

1996

Essays on the Intergenerational Transmission of  
Welfare Receipt and Inferences in Clustered  
Samples: Dissertation Summary

John Pepper

*University of Wisconsin-Madison*

---

# Essays on the Intergenerational Transmission of Welfare Receipt and Inferences in Clustered Samples

John Pepper

My thesis includes three chapters that collectively extend the literatures on welfare program participation and survey sampling. Chapters 2 and 3 reassess the effect that growing up in an AFDC household has on future welfare participation using data from the Panel Study of Income Dynamics (PSID), while Chapter 4 empirically examines a method for drawing inferences that account for the fact that the PSID includes multiple respondents from the same household.

## The Intergenerational Transmission of Welfare Receipt

Very soon after birth a change comes over [the young barnacle]. It attaches itself to the crab, loses the characteristics of the higher class, and becomes degraded in form and function. An irresistible hereditary tendency seizes upon it, and it succumbs. A hereditary tendency I say, because some remote ancestor left its independent, self-helpful life, and began a parasitic life... So we have the same in the pauper. *McCulloch (1888)*

Don't feed the alligators... Unnatural feeding and artificial care create dependency. When dependency sets in, these other-wise able alligators can no longer survive on their own. Now I know that people are not alligators, but I submit to you that with our current handout, non-work welfare system, we've upset the natural order. We've created a system of dependency.

Representative John Mica, Florida,  
*quoted by Pear (1995)*

For over a century, scholars and political leaders in the United States have argued that charity and public aid cause 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. Over a century later, President George Bush voiced similar concerns when he suggested that welfare is "passed from generation to generation like a legacy" (1992, State of the Union Address).

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 AFDC today induces participation by future generations.

Given recent efforts to reform the welfare system by both state and federal governments, these studies seem particularly relevant. However, despite the existing literature, the intergenerational effects of growing up in an AFDC household remain uncertain. While receiving AFDC today may induce future generations to participate, there may also exist unobserved factors that jointly determine whether parents and children receive welfare. 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 AFDC 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.

By not carefully addressing the selection problem, previous studies may have misstated the intergenerational effects of growing up in an AFDC household. In fact, this identification problem led Moffitt (1992) to conclude that the existing intergenerational welfare participation studies are essentially "noninformative."

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. Two contributions are made to the existing literature: First, this study investigates the

---

John Pepper received his Ph.D. from the University of Wisconsin—Madison. He is an Assistant Professor of Economics at the University of Virginia. Mr. Pepper's dissertation advisor was Charles F. Manski.

intergenerational effects of growing up in an AFDC household in light of the inherent uncertainties created by the selection problem; and second, by observing the behavior of parents and their daughters for multiple years, this study explores how the time spent receiving AFDC varies with the length of time that parents received welfare.

While the distinction between participation and length of participation has not been entirely ignored in the literature, researchers have focused on the effect 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." 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. Since most AFDC spells last less than two years (Gottschalk, McLanahan, and Sandefur, 1992), incomplete welfare histories may produce inconsistent estimates of lifetime inter-generational relationships (Gottschalk, 1992; Wolfe, Haveman, Ginther & An, 1994). It seems likely, for instance, that the intergenerational effects differ by when the daughter was exposed to parental receipt: infants and teenagers may be affected differently (Furstenberg, Levine, and Brooks-Gunn, 1990).

After formally defining the empirical questions and describing the selection problem, the main body of this analysis provides alternative estimates of the effect of growing up in a household that received AFDC on the distribution of time 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 statistically independent, I first present estimates that are consistent with the past literature. Apparently, growing up in a household that receives any AFDC increases the probability and expected duration of receipt. However, these point estimates rely on strong and questionable assumptions.

Explicitly recognizing the ambiguity created by the selection problem, I then apply a nonparametric bounding method developed by Manski (1995) to examine what can be learned under weaker assumptions. These alternative assumptions and the resulting estimates do not exhaust all of the possibilities. Rather, the assumptions used are motivated by the prevailing political discourse, social theories, and empirical conventions. For those who accept any of these commonplace hypotheses, the corresponding estimates apply.

A logical first step in examining intergenerational welfare participation is to ask what can be learned in the absence of any assumptions invoked to address the selection problem. Here, in the absence of prior information, the data reveal little about the intergenerational effects of growing up in a household that received AFDC for a specified duration.

Then, rather than taking the extreme positions of either making no assumptions, or making the strong assumptions necessary for identification, I explore a middle ground. Four sets of predictions are made. I begin by assuming that being exposed to AFDC as a teenager never decreases the time a daughter receives AFDC as a young adult. While 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 percent. The identifying power of a traditional instrumental variable or exclusion restriction assumption is then explored. In particular, I assume that the cross-state variation in AFDC benefits faced by parents does not affect the time that a daughter receives AFDC as an adult, but does affect the duration of her parents' receipt. Once again, we find that invoking this assumption is useful but 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 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 relies 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 utilizes 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 also suggest that being exposed to AFDC as a child

increases the likelihood of becoming a chronic recipient.

