DOI:10.18686/ahe.v7i12.8342

Does Receiving Temporary Assistance for Needy Families (TANF) Benefits Impact Elementary Children's Academic Achievement? A Propensity Score Method Analysis

Yao Wei

Geely University of China, ChengDu, 641423, Sichuan, China

Abstract: Since 1997, the federal Temporary Assistance for Needy Families (TANF) program has funded state welfare programs aimed at providing monetary assistance to low-income families with children. The program serves over 1.6 million children across nearly one million families (Ziliak, 2015). Such a program, executed well, benefits children directly by increasing the resources available to their families, but there may by indirect benefits as well. We seek to explore whether and to what extent receiving TANF benefits impact a child's academic achievement. Such a link is a plausible one, since school achievement is already linked to various negative outcomes TANF seeks to mitigate, like food insecurity (Reid, 2000). We seek to answer the following research question: What is the short-term impact of a family's receiving TANF benefits on an elementary school-aged child's academic achievement, as measured by math and reading test scores?

Keywords: TANF; Propensity Score Method; Elementary Children; Academic Achievement

1. Data

To address this research question, we use the Early Childhood Longitudinal Study, Kindergarten Class of 2010-2011 (ECLS-K:2011), a nationally representative longitudinal study of child development, school readiness, and early school experiences. We will focus our analysis on data collected in the 2nd grade year for the ECLS-K:2011 cohort, because this is the year with the least missing data for our outcomes of interest.

Our outcomes of interest are reading and math test scores collected at the end of the second grade year, standardized to have mean 0 and standard deviation 1. Whether a family reports having received TANF benefits or not over the last twelve months during their second grade interview is our binary treatment variable. We use a set of student-level, school-level, and family-level predictors in our analyses as well, listed in

2. Method

The ECLS-K:2011 dataset is an observational one, but is large and rich in potential covariates. With this sort of dataset, and in the absence of random assignment or a way to simulate it, propensity score methods are an ideal mode of analysis. Rather than draw causal conclusions from groups rendered equal in expectation via randomization, propensity scores are used to create groups that appear roughly equal along a set of observable characteristics. If we assume that these observable characteristics fully govern selection into treatment and that an individual's potential outcomes are independent of treatment status, any difference between two groups with equal propensity scores can be interpreted causally.

OLS estimation is not appropriate here because OLS requires us to assume that treatment is as good as randomly assigned. For OLS to be interpreted causally, the set of included controls must be such that adjusting for differences in these variables is sufficient to adjust for differences in the potential outcomes, but linear regression will become increasingly less credible when treated and control groups are less similar.

The solution here is to obtain estimations of probability that each observation is in the treatment group to construct an observationally similar comparison group from the set of control observations. We want to estimate propensity scores for "family received TANF in the last year" using a logit model for each child in the data set. ists the vector of control variables we use to estimate the propensity scores as well as brief rationales for each variable's inclusion.

Finally, we will run a robustness check, where we will do the exact same analysis, but use kindergarten test scores rather than second grade test scores, excluding families in the treatment group (i.e., who report they had received TANF benefits in the last twelve months when their child was in second grade) who also report they had received TANF benefits in the last twelve months when their child was in kindergarten. By making this exclusion, we ask whether receiving TANF benefits in first or second grade impacts kindergarten test scores for families who started receiving TANF when their child was in first grade or later. (Obviously, we expect to find no impact.) This exclusion to reduce the likelihood that findings from this analysis are due to the expected serial correlation of TANF status over time.

3. Challenges and Limitations of the Method

There are a number of challenges and limitations associated with this project. First, the TANF program is available to families with children for up to five years, so any results must be interpreted as an effect of having received TANF in the past year versus not, not necessarily as an effect of starting TANF or an effect of having TANF versus never having had it. In other words, some treatment families may have been receiving TANF benefits throughout the child's school career, and some control families may have received TANF benefits as recently as thirteen months before the data we are using was collected.

Second, there are some predictors we would like to include in our models but are unable to. For example, we would like to include some measure of parental involvement in a child's education in our model to predict receipt of TANF benefits, but receiving TANF benefits could also provide parents with the funds to, say, Uber to a parent-teacher night, indicating possible joint determination between TANF status and parent involvement. We could not include past TANF or EBT status as covariates due to concerns of serial correlation with our treatment across time. Similarly, we were unable to include variables like a public/private school indicator because they correlate perfectly to treatment: no students in our sample receive TANF and attend private school. Finally, some promising predictors (e.g., citizenship status, state of residence, teaching credential status) are suppressed in the public-use ECLS data.

A third challenge involves concern about having adequate common support in our dataset at low income levels, given TANF eligibility's high dependence on income. However, exploratory tabulations indicate that this is not a major concern: even at the lowest household income level (less than $5000 per year), TANF recipients comprise only a fifth of survey respondents.

