Learning from Administrative Data (LAD) Project[1]

Analysis Plan for the Connecticut Jobs First Study

October 16, 2020, Updated October 4, 2023

<u>Note to revision of pre-analysis plan (October 4, 2023):</u> This analysis plan has been revised given the acquisition of IRS data. The project team aspired to use information from IRS 1040 forms and W-2 forms but had not yet gained access to these data at the time of the drafting of the pre-analysis plan. The IRS data provide a more comprehensive measure of employment and earnings and also provide additional information to identify the children of study participants. The initial pre-analysis plan described identifying the children using the 2000 Full Count Decennial Census (henceforth 2000 Census). This approach disproportionately misses children who are older at random assignment and no longer coresiding with their parents in 2000. Identifying the children through the combination of 2000 Census and 1040 tax data significantly increases the child match rate. As detailed below, we assessed various methods of combining the IRS and 2000 Census data to find children to ensure that 1) the children match characteristics of children at baseline, and 2) there is balance in the number of children found across treatment and control groups. We updated the power calculations with the final child sample observations. In these updated calculations, we also make a more conservative assumption about sibling outcome correlation which results in a higher minimum detectable effect size for the impacts on children: previously, we had assumed that sibling outcomes were independent; now we assume that they are positively associated, with a correlation coefficient of 0.3. In a small number of cases we made other changes to the original analysis plan. We flag each of these in a footnote.

The Learning from Administrative Data (LAD) initiative seeks to extend the evidence of the long-term effects of various welfare-to-work programs evaluated in the 1990s, programs that included components such as earnings supplements, child care subsidies, work requirements, time limits on benefit receipt, and education services. Evaluations of these programs found that they typically increased mothers' employment and earnings and childcare use. Some increased incomes and others did not, as increases in earnings were offset by reduced benefits. Questions remain about the long-term effects of these programs on children. Did they increase parents' employment and/or family incomes enough to reduce poverty over the long-run or help end the "cycle of poverty" by improving the outcomes of the children of welfare recipients? This study will estimate the long-term effects of the Connecticut Jobs First evaluation.

Connecticut Jobs First examined the effects of one of the first welfare reform initiatives to impose statewide time limits on welfare receipt. These time limits were coupled with financial incentives designed to encourage work. This study focused on the two welfare offices of

[1] Any opinions and conclusions expressed herein are those of the authors and do not represent the views of the U.S. Census Bureau. The Census Bureau has ensured appropriate access and use of confidential data and has reviewed these results for disclosure avoidance protection (Project 7520931: CBDRB-FY23-CES018-019).

Manchester and New Haven, which served about one-fourth of the state's welfare caseload, and tested the effects of the Jobs First program relative to the original AFDC program. The findings showed that, on average, Jobs First led to increases in employment, earnings, and income over the four-year follow-up period, but did not reduce welfare receipt. By the end of the follow-up period, after families had reached the time limit, welfare receipt had decreased and the program's effects on income had diminished. Jobs First also had some positive effects on the outcomes for young elementary school-aged children but had mixed effects for adolescents.

This study will examine the long-term effects of the Connecticut Jobs First program. It will merge data for the study participants with multiple sources of demographic and administrative data in order to assess effects on employment, earnings, fertility, and mortality. The primary focus will be to assess effects on these outcomes during adulthood for individuals who were children during the program period, although effects will also be estimated for selected outcomes for the adult participants. The data include administrative records accessed through the U.S. Census Bureau.

I. Treatment

Jobs First represented several changes from the traditional AFDC program in place at the time of the study. First, Jobs First included a 21-month limit on benefit receipt, after which the family's cash grant was discontinued unless an extension or exemption was granted. Under the existing AFDC program, families did not face any time limits on benefit receipt. Second, Jobs First included unusually generous financial work incentives, allowing working recipients to retain their full grant (e.g., disregarding all earned income) until their earnings reach the federal poverty level. In contrast, AFDC recipients were subject to the standard earned income disregard (of $120 per month, through month 12 and $90 thereafter). Finally, Jobs First required recipients to participate in employment-related services targeted towards quick job placement. Families receiving AFDC, in contrast, were subject to Connecticut's preexisting welfare-to-work program, which had broader exemption criteria (for work activities) and a somewhat stronger focus on education and training.

