Supported Work Leads to Lasting Labor Market Success Among TANF Recipients*

Tania Barham† Brian C. Cadena‡ Lauren Schechter§

June 20, 2025

Abstract

This paper examines the impact of providing TANF recipients with tailored employment support services, including access to a fully subsidized six-month internship with a local employer that may become a permanent job. We use rich administrative data on the universe of Colorado TANF recipients to form an observably similar comparison group and implement a stacked difference-in-differences design. We estimate the effect of program access on participants' employment, earnings, and TANF receipt. Participation increases formal-sector employment and earnings immediately when services start, and, compared to other similar programs, fadeout after services end is only moderate. Program enrollment increases employment by 9.6 percentage points (19 percent) and earnings by $563 per quarter (32 percent) in the year following program exit, likely through the formation of more stable employer-employee matches. Program participation also increases participants' receipt of TANF benefits during the program, but it has no substantial effect on TANF receipt after services end. Estimates using a subset of early enrollees show that earnings effects persist for at least two years after participants have stopped receiving services. Assuming similar persistence for the full sample, a simplified marginal value of public funds calculation indicates that the program is about as efficient as a non-distortionary cash transfer.

JEL Classification: J24, J68, I38, H43

Keywords: subsidized employment, TANF, active labor market programs, staggered difference-in-differences

*We would like to acknowledge helpful comments from Patrick Turner, Chloe East, and seminar participants at the University of Notre Dame and Oregon State University. This research was supported financially by the Colorado Department of Human Services (CDHS), the Institute for Behavioral Sciences at the University of Colorado, and the Wilson Sheehan Lab for Economic Opportunities (LEO) at the University of Notre Dame. We are grateful to staff partners at CDHS and the various implementing agencies of the CW STEP program for their insights about the program, and to Paige Schoonover for excellent research assistance. The opinions and conclusions expressed herein are solely those of the authors and should not be construed as representing the opinions or policies of CDHS or the State of Colorado.

†Department of Economics and Institute of Behavioral Science, University of Colorado Boulder, tania.barham@colorado.edu

‡Department of Economics, University of Colorado Boulder and IZA, brian.cadena@colorado.edu

§Wilson Sheehan Lab for Economic Opportunities, University of Notre Dame, lschecht@nd.edu

1 Introduction

Approximately one in two children and one in four working-age adults in the United States participate in federal social safety net programs (Macartney and Ghertner, 2023). Aid to Families with Dependent Children (AFDC) was, historically, a key part of the social safety net, providing cash benefits to households with dependent children during times when their earnings were low. Beginning in the 1970s and 1980s, a series of reforms introduced welfare-to-work initiatives that were intended to increase AFDC recipients' labor market attachment and decrease their reliance on benefits.[1] After the 1996 Personal Responsibility and Work Opportunity Reconciliation Act replaced AFDC with Temporary Assistance for Needy Families (TANF), increasing work requirements and instituting lifetime limits for beneficiaries, labor force participation increased and benefit receipt decreased among low-income households with young children (Bitler and Hoynes, 2010). Cash welfare receipt continued to decrease steadily over the next few decades, from approximately four million families per month in 1996 to one million in 2021 (DHHS, 2024), with families relying instead on work-based social safety net programs such as the Earned Income Tax Credit (EITC). Increases in work and EITC receipt have arguably more than offset decreases in cash welfare, with even single-parent-headed families at the very bottom of the income distribution now better off when using consumption-based measures of poverty (Meyer and Sullivan, 2008, Han, Meyer and Sullivan, 2021).

Despite these recent improvements for poor single parent-headed families, earnings and employment rates among this population remain low compared to those of other working-age adults. Furthermore, in recent decades the shift away from cash welfare has resulted in the majority of social safety net spending accruing to households with at least some positive earnings and income above the poverty line. This more regressive distribution of aid raises

[1] These policy changes reflect a change in the stated goals of cash welfare programs. The AFDC program, and its predecessor, the Mothers' Pensions program, were implemented to increase the ability of poor single mothers — typically widows — to adequately care for their children at home rather than sending them to orphanages or training schools, and were not originally designed to encourage work. Aizer et al. (2016) and Aizer et al. (2024) evaluate the effects of Mothers' Pensions on children and mothers, respectively.

the question of whether using primarily work-based safety net programs leaves the families most in need of assistance vulnerable (Hoynes and Schanzenbach, 2018). While TANF serves families closer to the bottom of the income distribution than work-based programs like the EITC and Unemployment Insurance (UI) do, it is less effective as an automatic stabilizer when labor markets are weak, exhibiting no significant response in recent recessions (Bitler and Hoynes, 2016). In this context, programs that strengthen the labor force attachment of TANF recipients who face persistent barriers to employment may be particularly valuable to policymakers.

In this paper, we study whether participating in a supported work program improves labor market outcomes for TANF recipients, many of whom face substantial barriers that limit their ability to find employment on their own. The Colorado Works Subsidized Training and Employment Program (CW-STEP or STEP) offered individualized employment supports, including supportive services, job placement assistance, and work experience opportunities. Participants were matched to one or more activities based on their skills and interests, ranging from volunteer opportunities to paid temporary positions at private sector firms, non-profits, or government agencies. The state covered 100 percent of the cost of wages for these placements for a period of up to six months. Unlike many earlier supported work programs, STEP explicitly aimed to place participants in jobs for which the host-site employer was willing to hire the participant after the subsidized six-month internship if the program participant proved to be a successful match. The program could improve participants' employment and earnings through a few potential channels: by building participants' job skills, providing recent work history to reduce the effect of scarring, and by revealing information to potential employers about the participants' productivity and fit for a given job.

We estimate program impacts on participants' employment and earnings using data on the universe of Colorado residents who received a TANF payment during the program's initial operation period from 2018 to 2020. We combine formal-sector UI earnings data with

program tracking data to construct our primary employment and earnings measures, as some subsidized STEP internships do not appear in the UI earnings data. For completeness, we also present analysis using only formal-sector employment and earnings, with the caveat that relying on these measures likely understates the extent to which employment and earnings effects fade out after participants exit the program.

Using a stacked difference-in-differences design suitable for staggered program entry (Wing, Freedman and Hollingsworth, 2024), we compare changes in outcomes for multiple program entry cohorts of STEP participants to changes among matched comparison groups of non-participants who were eligible to join STEP because they received a TANF payment in the same calendar month. This method identifies the effect of STEP participation under an equal counterfactual trends assumption.

To support this assumption, we reweight each entry cohort's comparison group to improve the balance of baseline characteristics, including UI earnings histories and a rich set of baseline observables from TANF administrative data, such as TANF receipt histories, household structure, demographic characteristics, and recent TANF case notes. After reweighting, STEP participants and comparison group members have similar trends in pre-enrollment outcomes and are similar in both fixed and time-varying characteristics that have substantial predictive power for post-enrollment earnings trajectories. Notably, this similarity in pre-trends includes substantial declines in earnings and increases in TANF participation in the year prior to STEP enrollment, suggesting that the treatment and comparison groups experienced similar labor market shocks prior to becoming eligible to participate in STEP.