4. Results

Having assembled a dataset to predict both binary treatment (whether a family has received TANF assistance in the last year) and our outcome of interest (second grade math or reading score), we begin with a tabulation of our outcome variable to ensure that some basic requirements for propensity score methods are met. This tabulation shows many more untreated observations than treated, an encouraging sign. However, a person might reasonably worry that, with a binary treatment so closely tied to income, perhaps all our treated observations are low-income and all our control observations are high-income, leaving us with little common support. A tabulation by income level provides encouraging news here: in all $5000-dollar income bins, we have many more untreated observations than treated ones, providing evidence that propensity score methods can mechanically work in this data.

We proceeded with our analysis by generating propensity scores for each unit, which are simply predicted values from the logit model. Individuals whose propensity scores are closer to zero have observable characteristics more similar to those who have not received TANF benefits in the last year, while those with scores closer to one have observable characteristics more similar to those who have. We were unable to generate propensity scores for about 70% of observations due to missing values on at least one predictor in the ECLS dataset. This should be cause for alarm, since any systematic attrition from our sample undermines the validity of our causal estimates. To explore the possibility of non-random attrition, we generate an indicator for missing propensity scores (1 if a propensity score is missing) and run basic summary statistics on those for whom we were able to generate propensity scores and on those for whom we were not, shown in.

For both outcome variables (second grade math and reading scores), those without propensity scores scored lower than those with propensity scores, although the differences were small compared to the standard deviations of the distributions. Those without propensity scores seem to be lower-income, on average, than those with propensity scores, with about 2-3 percentage points more of their distribution occupying low-income bins compared to those with propensity scores. Overall, this exploratory analysis indicates that we tended to be missing data for lower-scoring and poorer individuals, although the set of individuals for whom we are missing data does not look totally incomparable to the set for whom we have data. Therefore, we will continue with our analysis, but we must be cautious about generalizing any causal estimates we find, given that we may have had nonrandom attrition.

We then estimate the TOT effect of having had TANF benefits in the last year on second grade math and reading test scores using 1-to-1 propensity score matching with replacement. We also estimated ATE effects for each of these, but report only TOT because its interpretation is simpler in this context and the estimates were very similar.) For both math and reading scores, our estimates are nonsignificant, with point estimates between -.003 and .02 standard deviations and standard errors many times that size Like the OLS with controls analysis, the propensity score matching analysis finds no evidence of an impact of a family having received TANF in the last year on a child's second grade math or reading test scores.

Finally, we use our propensity scores to create inverse probability weights to estimate a TOT effect. Treated units have a weight of 1, and untreated units are assigned weights such that those that look more similar to treated units have higher weights than those that do not. While not perfect, we consider the weighting process successful, since our weighted distributions look much more similar than our unweighted ones (Figures 2 and 3). We use these weights to adjust our initial OLS estimate, producing a regression-adjusted estimate of the TOT. We find no evidence that a family receiving TANF in the last year impacts a second grader's math or reading scores, with point estimates between -.1 and -.05 and standard errors of roughly .07 in both cases This estimate is in line with the OLS and propensity score matching estimates.

Given that our results are nonsignificant, a robustness check does not carry the same weight as it would if we had significant results, but we carry one out anyway. As expected, in a regression-adjusted propensity score weighting analysis, we find no evidence that a family's having received TANF benefits in the last year (i.e., when a child is in first or second grade) retroactively impacts that child's kindergarten math or reading scores, assuming that family did not also receive TANF benefits when their child was in Kindergarten.

References:

[1] Cooper, H., Robinson, J. C., & Patall, E. A. (2006). Does homework improve academic achievement? A synthesis of research, 1987–2003. Review of educational research, 76(1), 1-62.

[2] Crump, R. K., Hotz, V. J., Imbens, G. W., & Mitnik, O. A. (2009). Dealing with limited overlap in estimation of average treatment effects. Biometrika, 96(1), 187-199.

[3] Falk, G. (2014). Temporary assistance for needy families (TANF): Eligibility and benefit amounts in state TANF cash assistance programs. Congressional Research Servive Report. Retrieved from https://fas.org/sgp/crs/misc/R43634.pdf

[4] Manchester, C. F., & Mumford, K. J. (2010). Welfare stigma due to public disapproval. University of Minnesota and Purdue University, unpublished working paper.

[5] Moffitt, R. (1983). An economic model of welfare stigma. American economic review, 73(5), 1023-1035.

Our data also contains a number of respondents who claim to have received TANF benefits in the last year despite being quite wealthy. While these could be legitimate data (e.g., someone who had been receiving TANF benefits could have married a wealthy individual in the last 12 months), the presence of these data raises concerns that these individuals may have misunderstood the question or could be misrepresenting their TANF status or income level to the ECLS interviewer. Ultimately, these data are not a concern for this analysis because our propensity score trimming process eliminates anyone who claims to receive TANF benefits and who makes more than $35,000 per year. The TANF recipients we are left with after trimming have a plausible range of incomes (Falk, 2014).