(0, \sigma^2)$			
	C.) $E(Y(t)\mid X) = E(Y(t)\mid X, Z)$			
V.)	No assumptions	[−43.2, 48.0]	[−0.71, 0.81]	[−0.69, 0.83]
VI.)	The middle ground			
	A.) Ordered outcomes $t' > t \Rightarrow Y(t') > Y(t)$	[0.0, 35.8]	[0.00, 0.64]	[0.00, 0.63]
	B.) Instrumental variable: $E(Y(t)\mid X) = E(Y(t)\mid X, V)$ and $P(Z\mid X) \neq P(Z\mid X, V), \forall V$	[−33.3, 44.5]	[−0.55, 0.75]	[−0.51, 0.77]
	C.) ordered outcomes and instrumental variable	[4.6, 32.0]	[0.08, 0.57]	[0.09, 0.57]
	D.) Exogenous selection $E(Y(t)\mid X) = E(Y(t)\mid X, Z)$	13.9	0.23	0.34

political discourse, social theories, and empirical conventions. For those who accept any of these commonplace hypotheses, the corresponding estimates apply.

Table 9 summarizes a set of these estimates for the effects of growing up in a household that received AFDC for three to four years. The nonparametric bounds clearly reveal the limits of what can be learned under alternative identifying assumptions. The data reveal little in the absence of assumptions invoked to address the selection problem. For instance, without a priori information, the estimated effect of growing up in a household that received AFDC for $\{3, 4\}$ years as opposed to zero years on the expected duration of receipt lies between $[-43.2, 48.0]$ months for daughters who grew up in large, low-income households. However, if it is known that outcomes are ordered or that the local unemployment rate is an instrumental variable, these bounds can be narrowed. The bound under the ordered outcomes assumption is $[0, 35.8]$ months and under the instrumental variable assumption is $[-33.3, 44.5]$ months. The range of possible values is narrowed further to $[4.6, 32.0]$ months if one is willing to utilize both assumptions and if the length of time parents receive AFDC is known to be exogenous, the estimated effect equals 13.9 months.

While many of these assumptions fail to identify the intergenerational effects, the estimated bounds are informative. For those willing only to invoke either the instrumental-variable or the ordered outcomes assumptions, the inferences that can be drawn are much less certain than in the conventional literature. Here, the results are consistent with the notion that growing up in a household that receives welfare substantially increase the likelihood of future welfare dependence, but the results are also consistent with the interpretation that the effect is negligible or even, in certain cases, substantially negative. Still, these estimated bounds rule out extreme positions, and thus confine any debate to lie within a relatively narrow range.

For those comfortable with either the exogenous selection or joint IO assumptions, the resulting nonparametric estimates confirm a disturbing picture of the intergenerational effects of growing up in a household that received AFDC. A teenager exposed to welfare is more likely to receive and become dependent upon government assistance. It is important to keep in mind, however, that these conclusions rest in part on the assumptions imposed. The data alone cannot be conclusive.

REFERENCES

An, C. B., R. Haveman, B. Wolfe, "Teen Out-of-Wedlock Births and Welfare Receipt: The Role of Childhood Events and Economic Circumstances," this REVIEW 75(2) (1993), 195–208.

Antel, J., "The Intergenerational Transfer of Welfare Dependency: Statistical Evidence," this REVIEW 74 (1992), 467–473.

Arabmazar, A., and P. Schmidt, "An Investigation of the Robustness of the Tobit Estimator to Non-Normality," *Econometrica* 50 (1982), 1055–1063.

Blank, Rebecca M., "Analyzing the Length of Welfare Spells," *Journal of Public Economics,* 38(3) (1989), 245–273.

Blank, R., and P. Ruggles, "When Do Women Use AFDC & Food Stamps? The Dynamics of Eligibility VS. Participation," NBER working paper no. 4429 (1993).

Bush, George, "Transcript of President Bush's State of the Union Message," *Facts on File World News Digest,* 52a1 (1992).

Currie, J., and N. Cole, "Welfare and Child Health: The Link between AFDC Participation and Birth Weight," *The American Economic Review* 83(4) (1993), 971–985.

Dugdale, R. L., *"The Jukes": A Study in Crime, Pauperism, Disease and Heredity: also, Further Studies of Criminals,* 3rd ed. (New York: G. P. Putnam's Sons, 1877).

Fitzgerald, J., P. Gottschalk, and R. Moffitt, "The Impact of Sample Attrition on the Second Generation of Respondents in the Michigan Panel Study of Income Dynamics," *Journal of Human Resources* 33(2) (1998), 300–344.

Goldberger, A. S., "Abnormal Selection Bias," in S. Karlin, T. Amemiya, and L. Goodman (Eds.), *Studies in Econometrics, Time Series, and Multivariate Statistics* (New York: Academic Press, 1983).