Although the comparison group experienced a modest recovery from these shocks, STEP participants experienced even larger improvements in employment and earnings both during and after program participation. Difference-in-differences estimates indicate that STEP increased participants' employment rate by 25 percentage points (46 percent) and quarterly earnings by $766 (47 percent) during the first year after program entry — the time period when nearly all program services were delivered. Moreover, a substantial portion of the

program impacts persisted beyond the program period. In the year following exit from the STEP program, STEP participation increased formal sector employment by 9.6 percentage points and quarterly formal sector earnings by $563.[2] Additional analysis of early enrollees with a longer follow-up period shows that post-program impacts remain qualitatively similar more than two years after program exit.

Because STEP participants self-select into the program, we also show that the estimated program effects are robust to using alternative comparison groups that are likely positively selected on their willingness and ability to return to work. First, we restrict the comparison group to individuals without a recent sanction or re-engagement designation, which are indicators of failure to follow a case plan. Second, we limit the comparison group to other TANF recipients without documented barriers to work or exemptions from work requirements. Each of these subgroups is, on average, more work-ready than the treatment group—37% of STEP participants have a re-engagement designation and 61% have a documented barrier to work during the TANF spell in which they enroll in STEP. This robustness check helps address the concern that, even after reweighting the comparison group, STEP participants are positively selected on post-enrollment potential outcomes. Furthermore, we show that the results are robust to changing the set of observable characteristics used to reweight, to limiting the analysis to counties where STEP internships are almost always captured in UI earnings data, and to dropping participants whose service receipt extends into the second year after enrollment.[3]

STEP participation also led to more stable employment outcomes. Participants were 4 percentage points (22 percent) more likely than their non-participating TANF counterparts to remain continuously employed with the same employer during the year following program exit. They were also more likely to work in industries with longer average job duration. These

[2]The estimated program effects in the second year post-enrollment are nearly identical when using outcome variables that do not include any imputed earnings because only a very small portion of the treatment group receive services then.

[3]Andersson et al. (2022) find that estimates of the impact of Workforce Investment Act (WIA) participation are insensitive to conditioning on variables other than earnings histories. We find similar insensitivity when evaluating STEP.

patterns suggest that the program helped participants form quality matches with employers who may not have hired them otherwise. These matches may have formed by reducing employers' costs of trying out an employment match with low expected productivity, by providing participants with work experience in more stable sectors, or through a combination of both mechanisms.

Unsurprisingly, STEP participation increased TANF receipt during the first year after enrollment, as continued TANF participation was a condition of STEP participation. While these additional benefits likely replaced lost income in the wake of participants' recent negative earnings shocks, participants may have used up more of their lifetime TANF limit during a period of stable employment than they would have preferred, making the effect of additional benefit receipt on participants' long-term well-being ambiguous. We are unable to distinguish whether the program increased participants' ability to meet work requirements and retain their benefits or whether they were simply less able than the comparison group to voluntarily exit the TANF caseload and conserve future months of eligibility.[4] We find no lasting impact on TANF receipt after participants have exited the program.

This paper contributes to two key strands of the literature. First, it adds to our understanding of the effectiveness of transitional jobs programs by demonstrating that STEP's program model, which features individualized placements with the potential for conversion to unsubsidized employment with the host-site employer, may be particularly effective for its target population of low-income parents with young children. Second, it demonstrates an effective way to increase the employment and earnings of social safety net beneficiaries while preserving access to benefits. Existing work typically finds that other policy changes or interventions that focus on strengthening incentives to work reduce benefit receipt and fail to produce long-term earnings gains.

Prior work in the first literature on the effectiveness of transitional jobs programs produced smaller and less persistent program impact estimates. Although prior studies evalu-

[4]Grogger (2003) finds that some recipients do strategically and voluntarily exit welfare caseloads to conserve future eligibility in response to time limits.

ating similar subsidized employment programs have typically used Randomized Controlled Trials (RCTs), we believe our difference-in-differences estimates—supported by matching on a rich set of covariates, pre-trend balance, and robustness checks—provide credible evidence that allows for meaningful comparison of program effectiveness across program models. Unlike STEP, earlier subsidized employment programs commonly placed all participants in similar temporary positions at a single host site. These early programs produced only short-run gains, with impacts on earnings and employment fading after the programs ended.[5] One more recent program with standardized placements that targeted TANF recipients was the "Paid Work Experience" treatment arm of the Los Angeles Subsidized and Transitional Employment Demonstration (LA STED) (Anderson et al., 2019). Participants were placed in nonprofit jobs that were explicitly not intended to lead to permanent employment, and earnings impacts were smaller than our in-program estimates from STEP while participants were in their transitional jobs and faded out completely after participants exited subsidized employment.

A limited literature examines subsidized employment programs that offered the potential for permanent job placement but that targeted different populations. Two sites in the Enhanced Transitional Jobs Demonstration (ETJD) featured this design element but focused on non-custodial parents (Milwaukee) and formerly incarcerated individuals (Fort Worth) (Barden et al., 2018). ReHire Colorado (Barham, Cadena and Turner, 2023) is programmatically similar to STEP and is broadly available to low-income individuals across the state but serves relatively few TANF recipients, instead prioritizing older workers, non-custodial parents, and military veterans. STEP produced meaningfully larger and more persistent effects on employment than these comparable programs. The most similar program targeting TANF recipients is the "On-the-Job Training" arm of the Los Angeles Subsidized and Transitional Employment Demonstration (LA STED) (Anderson et al., 2019). This program also placed participants in jobs with the potential for permanent hire but had lower

[5]See Bloom (2010) for a review of the earlier iterations and (Barden et al., 2018) for a review of some more recent programs with similar models within the Enhanced Transitional Jobs Demonstration Project.

placement rates, a shorter fully subsidized trial period, and smaller post-program impacts. We provide evidence consistent with the interpretation that reducing the cost to employers of testing out riskier employee matches is a key mechanism behind STEP's effects. These programmatic differences may have limited that LA STED's ability to take full advantage of the information revelation channel.

We contribute to the second literature by identifying a cost-effective intervention that leads to substantial increases in employment and earnings among public benefits recipients. These improvements are an important objective for policymakers, as social safety net programs typically have large positive impacts on children that are partially offset by modest distortionary impacts on adult labor supply (Guldi and Schmidt, 2017, Aizer, Hoynes and Lleras-Muney, 2022). Recent work evaluates whether work requirements have been effective in offsetting disincentives to work that are inherent in social safety net programs. Evidence from SNAP suggests that work requirements result in reduced access to benefits without increasing labor force participation (Gray et al., 2023, Cook and East, 2024). Similarly, Richard and Bart (2024) find that sanctions for TANF recipients failing to meet work requirements produce a small labor supply response that is insufficient to offset lost benefit income. Further, increasing the duration of benefit ineligibility after a sanction results in long-term reductions in employment and earnings. This evidence from both SNAP and TANF implies that work requirements and the associated penalties for non-compliance are not an effective way to increase long-term self-sufficiency among the eligible population.