Jobs First also included a few other changes to the traditional welfare rules. For example, the program imposed a partial "family cap," under which the birth of a new child generated an increase in benefits of only half as much as under traditional rules. Jobs First participants received two years of transitional Medicaid coverage after leaving welfare while employed (as opposed to the one year of coverage provided under prior law). Finally, Jobs First also changed some key rules about the interaction between child support payments and welfare benefits, passing through all child support paid to the custodial parent and disregarding more of that income when calculating cash grant amounts.

II. Study Design

The study is a randomized controlled trial. Between January 1996 and February 1997, 4,803 welfare applicants and recipients were assigned at random to the Jobs First group or the AFDC group. Study participants were subject to the rules of each program for as long as they received welfare.

III. Research Questions

The study will address the following confirmatory research questions:

1. Did the Jobs First program affect the earnings of participants' children, measured at age 21 and older?

2. Did the Jobs First program affect the earnings of participants' children who were under age 6 at randomization, measured at age 21 and older?[2]

Exploratory Research Questions include:

1. Did Jobs First lead to effects on employment, teen fertility, and mortality in adulthood, for individuals who were children during the study period?
2. Did Jobs First generate long-run impacts on employment, earnings and mortality of adult participants?
3. Did Jobs First generate long-run impacts on participation in government assistance programs (e.g., Medicaid, subsidized housing) among adult participants and their children?
4. Did Jobs First impact geographic mobility and neighborhood characteristics for adult participants and their children?
5. Do the long-term effects vary across subgroups, including groups defined by children's gender and age at randomization and parents' employment history, welfare receipt history, and education level? Are the average effects estimated for the full sample similar across the distribution of outcomes, e.g., does the program lead to larger changes in earnings at the bottom of the earnings distribution?
6. Do the long-term effects on earnings and employment of individuals who were children during the study period vary by age in adulthood?

[2] The analysis plan previously defined this subgroup as children under 8 at randomization. We initially made this decision because the combination of the experiment timing (1996) and the end of the LEHD (2015) yielded a data window insufficient to capture child earnings outcomes for the youngest children. For example, A six-year-old in 1996 would only be about 24/25 in 2015Q1 (the last quarter of our LEHD data). The availability of the IRS data helps to alleviate this concern so we now define the subgroup as under 6.

IV. Sample

The program recruited and randomly assigned 4,803 welfare applicants and recipients between January 1996 and February 1997 to either the Jobs First group (2,396), or the control group (2,407) subject to existing AFDC rules.

The analysis sample of the children of the participants will include all individuals age 0 to 17 at the time of randomization who were identified as own children of the matched adult participants (see more information below for the matching process).

V. Data Sources

Data sources to be used for the long-term analysis include the following:

IRS 1040, W-2 and 1099 records: The project has access to 1040 data for tax years 1994, 1995, and 1998 to 2020, W-2 data for 2005 to 2020, and data on selected information returns (e.g. flags for receipt of a 1099) for 2003-2020. These data will be used to measure employment and earnings and will capture earnings in states not included in the LEHD data (which was previously the only source of information available to the project to capture earnings). The records will also be used to capture geographic location and mobility. The IRS data will also be used to expand the identification of the child sample, given that the original study data do not include personally identifiable information (PII) for the children of study participants. In our original approach we constructed our child sample by locating the participants in the 2000 Census and identifying the children co-residing with the participant. With the IRS data we can augment that approach by additionally identifying children who were claimed by participants in the 1040 returns.

Longitudinal Employer-Household Dynamics (LEHD) program: These data capture employment covered by state unemployment insurance systems for states participating in the LEHD program. Additionally, we have earnings for states that allowed us to use their data, including Connecticut. In exploratory analysis, these data will be used to compare employment and earnings outcomes using LEHD vs IRS data.

Benefit and health data: Medicaid enrollment will be obtained from CMS; public housing program participation information will be obtained from HUD; and death records will be obtained from Census Numident.[3]

[3] The previous analysis plan listed Medicaid utilization as an outcome. We only have access to Medicaid enrollment so this has been corrected.

2000 and 2010 Decennial Census (full count): The project will use household composition, relationship and location information from these sources for identifying children of Jobs First participants and for measuring geographical mobility and demographic outcomes.