Gottschalk, P., "AFDC Participation Across Generations," *American Economic Review* 80(2) (1990), 367–371.

——— "The Intergenerational Transmission of Welfare Participation: Facts and Possible Causes," *Journal of Policy Analysis and Management* (1992), 254–272.

——— "Is the Correlation in Welfare Participation Across Generations Spurious?" *Journal of Public Economics* 63(1) (1996), 1–25.

Hardle, Wolfgang, *Applied Nonparametric Regression* (New York: Cambridge University Press, 1990).

Haveman, R., and B. Wolfe, *Succeeding Generations: On the Determinants of Children's Attainments* (New York: Russell Sage Foundation, 1994).
Hill, M. S., *The Panel Study of Income Dynamics: A User's Guide.* (Newbury Park, CA: Sage Publications, 1992).
Katz, Michael B., *In the Shadow of the Poorhouse* (New York: Basic Books, Inc., 1986).
Levine, P. B., and D. J. Zimmerman, "The Intergenerational Correlation in AFDC Participation: Welfare Trap or Poverty Trap?" Institute for Research on Poverty Discussion Paper #1100-96 (1996).
Manski, C. F., *Identification Problems in the Social Sciences* (Cambridge, MA: Harvard University Press, 1995).
——— "Monotone Treatment Response," *Econometrica* 65 (1997), 1311–1334.
Manski, C. F., and D. S. Nagin, "Bounding Disagreements About Treatment Effects: A Case Study of Sentencing and Recidivism," in A. Raftery (Ed.), *Sociological Methodology,* 28 (1998), 99–137. (Basil-Blackwell, 1998).
Manski, C. F., and J. V. Pepper, "Monotone Instrumental Variables: With an Application to the Returns to Schooling," *Econometrica,* 68(4), 997–1012.
McCulloch, O. C., *The Tribe of Ishmael,* Proceedings of the National Conference of Charities and Corrections (Buffalo, NY, 1888).
McLanahan, Sara, "Family Structure and Dependency: Early Transitions to Female Household Headship," *Demography* 25 (1988), 1–16.
Moffitt, Robert, "An Economic Model of Welfare Stigma," *American Economic Review* 73 (1983), 1023–1035.
——— "Incentive Effects of the U.S. Welfare System: A Review," *Journal of Economic Literature* 30(1) (1992), 1–61.
Morgan, J. N., G. J. Duncan, M. S. Hill, and J. Lepkowski, *Panel Study of Income Dynamics, 1968–1989,* [waves I–XXII]. Computer file produced by University of Michigan, Survey Research Center and distributed by Inter-University Consortium for Political and Social Research (1992).
Pepper, J. V., "Dynamics of the Intergenerational Transmission of Welfare Receipt in the United States," *Journal of Family and Economic Issues* 16(2/3) (1995), 265–279.
U.S. Bureau of the Census, Current Population Reports, Series P-60, No. 115, *Characteristics of the Population Below the Poverty Line: 1976* (Washington, D.C.: U.S. Government Printing Office, 1978).
——— Current Population Reports, Series P-60, No. 171, *Poverty in the United States: 1988–1989* (Washington, D.C.: U.S. Government Printing Office, 1990).
Wolfe, B., R. Haveman, D. Ginther, C. B. An, "The 'Window Problem' in Studies of Children's Attainments: A Methodological Exploration," *Journal of the American Statistical Association* 991(435) (1996), 970–982.

APPENDIX A: TREATMENT EFFECT BOUNDS GIVEN THE ORDERED OUTCOMES ASSUMPTION

A naive bound on the treatment effect (equation (1)) can be derived by differencing the appropriate upper and lower bounds for $E(Y(t)|X)$ and $E(Y(t')|X)$. In particular, the upper (lower) bound equals the difference between the upper (lower) bound for the $E(Y(t')|X)$ and the lower (upper) bound for $E(Y(t)|X)$. In the absence of any prior information, these treatment effect bounds are sharp (Manski, 1995). However, using the ordered outcomes assumption, these "naive" treatment effect bounds can be narrowed. In this appendix, I derive a sharp bound for the treatment effects defined in equation (1) through (3) given ordered outcomes.