While many of these alternative assumptions fail to identify the intergenerational effects, the estimated nonparametric bounds are informative. At a minimum, these estimates rule out extreme positions, and thus confine any debate to lie within a relatively narrow range. Furthermore, the results strengthen the evidence that growing up in a household that receives any AFDC substantially increases the probability of future welfare receipt and dependence. All of the alternative estimates are consistent with this notion; and given certain assumptions, the effect is necessarily positive and is often substantial. For instance, if it is known that the state AFDC benefit levels faced by the family of origin is an instrumental variable, then the estimates do not rule out the possibility that the intergenerational effects are positive and substantial. If instead the length of time parents receive AFDC is known to be exogenous, the estimated effects are positive and substantial.

### Inferences from Clustered Samples

Empirical analyses often use data consisting of independent clusters of dependent random variables. In fact, most large surveys in the social sciences, including the PSID, use some type of clustered sampling scheme. Similarly, data from medical experiments are often clustered (Rao and Scott, 1992). These data arise when multiple observations exist on the same respondent, as in panel data, and when respondents share a common factor, such as a neighborhood or family. In the presence of clustered data, methods that rely on random sampling to measure the precision of an estimator may be incorrect.

In this chapter, I empirically investigate the effect that clustered sampling has on inferences made using data from the PSID, one of the most important and widely cited surveys in the social sciences. In 1968, The University of Michigan's Survey Research Center (SRC) selected a random sample of approximately 2,900 households to interview for the PSID. Each year since, the members and offspring of these families have been surveyed. Thus, the 1968 representative wave of the PSID includes socioeconomic data on 9,461 individuals and the 1992 panel includes information on 20,078 individuals and 4,052 households.<sup>1,2</sup>

Since each wave of the PSID includes multiple individuals and households that can all be connected to an original 1968 family, these data are clustered. That is, for each independent 1968 household, the SRC collects potentially dependent information on the associated individuals and derivative households. Furthermore, to reduce surveying costs the SRC

utilized a complex geographic clustering scheme so that groups of respondents share the same 1968 block, city, or county. Previous analyses (see, for example, Kish and Frankel, 1974; Scott and Holt, 1982; Moulton, 1990; and, Rao and Scott, 1992) suggest that clustered sampling can have important implications for methods of statistical inference including classical hypothesis tests and asymptotic confidence intervals. These studies find that in practice clustered samples tend to decrease the precision of estimators relative to random sampling. In fact, Hill (1981) examines the effect of the survey design used to create the PSID, including both geographic clustering and stratification, on inferences in a linear mean wage regression model. She finds that in general the estimated standard errors computed under the random sampling assumption are biased downward and suggests as a "conservative" rule of thumb inflating these estimates by a factor of 1.5.

Basic insights into the potential effects of clustered sampling on statistical inference are revealed by examining the variance of the sample mean. Suppose that  $N$  observations of the random variable  $Z$  are drawn from  $C$  independent clusters (e.g., households). Furthermore, assume that  $Z$  has a finite variance ( $\sigma^2$ ), that each cluster has an equal number of observations ( $M$ ), and that there is a common intraclass correlation ( $\rho$ ). Given these assumptions, the variance of the sample mean is

$$N^{-1}\sigma^2(1+(M-1)\rho) \quad (1)$$

(Cochran, 1963). In contrast, the variance of the sample mean in random sampling equals  $N^{-1}\sigma^2$ . So Equation (1) highlights the well-known result that the variance of the sample mean increases with the intraclass correlation and with the number of units per cluster.

Arguably, the intrafamily correlation in the PSID is not zero and the traditional methods of inference made under the assumption of random sampling are inappropriate. In general, however, most researchers continue to treat respondents from the same 1968 household or geographic area as independent observations and thus implicitly ignore the potential intraclass correlations. A survey of the articles published over the last decade (1985 to 1995) in the *American Economic Review*, the *Quarterly Journal of Economics*, the *Journal of Labor Economics*, and the *Journal of Human Resources*, reveals that nearly 80 percent of the analyses that use the PSID assume that respondents who share the same 1968 household are independent observations. Researchers who do account for the clustering of individuals sharing the same 1968

household or geographic region often either arbitrarily exclude observations from the analysis or impose strong prior information about the intracluster dependencies. The most commonly used approach excludes all but a single observation from each 1968 household. Solon (1992), for example, only includes the eldest son in his analysis of intergenerational earnings mobility in the United States. Although inferences drawn under the random sampling assumption with only a single observation per cluster are appropriate, this approach is both unsatisfying in that it results in less precise estimators and unnecessary in that there exist general methods for drawing inferences with dependent data.

Panel data models can also be used to account for the intracluster dependencies. By parameterizing unobserved family specific effects, researchers model the correlation between different family members. In general, however, panel data methods have not been used to account for intra-family correlations but instead are used to account for individual effects over time.