Census Household Composition Key: This SSA-derived dataset links parents to children born in 1997 or later and will be used to measure teen fertility among the children of the participants.

Original study data: The long-term analysis will also use data from the original evaluation, including administrative records on earnings and benefit usage during the original follow-up period and baseline data for subgroups and regression adjustment.

VI. PIKing, Matching and Identifying Children

Our analysis sample of the adult participants will include all Jobs First adult participants who are assigned a PIK by Census.

To obtain the analysis sample for children, the project first used only the 2000 Census given that we did not have access to IRS records. Using these data, Jobs First adult participants were first matched to respondents in the 2000 Census using their PIK. The analysis sample of the children of the Jobs First participants included all individuals age 0 to 17 at the time of randomization who were identified as own children of the matched adult participant in the 2000 Census.[4]

After the team obtained IRS data (in particular the 1040 data), these records were used to augment the matching conducted using the 2000 Census. As we have many years of tax data and a child may appear across tax years with different tax filers, there are different ways to match children to participants using the 1040s. We used a unique feature of another demonstration project (National Evaluation of Welfare to Work Strategies or NEWWS) to assess various matching algorithms for finding the children of the original program participants. For the NEWWS experiment, the MDRC data include the number and age of all children in the recipiency unit at random assignment. We make use of this information to test whether the children found living with participants in the 2000 Census and/or as dependents in the 1040 data could plausibly be the "same" children that were in the families at baseline. An additional complication with the IRS data is that it only covers those who file a 1040. Therefore, if the interventions affected filing rates, this may affect treatment and control balance in the number of children found. The matching algorithm we have selected identifies the child sample as those who are 1) under 18 at randomization and 2) among all those who have ever claimed them as dependents in all of the 1040 forms we have available, a participant in the experiment was the first claimant or they were identified as own children of the matched adult participant in the 2000 Census. This algorithm has two advantages: 1) it performs best in finding children that

[4] The plan previously included children born within two years of randomization with an interest in maximizing sample size (power). On reflection, concerns about possible endogeneity of children born after random assignment led us to decide to limit to ages 0-17 at random assignment.

match the characteristics of those in the NEWWS roster, and 2) we cannot reject balance in the fraction of children found between the treatment and control groups.

The resulting sample sizes for children in the Jobs First sample are as follows. The 2000 Census leads to a child sample of 5,400 observations (rounded per disclosure review). Combining IRS and 2000 Census, our final child sample is 7,400 (rounded). For comparison there are 2396 adult participants in the Jobs First treatment sample and 2407 adult participants in the control.

Table: Sample Sizes for Connecticut Jobs First

Treatment	Decennial Sample Size	Total Sample Size
Control	2600	3600
Treatment	2800	3800
Total	5400	7400

Notes:

Sample description: Children of Participants in the Connecticut Jobs First experimental sample.

The child sample is defined as all individuals who were claimed first by a participant in the 1040 data or who were found in the 2000 decennial census and were between ages 0 and 17 at the time of randomization.

The column labeled Decennial Sample Size shows the number of children found in the 2000 decennial census. The column labeled Total Sample Size is the sum of children found in the 2000 decennial census plus those found via 1040 tax records.

All results were approved for release by the U.S. Census Bureau, authorization number CBDRB-FY23-CES018-019

VII. Outcomes

Long-term earnings for individuals who were children during the study period will be calculated using the IRS W-2 records and measured as annual earnings across years when they are age 21 and above. Employment is defined using information on whether the individual has any W-2 earnings reported for a given year, also measured for years where the individual is age 21 or older.

For the adult experiment participants, earnings will be defined as annual earnings for the years of available data, 2005-2020. Annual employment rates will be defined using information on any earnings during the year.

LEHD data is available for Connecticut providing quarterly earnings (exploratory).[5] We also have access to an indicator for employment in any state that reported data to the LEHD. The start year for the data varies by state and we currently have data through 2015Q1. For the LEHD earnings and employment outcomes, we include all years in which the individual is 21 or older (for children) or under age 60 (for adults). We will use the LEHD data for exploratory purposes only.

[5] The other states are: AZ, CA, CO, DC, DE, IL, KS, MD, ME, ND, NE, NM, NV, OH, OK, OR, TN, WI, and WY.