Furthermore, a body of earlier research evaluated a wide range of welfare-to-work programs implemented before PRWORA in the 1980s and early 1990s. These programs typically combined changes in the incentive structure of cash welfare programs with some type of employment supports or job search assistance. Consistent with our findings, the most effective programs were "jobs-first" models that prioritized job placement assistance over basic education and training. Earnings effects were largest for programs that placed participants into

higher-quality jobs by taking their individual skills and experience into account.[6] Many of these earlier experiments were designed to simultaneously reduce welfare receipt and increase earnings. They altered work requirements, time limits, or the benefit reduction rate from increases in earnings in addition to providing employment-related services. These welfare-to-work experiments featured bundled interventions, making it difficult to isolate the effect of employment services alone.

In contrast, this study evaluates the impact of providing supported work to cash welfare beneficiaries under a standard set of post-PRWORA benefit eligibility rules that include work requirements, time limits, and generous earned income disregards in benefit calculations.[7] We find that this approach substantially improves participants' labor market outcomes both during and after the period of subsidized employment. These gains are accompanied by a short-term increase in benefit receipt during program participation, but there is no sustained change over the longer term.

These results also contrast with the small or even negative effects of other interventions intended to increase labor market attachment among the TANF population specifically, such as providing placement with a temporary-help firm Autor and Houseman (2010).[8] Supported work programs like STEP that are available to beneficiaries of public benefits may therefore

[6]Greenberg, Deitch and Hamilton (2010) summarizes the findings of 28 such welfare-to-work experiments conducted by the Manpower Demonstration Research Corporation (MDRC). To our knowledge, none featured subsidized employment. Examples of successful jobs-first program models include Portland's JOBS program and various GAIN programs in the National Evaluation of Welfare-to-Work Strategies (Freedman, 2000).

[7]Earnings from STEP internships are automatically disregarded in benefit calculations. However, Colorado offers a 100% earnings disregard for the first six-month recertification period after recipient obtains a new job, and a 67% disregard for existing earned income sources after that. In practice, this policy means that both STEP participants and non-participating TANF beneficiaries in the comparison group likely faced similar benefit reduction rates for earned income.

[8](Autor and Houseman, 2010) also studied a "direct hire" intervention that helped participants find unsubsidized jobs with local employers, and it had estimated effects more similar to STEP's. Even with similar estimated impacts, STEP's use of temporary subsidies rather rather than direct placement allows STEP to serve participants who may not have been competitive candidates for direct hire by giving employers the ability to learn to learn the match quality with minimal downside risk. Further, the prior study uses an IV methodology that relies on the relative propensity of different placement agencies to use temporary-help rather and direct hire placements. It is therefore best suited to comparing the relative effectiveness of these two types of interventions, and comparisons to the non-placed portion of the caseload may introduce selection bias.

provide an "all of the above" path forward for policymakers trying to improve employment and earnings without necessarily reducing cash assistance.

2 Background and Program Description

2.1 Program Design

The Colorado Works Subsidized Training and Employment Program (STEP) is a TANF-funded program that provides case management, job coaching, and other support to enable participants to engage in a period of supported work with a local employer. The program was administered jointly by the Colorado Department of Human Services (CDHS) and the Colorado Department of Labor and Employment (CDLE) and was implemented through local contractors, including county workforce centers and in some cases a workforce division of the county TANF office. Our analysis uses data on the universe of Colorado residents who received a monthly TANF payment (Basic Cash Assistance) in a county where STEP services were available in its initial operating period from January 2018 to June 2020.[9] Any individual in a participating county who received monthly TANF benefits during this time was eligible to participate in the STEP program.[10]

The program provided an individualized set of supportive services and work experiences designed to improve a participant's likelihood of gaining unsubsidized employment, including the ability to work in a temporary subsidized job at a local host-site employer. There were no direct costs to the employer, as the program used state TANF funds to cover the full cost of the employee's wages for up to six months of part-time or full-time employment. In some cases, the local contractor or a staffing agency served as the employer of record, which also eliminated any administrative burden of hiring the participant; in other cases,

[9]The program ended in July 2020 but was restarted in July 2022. As noted in Section 3.2, we drop the relatively few participants who enrolled after March 2020, as their experience in the program was likely different than that of the typical participant.

[10]As of January 2018, 74% of adults on the Colorado TANF caseload lived in a STEP county.

local contractors reimbursed the host site employer for wages paid during the internship.[11] Unlike many prior transitional jobs programs that placed participants into a limited set of pre-selected jobs, STEP placements were tailored to fit each participant's skills, education level, and interests. Most placements were structured as trial positions that could ideally lead to a hire into a regular position at the host site if successful. The STEP operations guide makes clear that employers are under no obligation to hire successful participants but implies that, ideally, the host-site employer would consider the participant for an open role before conducting a full search.[12] If case managers determined that participants needed additional training or experience before starting a paid internship, they could also place the participants in a volunteer position or in a more standardized work experience (e.g. retail work). These participants retained eligibility for a paid individualized placement after they had demonstrated important soft skills and reliability in their initial placement(s).

STEP participants were also eligible for supportive services to ensure that they were prepared for and successful in their supported work experiences. These services included job coaching, financial assistance to cover transportation costs or to purchase work tools, and assistance with childcare. However, implementing partners were directed to spend STEP funds on these additional supports only after exhausting available TANF funding, which implies that both STEP participants and those receiving standard TANF benefits were eligible for similar supportive services. STEP participants could also receive job training toward a certification as a part of their transitional job, but host site employers could not require participants to obtain a certification prior to the start of the internship.[13]

Importantly, any earnings from STEP internships were fully disregarded when calculat-

[11] According to conversations with implementing partners, the wage rates were often set at Colorado's minimum wage, which was $10.20/hr as of January 1, 2018 when the program began and which rose to $11.10 on January 1, 2019 and to $12.00 on January 1, 2020.

[12] A small share of placements (about 5 percent) were structured explicitly as on-the-job training where a participant who successfully completed training would be assured of a regular position. Most placements were instead "work experience" or "subsidized employment."

[13] As an example, an employer could initially hire the participant in a lower position (Care Aide) and use STEP funds to pay the costs of the participant earning a certificate to qualify for a higher-level position (CNA).

ing participants' TANF income eligibility and thus could not result in a reduction in TANF benefits during their participation in the program. In practice, this is not dissimilar to the treatment of earnings from any new job under Colorado TANF rules. Colorado's earned income disregard – the proportion of earnings excluded from the calculation of the participant's monthly benefit – is 100% in the first six-month recertification period after the participant finds a new job, and 67% in subsequent periods. STEP participants were required to remain enrolled in TANF for the duration of their program participation, which did count against the federal lifetime TANF limit of 60 total TANF months.