While these models can be used to formalize the intracluster relationships, they only apply to particular parametric specifications and typically require prior information about the form of the within-cluster dependence. Certainly, a more general approach seems useful. Focusing on the limiting distribution of method of moments estimators, I begin by formalizing a method of statistical inference in the presence of clustered samples. This method, which generalizes the White (1980) variance estimator, allows for both arbitrary intracluster dependence and applies to the large class of method of moments problems.

Then I apply this method to various parametric regressions using clustered data from the PSID. While the specific results can not be generalized to other models, they do provide compelling evidence that in clustered samples inferences made under the assumption of random sampling can be misleading. Consistent with the previous literature, important differences are revealed in comparisons between the estimated asymptotic variances derived assuming random and clustered sampling. In general, the estimates derived under random sampling appear to be biased downward. This conclusion should not come as a surprise: statistical inferences are affected by the underlying sampling process. Certainly, these results imply that researchers using clustered samples should be wary of drawing inferences under the assumption that the sampling process is random.

## NOTES

1. Drawn from the 48 contiguous states and the District of Columbia, the 1968 panel includes two subsamples of families: The Survey Research Center's subsample of 2,930 families is representative of the households in the United States in 1968, while the Survey of Economic Opportunity subsample of 1,872 families over-represents the low-income minority population. Thus, in total the 1968 wave of the PSID includes socioeconomic data on 18,224 individuals, and the 1992 panel includes information on 41,420 individuals and 7,561 households. However, to avoid complications that arise from unrepresentative stratification, this analysis focuses on the S.C. subsample of 2,930 families. In 1990, the S.C. added a sample of 2,430 Latino families, which are also excluded from this analysis.

2. Of the 20,078 individual respondents in the 1992 PSID, 13,933 are members or children of members of an original 1968 household. The remaining 6,145 respondents lived in a "PSID" household for at least one year. These respondents are referred to as nonsample members and are only followed as long as they reside in a household with a PSID member.

## References

- Antel, J. 1992. "The Intergenerational Transfer of Welfare Dependency: Statistical Evidence," *Review of Economics and Statistics*, 74, 467-473.
- Bush, George. 1992. "Transcript of President Bush's State of the Union Message," *Facts on File World News Digest*, 52a1.
- Cochran, W.G. 1963. *Sampling Techniques*. Second Edition, New York: Wiley.
- Furstenberg, F., J. Levine, and J. Brooks-Gunn. 1990. "The Children of Teenage Mothers: Patterns of Early Childbearing in Two Generations," *Family Planning Perspectives*, 22(2), 54-61.
- Gottschalk, P. 1992. "The Intergenerational Transmission of Welfare Participation: Facts and Possible Causes," *Journal of Policy Analysis and Management*, 254-272.
- Gottschalk, P., S. McLanahan, and G. Sandefur. 1992. "The Dynamics and Intergenerational Transmission of Poverty and Welfare Participation," IRP Conference Paper.
- Hill, M. 1981. "Some Illustrative Design Effects: Proper Sampling Errors Versus Simple Random Sample Assumptions." In *Five Thousand American Families—Patterns of Economic Progress*, Vol. IX, edited by Martha S. Hill, Daniel H. Hill and James M. Morgan. Ann Arbor: Institute for Social Research, University of Michigan.
- Katz, Michael B. 1986. *In the Shadow of the Poorhouse*, New York: Basic Books, Inc..
- Kish, L. 1965. *Survey Sampling*. New York: Wiley.
- Kish, L., and M. R. Frankel. 1974. "Inferences from Complex Samples," *Journal of the Royal Statistical Association*. Ser. B, 36, 1-37.
- Manski, C. 1995. *Identification Problems in the Social Sciences*, Cambridge, MA: Harvard University Press.
- Moffitt, Robert. 1992. "Incentives Effects of the U.S. Welfare System: A Review," *Journal of Economic Literature*, 1-61.

- Moulton, B.R. 1990. "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units", *Review of Economics and Statistics*, 72, 334-8.
- McCulloch, Oscar C. 1888. *The Tribe of Ishmael*, Proceedings of the National Conference of Charities and Corrections, Buffalo, NY.
- Pear, R. 1995. "House Backs Bill Undoing Decades of Welfare Policy," *New York Times*, March 24, 1.
- Rao, J.K., and A.J. Scott. 1992. "A Simple Method for the Analysis of Clustered Binary Data". *Biometrics*, 48, 577-85.
- Scott, A.J., and D. Holt. 1982. "The Effect of Two-Stage Sampling on Ordinary Least Squares Methods." *Journal of the American Statistical Association*, 77, 848-54.
- Solon, G. 1992. "Intergenerational Income Mobility in the United States," *American Economic Review*, 82(3), 393-408.
- White, H. 1980. "A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity", *Econometrica*, 48(4), 817-38.
- Wolfe, B., R. Haveman, D. Ginther, and C. B. An. 1994. The "Window Problem" in Studies of Children's Attainments: A Methodological Exploration. Unpublished manuscript.