Proposition 1: Let the ordered outcome assumption (equation (9)) hold. Let $f(\cdot)$ be a specified weakly increasing function, $K_0 = \inf_u = f(u)$ and $K_1 = \sup_u f(u)$. Let $T(t', t) = E(f(Y(t'))|X) - E(f(Y(t))|X)$. If $t' > t$, then,

$$\begin{aligned}0 \le T(t', t|X) \le {} & [E(f(Y(z))|X, Z \ge t') - K_0] * P(Z \ge t'|X) \\ & + [K_1 - K_0] * P(t' > Z > t) + [K_1 - E(f(Y(z))|X, Z \le t)] \\ & * P(Z \le t|X).\end{aligned} \tag{A1}$$

Proof: Let $t' > t$. Using the Law of Iterated Expectations, write

$$\begin{aligned}T(t', t) &= E(f(Y(t'))|X) - E(f(Y(t))|X) \\ &= [E(f(Y(t'))|X, Z \ge t') - E(f(Y(t))|X, Z \ge t')] \\ &\quad * P(Z \ge t'|X) + [E(f(Y(t'))|X,\ t' > Z > t) \\ &\quad - E(f(Y(t))|X, t' > Z > t) * P[t' > Z > t|X] \\ &\quad + [E(f(Y(t'))|X, Z \le t) - E(f(Y(t))|X, Z \le t) \\ &\quad * P(Z \le t|X).\end{aligned} \tag{A2}$$

By monotonicity of $Y(\cdot)$ and $f(\cdot)$,

$$\begin{aligned}z \ge t' &\Rightarrow K_1 \ge f(Y(z)) \ge f(Y(t')) \ge f(Y(t)) \ge K_0 \\ &\Rightarrow f(Y(z)) - K_0 \ge f(Y(t')) - f(Y(t)) \ge 0 \\ &\Rightarrow E(f(Y(z))|X, Z \ge t') - K_0 \ge \\ &\qquad E(f(Y(t'))|X, Z \ge t') \\ &\qquad - E(f(Y(t))|X, Z \ge t') \ge 0 \\ t' > z > t &\Rightarrow K_1 \ge f(Y(t')) \ge f(Y(z)) \ge f(Y(t)) \ge K_0 \\ &\Rightarrow K_1 - K_0 \ge f(Y(t')) - f(Y(t)) \ge 0 \\ &\Rightarrow K_1 - K_0 \ge E(f(Y(t'))|X, t' > z > t) \\ &\qquad - E(f(Y(t))|X, t' > z > t)) \ge 0; \\ t > z &\Rightarrow K_1 \ge f(Y(t')) \ge f(Y(t)) \ge f(Y(z)) \ge K_0 \\ &\Rightarrow K_1 - f(Y(z)) \ge f(Y(t')) - f(Y(t)) \ge 0 \\ &\Rightarrow K_1 - E(f(Y(z))|X, Z \le t) \\ &\qquad \ge E(f(Y(t'))|X, Z \le t) \\ &\qquad - E(f(Y(t)|X, Z \le t) \ge 0.\end{aligned} \tag{A3}$$

The bounds in equation (A3) are sharp. Thus, inserting equation (A3) into equation (A2) yields the sharp bounds in equation (A1).

These bounds improve upon the naive treatment effect bounds described in section V. The upper bound in equation (A1) equals the difference between the upper bound for $E(f(Y(t')|X)$ and the lower bound of $E(f(Y(t)|X)$. The naive lower bound, however, can be negative. In particular, the lower bound derived by differencing the lower bound of $E(f(Y(t')|X)$ and the upper bound of $E(f(Y(t)|X)$ equals,

$$\begin{aligned}&(E(f(Y(z))|X, Z \le t') * P(Z \le t'|X) + K_0 * P(Z > t'|X)) \\ &\quad - (E(f(Y(z))|X, Z \ge t') * P(Z \ge t|X) + K_1 * P(Z < t|X)).\end{aligned} \tag{A4}$$

Because $t' > t$, equation (A4) can be simplified to

$$\begin{aligned}&(K_0 - E(f(Y(z))|X, Z > t')) * P(Z > t'|X)) \\ &\quad + (E(f(Y(z))|X, Z < t) - K_1) * P(Z < t|X).\end{aligned} \tag{A5}$$

Because $K_0 \le E(f(Y(z))|X, Z > t'))$ and $(E(f(Y(z))|X, Z < t) \le K_1$, this lower bound is less than or equal to zero.

The sharp lower bound of 0 in equation (A1) is invariant to the data. Thus, without data, the ordered outcomes assumption restricts the possible values of the treatment effect to lie between 0 and $(K_1 - K_0)$. With data, the upper bound may be lowered. The lowest upper bound, namely zero, occurs if $E(f(Y(z))|X, Z \ge t') = K_0$, $E(f(Y(z))|X, Z \le t) = K_1$ and $P(t < z < t'|X) = 0$. In this case, the treatment effect is zero. The maximum upper bound of $K_1 - K_0$ occurs if $E(f(Y(z))|X, Z \ge t') = K_1$ and $E(f(Y(z))|X, Z \le t) = K_0$. In this case, the data are uninformative.