This program model was expected to improve future labor market outcomes through multiple channels. First, participants were offered coaching, supports, and barrier mitigation to improve their overall work readiness and to assist in their job search process. Second, STEP internship placements were expected both to improve participants' human capital through work experience and on-the-job training and to provide recent work history to prevent participants from experiencing the scarring effects of a long stretch of unemployment.[14] Finally, the subsidized and temporary nature of the internship was intended to encourage employers to hire workers they may not have selected through their regular recruiting processes. To the extent that employee quality or employer-employee match quality is revealed only after the employee has begun working the job (Altonji and Pierret, 2001, Pries and Rogerson, 2005), the program was intended to uncover productive matches that would have otherwise gone undiscovered.[15]

[14]A substantial literature documents the negative impact of job loss on future earnings (Ruhm, 1991, Jacobson, LaLonde and Sullivan, 1993, Stevens, 1997, Arulampalam, 2001, Gangl, 2006), although Krolikowski (2018) makes the important point that selecting the correct comparison group can substantially change the estimated magnitude. Rose and Shem-Tov (2023) recently confirmed the displacement costs for low-wage jobs specifically.

[15]Barham, Cadena and Turner (2023) study a similar program with some overlap in program contractors and conclude that this final mechanism was likely the most important.

2.2 Recruitment, Enrollment, and Service Receipt Timing

TANF case managers recruit potential participants for the STEP program based on their individual goals, current employment situation, and interest in the program. TANF recipients in Colorado are subject to work requirements, and TANF case managers often assist them with job search activities, enrollment in job training and formal education programs, and resolution of any significant employment barriers such as transportation or child care challenges to satisfy those requirements. Participants enroll in STEP after receiving a referral from their TANF case manager and, in some cases, after meeting with a representative of the STEP contractor.[16] TANF recipients who choose not to enroll in the STEP program must find unsubsidized employment or participate in other job training or education programs in order to meet work requirements so that they remain eligible for TANF benefits at every six-month re-certification window.[17]

Table 1: Descriptive Statistics, STEP Program Events

	Mean	SD	N
Months from TANF Start to STEP Enrollment	7.20	8.35	912
Months from STEP Enrollment to STEP Placement (If Placed)	0.69	1.40	588
Total Months in STEP Program	4.62	3.75	912
Placed in Transitional Job	0.64	0.48	912
Months in First Placement (If Placed)	2.66	1.65	588

Note: Data come from CDLE STEP program data and CDHS TANF benefits data, described in Section 3. The sample includes STEP participants who applied between January 2018 and March 2020.

Table 1 indicates that STEP participants enrolled in the program, on average, about seven months after they started a new TANF spell, coinciding with the typical recertification timeline. Time from enrollment to internship placement averaged less than one month, but the timing varied substantially by county—some counties enrolled participants only after finding them a placement and others started the search for a placement after enrollment.

[16]For the first year of the program's operation, referrals were not captured in program tracking data, and no TANF caseworker identifiers are captured in the CBMS benefits data. For these reasons, we do not use referral/non-referral as a source of identifying variation.

[17]Figure A-2 shows bunching in the distribution of TANF spell length at 6-month intervals, indicating that many recipients likely exit the caseload due to failure to re-certify.

64 percent of enrollees were successfully placed in an internship, and nearly all participants exited the program within one year (see Appendix Figure A-1 for further details).

3 Data

3.1 Data Sources and Key Variables

We combine three administrative data sources to form a panel dataset at the individual level covering the period from January 2015 to December 2021. First, the Colorado Benefits Management System (CBMS) provides data on the universe of Colorado adults—more than 30,000 individuals—who received a TANF benefit payment while STEP was in operation. For both STEP participants and members of the comparison group, these data provide demographic characteristics (e.g. age, education, county of residence, and household structure), which we measure as of each potential STEP enrollment month. We also have CBMS data on participation in TANF work activities that satisfy work requirements (e.g. education, job training, part-time and full-time employment), transportation, medical, childcare or other barriers that may make it difficult for the client to find work, and sanctions or case plan re-engagement designations indicating failure to follow the case plan.[18] These TANF case variables are measured during the current TANF spell up to and including each potential enrollment month, as these data are typically updated when a recipient's case plan is reviewed or changed. As we will detail in Section 4, we use these data to reweight the comparison group based on observables and to show balance between STEP participants and the STEP-eligible, non-participating TANF recipients that make up the comparison group. The CBMS data also report monthly TANF benefit payment amounts, and we analyze both binary TANF participation and monthly payment amounts as secondary outcomes.

Second, STEP participants are identified from the Colorado Department of Labor and

[18]We observe this information only during months when the individual is receiving TANF, but TANF receipt is the primary program eligibility criterion, so these characteristics are available at baseline for the full sample.

Employment (CDLE) program tracking data that provides information on dates of key STEP program events, including enrollment, placement, and exit. We code all enrollees as members of the treatment group, regardless of whether they were placed in an internship, because they all became eligible to receive individualized work supports. These data are merged with the CBMS data by social security number, name, and/or birth date. The match rate between the program tracking data and CBMS data is high. Only three percent (35/1,201) of STEP participants identified in the CDLE program data could not be matched with the CBMS data - these observations are removed from the sample.

Finally, we use Unemployment Insurance (UI) earnings data from CDLE to construct labor market outcomes. These data record quarterly earnings from formal sector employment in Colorado but do not capture earnings from employment as an independent contractor, from informal work, or from any work performed outside of Colorado. We also construct quarterly indicators for formal-sector employment, which we code as 1 when an individual has positive earnings and 0 otherwise.