Outcomes (exploratory) related to fertility and mortality for the children will be measured as follows: had a birth before age 20 (Census Household Composition Key, for girls) and mortality as of age 25 (Census Numident).

VII. Methods

To address the confirmatory research question related to earnings, we will use linear regressions to estimate the impact of Jobs First on study participants' children. We will cluster standard errors at the family level to adjust for correlations across multiple children. The model will also control for the common set of baseline characteristics across the MDRC welfare experiments, such as parental age, race, education, and employment and welfare history.

Subgroup analysis will be conducted using similar models with the subgroup indicator(s) of interest interacted with treatment status. In addition to heterogeneity of impacts by subgroups, we can also explore how experiments affected the distribution of earnings of parents and children using quantile regression models.[6]

No outcome data will be imputed. Because outcomes are measured using administrative data, when individuals in our sample cannot be linked to the outcome data, we assume a zero value for the outcome (e.g., if an individual cannot be found in the earnings data, we assume the individual had no earnings). For missing pre-test data, we will use dummy variable adjustment (i.e., setting missing cases to a constant and adding "missing data flags" to the impact analysis model). We will not cap W-2 earnings for our confirmatory analysis but will assess the robustness of the findings to outliers and may winsorize the data in exploratory sensitivity tests. Given the many exploratory hypotheses, we will apply appropriate corrections to account for multiple hypothesis testing that control the family-wise error rate, such as that of Holm (1979).[7] This correction method will be applied separately to the set of full sample impact estimates and to the set of tests for subgroup differences.

An analysis of minimum detectable effects suggests that the study is reasonably powered to detect effects. The table below presents sample sizes and MDEs for the sample of adult participants and children.[8]

Our sample sizes are large enough that the MDEs are reasonable. For the child sample, for example, the study would be able to detect an increase in annual earnings of $650 and an

---

[6] See Bitler, Marianne, Jonah Gelbach and Hilary Hoynes (2006). "What Mean Impacts Miss: Distributional Effects of Welfare Reform Experiments," *American Economic Review*, Volume 96, Number 4, pp. 988-1012.

[7] Sture Holm, "A Simple Sequentially Rejective Multiple Test Procedure", *Scandinavian Journal of Statistics*, 6(2), pp. 65-70, (1979).

[8] As stated above, our analysis sample of adults will include all adults who are assigned a PIK. Because the PIK rates are over 99 percent we calculate the MDE for adults using the full Jobs First sample.

increase in employment rate of 3 percentage points. Assuming average earnings of $11,829 for the control group, this earnings effect represents a 5 percent increase.[9] MDEs for the adult sample are also reasonable. In addition, effects smaller than these would arguably not be policy relevant. Finally, minimum detectable effect sizes range from a 0.063 to 0.068, both considered small in the literature.[10]

	Sample size	Annual earnings	Employment	Effect size
Children	7,400	$650	0.030	0.063
Adults	4,803	$749	0.034	0.068

Notes: MDEs calculated assuming test with 80% percent power, 10% significance level with 2-tailed test, and an R-squared for the impact model of 0.10. The MDE for children assumes an intraclass correlation of 0.3.

VI. Correspondence with Ethical Standards for Research

The research has been reviewed and approved by the MDRC Institutional Review Board (IRB). The research design assures minimal risk to study participants. The research will be conducted with de-identified data at the Census Bureau's Federal Statistical Research Data Centers, which will ensure that the privacy and the identity of study participants is protected. A disclosure avoidance review process will further ensure that no participant can be identified in the results that are reported from the study. Due to the low risk and the benefit of conducting this research, the MDRC IRB waived the requirement for study participant children to be consented for this study.

[9] This earnings example is from average total earnings for adults in the Jobs First study in years 3-4, shown in Table 4.1 in https://www.mdrc.org/sites/default/files/full_413.pdf. The control group earnings of $7,783 are converted to 2019 dollars (from current 2000 dollars) using the CPI-U (https://www.bls.gov/data/inflation_calculator.htm).

[10] Cohen, Jacob (1988) *Statistical Power Analysis for the Behavioral Sciences* 2nd edition (Hillsdale, NJ: Lawrence Erlbaum). Lipsey, Mark W. (1990) *Design Sensitivity: Statistical Power for Experimental Research* (Newbury Park, CA: Sage Publications).