Earnings from paid STEP internships are sometimes recorded in the UI data, and we include them when building both quarterly earnings and quarterly employment outcomes. However, comparing UI earnings records to placement records reveals that a substantial portion of STEP placements did not result in formal-sector earnings (see Appendix Figure A-4), potentially due to program contractors reimbursing host site employers for wages paid to participants in a manner other than formal W-2 employment. Using outcome variables constructed using UI earnings data alone would result in under-estimates of in-program earnings effects and thus understate the extent to which effects fade out after participants exit the program. Therefore, our preferred outcome variables include imputed placement-based earnings for STEP participants who had a subsidized internship that appears not to have generated earnings records in the UI data. Specifically, we identify these participants as individuals who had at least one quarter during which the program tracking data records them as being in a subsidized internship but who had zero earnings recorded in the UI

data for that quarter. For those individuals, we impute their STEP internship earnings by assigning them the average daily subsidized wages paid to STEP participants for each day they were listed as being in the STEP placement.[19] We then aggregate these imputed earnings to the quarterly level and add them to any observed UI earnings. We also construct an employment indicator that codes an individual as employed in a quarter if they have any UI earnings or if they are recorded as working in a subsidized job for any portion of that quarter. Although these data sources necessarily treat income from sources other than formal-sector employment and STEP placements as zero, they nevertheless provide a reliable measure of the extent to which STEP participation improved formal-sector earnings and earnings from subsidized placements.[20]

3.2 Analysis Sample

We limit the analysis sample to individuals whose outcomes are consistent with program eligibility – specifically, to individuals who received a TANF payment during the program's operation and did not have recorded formal-sector earnings that far exceeded the TANF eligibility thresholds in the months when they received TANF benefits, as these cases likely reflect incorrect identity matching or earnings amounts that were entered incorrectly by employers. We also limit the comparison group to individuals residing in counties where STEP operated. The final sample consists of 912 STEP participants and 24,991 program-eligible TANF recipients in the comparison group. See Appendix Table A-1 for treatment and comparison sample sizes in each enrollment month and Appendix Section A.2 for details on sample restrictions.

[19]To account for any changes over time in the placement-based earnings, we use separate averages for each fiscal year. The daily averages are roughly equivalent to working for state minimum wage at 30 hours per week, which CDHS staff described as the modal STEP internship experience.

[20]We assume that STEP participants and non-participants engaged in other informal work at similar rates. If program participation causes STEP participants to substitute away from other informal work

3.3 Treatment Group Characteristics

The first column of Table 2 shows that the characteristics of STEP participants broadly mirror the TANF recipient population from which they are drawn.[21] Household and demographic variables measured as of the STEP enrollment month in the CBMS data show that most participants are women with children in households without another adult. More than half have at least one child under 5, and around two-thirds have at most a high school diploma or GED. UI earnings histories indicate that roughly 44 percent were employed in the quarter prior to STEP enrollment, and during the three years leading up to enrollment, participants earned on average about $1,700 per quarter, or roughly $3,400 per quarter among those who were employed. The majority of STEP participants had an employment barrier, such as childcare or transportation difficulties, noted by their TANF caseworker during the TANF spell in which they enrolled in STEP.

The remainder of the table shows unweighted means of the same characteristic for the comparison group (column 2), differences in means (column 3), and t-statistics from a test of equal means (column 4). There are meaningful differences in these baseline characteristics—relative to the comparison group, treatment group members are slightly older, more likely to be the only adult in their household, less likely to have an infant, and less likely to have a barrier to work identified by their caseworker. They were also more likely to have had some employment over the prior year. On the other hand, they had slightly lower earnings and substantially lower TANF benefits prior to enrollment. Some of these differences suggest the potential for positive bias—for example, participants without infant care needs or other work barriers may have an easier time recovering their labor force attachment and earnings. Alternatively, because the treatment group had stronger work histories, the comparison group may have had more room to improve their labor market outcomes and thus more positive trends in untreated outcomes, which would lead to negative bias. These differences

arrangements, the imputation procedure would over-estimate in-program effects and overstate fade-out of earnings effects after services end.

[21] See Appendix Table A-2 for further details.

Table 2: Sample Characteristics

	(1) Mean (Trt.)	(2) Mean (Comp).	(3) Diff. Means	(4) t (Diff.)
Age	33.21	31.96	1.25	4.52
Female	0.87	0.83	0.05	4.01
Only Adult in Household	0.85	0.80	0.05	4.42
Number of Children in Household	1.94	1.94	0.00	0.08
Infant in Household	0.15	0.19	-0.04	-3.13
Child Under 5 in Household	0.58	0.60	-0.02	-0.99
Single Parent of Infant	0.12	0.14	-0.02	-1.85
Single Parent of Child Under 5	0.49	0.45	0.04	2.15
Less than High School Equivalent	0.13	0.22	-0.09	-7.98
GED	0.13	0.12	0.01	0.59
High School Diploma	0.41	0.40	0.01	0.55
More than High School Equivalent	0.28	0.20	0.07	4.74
Education Data Missing	0.05	0.05	-0.01	-0.77
TANF Employment Barrier Identified	0.61	0.65	-0.04	-2.62
TANF Work Requirements Exemption	0.06	0.10	-0.04	-4.52
TANF Sanction	0.10	0.18	-0.08	-7.78
TANF Case Plan Re-engagement	0.37	0.46	-0.08	-5.10
Total Earnings, 3 Years Prior (Thousands)	20.53	20.58	-0.05	-0.06
Unemployed Full Year Prior to T=0	0.29	0.32	-0.03	-1.90
Total TANF Benefits, 3 Years Prior	4599.33	5287.70	-688.36	-4.41

Note: Data come from administrative UI earnings data from CDLE, administrative TANF data from CDHS, and fields collected by the TANF caseworker at TANF application. The treated sample includes STEP participants who enrolled in the program between January 2018 and March 2020, and the comparison sample includes other TANF recipients who received benefits during the same time period. A more detailed version is reported in Appendix Table A-2.

and the resulting potential selection bias motivate the use of propensity score reweighting as a component of our stacked difference-in-differences methodology.

4 Methodology

To estimate the causal effect of STEP participation on employment, earnings, and TANF receipt, we implement a propensity-score-weighted difference-in-differences design using the stacked difference-in-differences estimator characterized by Wing, Freedman and Hollingsworth (2024). We use this method to compare the trajectories of STEP participants to those of observably similar non-participants who received TANF benefits over the same time period. This method considers each potential enrollment month as a separate natural experiment with some observations treated and others remaining untreated. It then stacks these natural experiments into a single dataset and weights the portion from each enrollment month appropriately.

4.1 Stacked Difference-in-Differences Estimation

The stacked difference-in-differences estimation requires building separate panel datasets for each calendar month (m) during which at least one participant enrolled in STEP. Time is measured relative to the enrollment month so that $t = 0$ in month m. Observations in each panel include the subset of treatment group members who enrolled in m and a comparison group composed of all never-treated individuals who also received TANF benefits during that enrollment month, making them eligible to have enrolled in the STEP program during m. Each cohort's panel dataset includes only the relative time periods that are available for all STEP participants: 12 quarters prior to and 7 quarters after enrollment for earnings and employment outcomes, and 36 months prior to and 21 months after enrollment for TANF receipt. The cohort-level panel datasets are then appended ("stacked") to create the analysis dataset, in which individuals in the comparison group can be re-used in multiple

time periods, but each individual in the treatment group appears only once. Wing, Freedman and Hollingsworth (2024) explicitly account for this recycling of comparison units in defining the properties of the estimator and show that clustering standard errors at the person level is sufficient to deal with both the re-appearance of comparison units and the typical within-person correlation in untreated outcomes over time.

Stacking weights w_{im},

$$w_{im} = \begin{cases} \frac{N_m^{D=1}/N^{D=1}}{N_m^{D=0}/N^{D=0}} & \text{if } D_i = 0 \\ 1 & \text{if } D_i = 1 \end{cases}, \tag{1}$$

are constructed such that individual i who either enrolled in or was eligible to enroll in the STEP program in month m is weighted by cohort m's relative share of the treated observations in the analysis dataset divided by the relative share of comparison observations in the analysis dataset. The number of treated and comparison individuals in each cohort and the corresponding cohort weights w_{im} are reported in Appendix Table A-1. Balancing the dataset in event time ensures that newly-treated individuals are not erroneously compared to already-treated ones, which eliminates the well-documented problems with the two-way fixed effects strategy that was previously popular in staggered difference-in-differences designs (Goodman-Bacon, 2021). Reweighting by w_{im} ensures that each treated observation in the dataset receives equal weight and that the relative cohort weights are the same across the treatment group and the comparison group.

We use the stacked dataset to estimate the Average Treatment effect on the Treated (ATT) using a difference-in-differences regression, weighted by w_{im},

$$\begin{aligned} Y_{imt} = \beta_0 + \beta_1 D_i + \beta_2 InProgram_{mt} + \beta_3 (InProgram_{mt} * D_i) \\ + \beta_4 PostProgram_{mt} + \beta_5 (PostProgram_{mt} * D_i) + \epsilon_{imt}, \end{aligned} \tag{2}$$

where D_i is an indicator for STEP enrollment by individual i, m is the potential STEP en-

rollment month for a given cohort, and t is the relative quarter or month since the potential enrollment month, such that $t = 0$ in calendar month m. An individual i in the comparison group can have multiple observations for the same relative time period t and different potential enrollment months m. An individual i in the treatment group appears only once for each relative time period t. $InProgram_{mt}$ is an indicator denoting relative quarters $t =$ 0–3 (months $t = 0$–9) after (potential) enrollment in month m, during which most STEP participants are receiving services including subsidized placements and other supports; and $PostProgram_{mt}$ is an indicator for relative quarters $t = 4$–7 (months $t = 10$–21) when the vast majority of participants have exited the program. The coefficient β_3 therefore estimates the effect of participating in the STEP program in the first year post-entry while β_5 estimates the post-program-exit effect in the second year after entry.[22] We also estimate the analogous event study regression to reveal more detailed dynamic effects in each individual quarter (month) relative to STEP enrollment. We omit individual and cohort-level fixed effects, as Wing, Freedman and Hollingsworth (2024) show that they are unnecessary in light of the stacking and weighting procedure and can even re-introduce some of the weighting problems this procedure is intended to solve.

4.2 Adjustments for Selection on Observables

The comparison group is composed of TANF recipients who were receiving benefits at the same time as the treatment group and thus were eligible to enroll in STEP, but who did not enroll. Therefore the treatment and comparison groups have broadly similar demographics, such as age, gender, household structure, education, and work history (see Table 2). Receiving a monthly TANF payment also requires both the treatment and comparison groups to have similarly low current incomes in the month when the treatment group entered the program.

[22] Although the program allows for only six months of subsidized employment, the time from STEP enrollment to the beginning of the internship placement varies across participants. We err on the side of counting quarter three as an "in-program" period when in reality many participants exit by the end of quarter two.

To further increase the comparability of STEP participants and the comparison group, we reweight the comparison group within each cohort to match the treatment group's observables. We use probit regressions to estimate each TANF recipient's probability $\hat{p}_{im}$ of enrolling in STEP in each potential enrollment month m given their baseline characteristics measured in that month, including detailed earnings and TANF receipt for the past three years, age, completed education, gender, family structure, current enrollment in formal schooling, barriers to employment, and county.[23] We then re-scale each cohort's propensity scores $\hat{p}_{im}/(1-\hat{p}_{im})$ such that they sum to 1 within the comparison group for each enrollment month. We combine these re-scaled propensity scores with the original stacking weights w_{im} to construct covariate-adjusted stacking weights $\hat{pw}_{im}$,

$$\hat{pw}_{im} = \begin{cases} w_{im} * \frac{\hat{p}_{im}/(1-\hat{p}_{im})}{\sum_{\forall i \in m, D_i=0} \hat{p}_{im}/(1-\hat{p}_{im})} & \text{if } D_i = 0 \\ 1 & \text{if } D_i = 1, \end{cases} \tag{3}$$

such that the rescaled propensity scores reweight the comparison group only within each cohort m and do not change the relative weights across cohorts.

Appendix Figure A-3 presents the common support between the treatment and comparison groups. It indicates that for all values of $\hat{p}_{im}$ among STEP participants, there are many individuals in the comparison group with similar values. Access to the universe of potential program participants facilitates finding these close matches by providing a comparison group much larger than the treatment group. The substantial overlap in predicted enroll-

A small share —less than 10 percent—of participants receive services in quarter 4, and many of these have exit dates within the first month of quarter 4. Thus, a small portion of the estimated treatment effects in quarter 4 likely reflects "in program" effects. See Appendix Figure A-1 for details.

[23]Probit regressions are pooled within each potential enrollment quarter rather than estimated separately for each potential enrollment month to avoid perfectly predicting treatment status in months with very few STEP enrollees; see cohort-specific sample sizes in Appendix Table A-1 and probit results in Appendix Table A-4. Thus, an individual in the comparison group may have up to three observations in the same quarter-level propensity score estimation sample if they are enrolled in TANF for the entire quarter, but their predicted probability of STEP enrollment still varies across months in that quarter. Some individuals in the comparison group are assigned missing values of $\hat{p}_{im}$ because they differ from all treated observations in the quarter containing month m on some discrete characteristic. We allow these observations to be automatically excluded from the comparison group.

ment probabilities suggests that there are many non-participants whose earnings and benefit histories are sufficiently similar to those of each STEP participant, resulting in weighted average counterfactual outcomes that take advantage of the rich comparison pool while still ensuring similarity with the treated group.

4.3 Plausibility of Equal Counterfactual Trends Assumption

The stacked difference-in-differences design will correctly identify the causal effect of program participation under an equal counterfactual trends assumption — non-participants' change in outcomes must be exactly what would have occurred among the treatment group had they not participated in STEP. Because participation is voluntary, it is possible that the trends in untreated outcomes would have been different between the treatment and comparison groups had the treatment group not received services.

Table 3: Sample Characteristics, Reweighted Comparison Group

	(1) Mean (Trt.)	(2) Mean (Comp.)	(3) Diff.	(4) t (Diff.)
Age	33.21	33.25	-0.03	-0.11
Female	0.87	0.87	0.00	0.27
Only Adult in Household	0.85	0.85	0.00	0.10
Number of Children in Household	1.94	1.94	0.00	-0.07
Infant in Household	0.15	0.16	0.00	-0.33
Child Under 5 in Household	0.58	0.58	0.00	-0.21
Single Parent of Infant	0.12	0.12	0.00	-0.35
Single Parent of Child Under 5	0.49	0.49	0.00	-0.10
Less than High School Equivalent	0.13	0.12	0.00	0.11
GED	0.13	0.13	0.00	0.23
High School Diploma	0.41	0.40	0.00	0.10
More than High School Equivalent	0.28	0.29	-0.01	-0.64
Education Data Missing	0.05	0.05	0.00	-0.13
TANF Employment Barrier Identified	0.61	0.61	0.00	0.15
TANF Work Requirements Exemption	0.06	0.06	0.00	0.50
TANF Sanction	0.10	0.10	0.00	0.24
TANF Case Plan Re-engagement	0.37	0.37	0.00	0.08
Total Earnings, 3 Years Prior (Thousands	20.53	20.45	0.08	0.09
Unemployed Full Year Prior to T=0	0.29	0.29	0.00	0.10
Total TANF Benefits, 3 Years Prior	4599.33	4552.65	46.68	0.29

Note: Data come from administrative UI earnings data from CDLE, administrative TANF data from CDHS, and fields collected by the TANF caseworker at TANF application. The treated sample includes STEP participants who enrolled in the program between January 2018 and March 2020, and the comparison sample, reweighted by $\hat{p}_{im}$, includes other TANF recipients who received benefits during the same time period. A more detailed version is reported in Appendix Table A-3.

To address this concern, we provide two initial pieces of descriptive evidence that support the plausibility of the equal trends assumption. First, the outcome trends are remarkably similar for STEP participants and for the comparison group prior to month 0 even when using only the unadjusted stacking weights (see Figure 1 below). Second, Table 3 shows that the reweighting procedure successfully eliminates differences in observable baseline characteristics, which reduces concerns that differences in these variables may drive a post-treatment divergence in outcomes. The third column shows differences in means between the treatment and reweighted comparison group, and differences in all characteristics are minimal. No difference is statistically different from zero (see column 4) indicating that it is possible the equal counterfactual trends assumption may hold, as the treatment group and comparison group are balanced on characteristics that may affect post-enrollment employment and earnings.

Moreover, there are plausibly exogenous reasons why the treatment group chose to participate while the comparison group chose not to. Case managers report that STEP participation is determined primarily by the participant's interest in the program and the ability of the caseworker to find a suitable placement for the participant. There is likely some inherent randomness in how the program is presented to participants and in which possible placements were available to potential participants at the time they considered enrolling. On the other hand, recipients who are due to recertify their compliance with TANF work requirements may be particularly drawn to STEP, which could provide exogenous variation in program enrollment.[24]

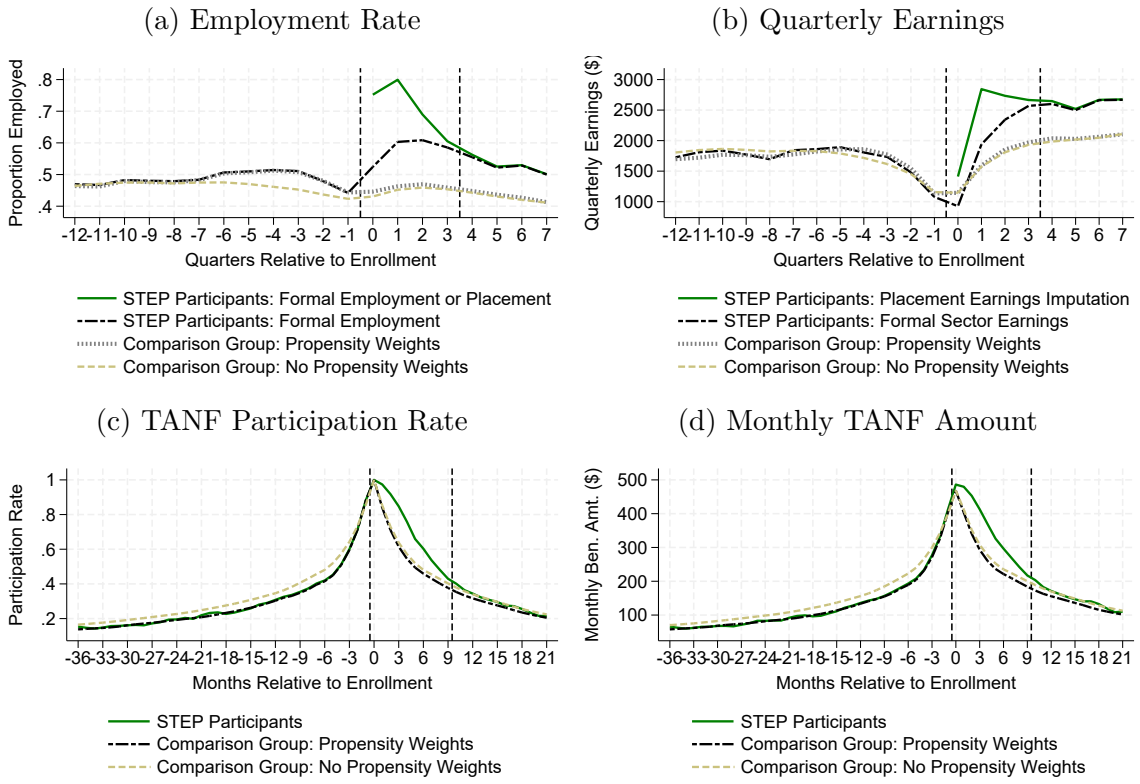
Figure 1 uses the stacked dataset to present trends over time in earnings, employment, and TANF receipt among STEP recipients, as well as a weighted average of the comparison group's outcomes in the same relative time period. The horizontal axis is measured relative

[24]Recall that the condition for inclusion in the comparison group in a given enrollment month m is current TANF receipt in that month. This requirement means that TANF beneficiaries eligible to enroll in STEP in m have varying recertification dates based on when their TANF spell started or when they last recertified their compliance with work requirements. We intentionally exclude variables related to TANF spell start date from the estimation of the propensity score to preserve this variation.

to each participant's STEP enrollment date. Quarter or month zero indicates the enrollment period; negative values indicate time periods prior to enrollment, and positive values indicate periods after enrollment. In panels (a) and (b) the black dashed line provides trends for the treatment group using only UI earnings records. The solid green line, which begins in period 0 by definition because no one has earnings from an internship prior to enrollment, includes imputed earnings from STEP placements as described in Section 3.1. In panels (c) and (d), the trends for the treated group are shown with a solid green line. All four panels include two versions of the comparison group's trend lines to highlight the value of reweighting the comparison group to match baseline obervables. The first (black dashed line, main propensity weights) reweights using propensity scores estimated on the full set of available baseline characteristics (see Appendix Tables A-3 and A-4 for further details). The second (gold dashed line, no propensity weights) does not adjust for any observable characteristics and uses only the cohort stacking weights described in Equation 1.

The treatment group experiences notable decreases in employment and earnings (panels a and b), along with corresponding increases in TANF (panels c and d) in the 6–9 months prior to period 0. This timing is consistent with the fact that the typical STEP enrollee had been receiving TANF payments for six months prior to enrollment (see Table 1). Importantly, these trends are broadly shared by members of the comparison group, and this lack of differential pre-trends supports the equal counterfactual trends assumption. Further, reweighting results in even closer-matched pre-trends. Despite similar pre-trends, it is possible that, even in the absence of the program, trends in untreated outcomes would have diverged after the treatment group enrolled. The trends prior to period 0 demonstrate that both groups experienced negative shocks just prior to when the treatment group enrolled in STEP, and some improvement in employment and earnings was likely to occur among both groups as they recovered from those shocks. If STEP participants have characteristics that make their labor market recovery more likely or quicker, then difference-in-differences estimation of the ATT could still be biased despite parallel pre-trends. Reweighting the

Figure 1: Outcome Trends



Notes: Data come from administrative UI earnings data from the CUBS data system, administrative TANF data from CDHS, and baseline survey data collected at application. Each quarterly sample includes 912 STEP participants who applied between January 2018 and March 2020. Quarter or Month 0 represents the quarter or month in which an individual enrolled in STEP and is thus a different calendar quarter from person to person. Formal employment is defined as having UI-covered earnings in Colorado greater than $0 in a given quarter. When earnings from a STEP-sponsored transitional job are reported to the UI system, they are included in formal sector employment/earnings. The imputation for placement-based earnings not reported to the UI system is described in section 3.1. TANF receipt is defined as receiving any payment in a given month. Average quarterly earnings and monthly TANF amounts include zeroes.

comparison group to match the average characteristics of the treatment group mitigates this potential concern.

Appendix Tables A-2 and A-3 provide the full set of baseline characteristics used to reweight the comparison group and demonstrate that reweighting almost completely eliminates treatment/comparison differences in observable baseline characteristics. Previous work by Grosz (2020) finds that conditioning on detailed earnings histories is sufficient to identify the effects of a large community college program in observational data and yields estimates similar to those from randomized lotteries. Similarly, Andersson et al. (2022) find that conditioning on characteristics beyond earnings histories does not affect estimates of the effects

of WIOA-sponsored job training programs. Because our conditioning set contains similarly detailed earnings histories as well as detailed benefits histories (given that TANF receipt is the primary STEP eligibility criterion), it is plausible that our analysis also identifies the causal effect of STEP participation.

5 Results

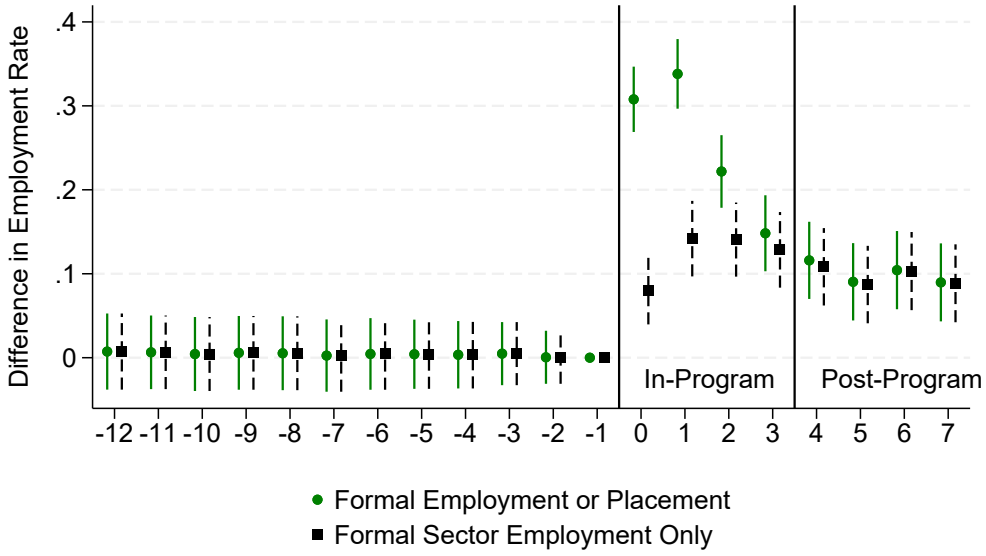
5.1 Effects on Employment and Earnings

5.1.1 Main Effects

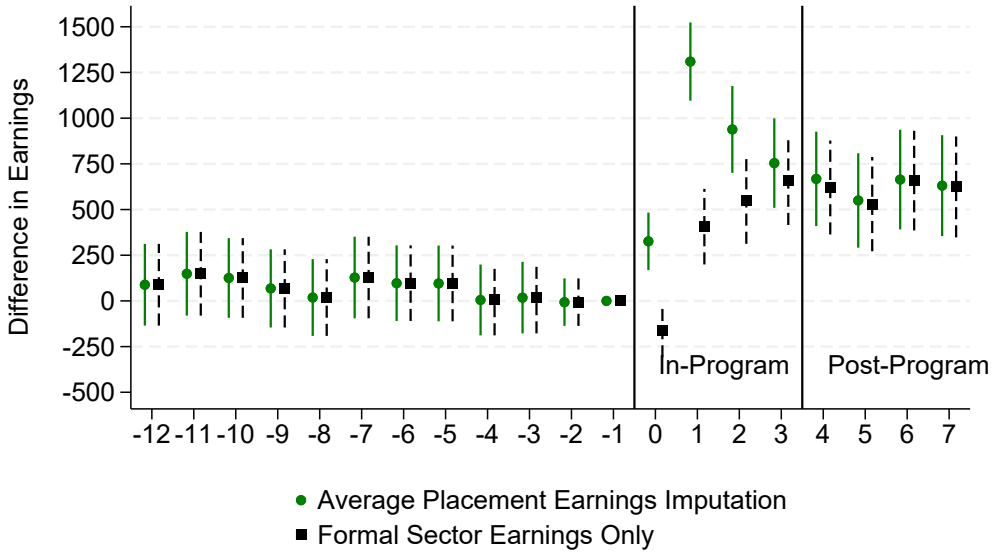
We begin by presenting estimates from an event-study version of Equation 2 in Figure 2, using time period (-1) as the reference period. Each displayed coefficient is therefore the difference in treatment/comparison differences between the period listed on the horizontal axis and period (-1). For completeness, we provide estimated treatment effects using only formal-sector earnings (black squares) in addition to the main estimates imputing employment and earnings for individuals whose STEP placements do not appear in the UI earnings data (green circles).

Figure 2: Event-Study Effects on Employment and Earnings

(a) Employment Rate



(b) Quarterly Earnings



Notes: Data source is administrative UI earnings data from CDLE. Panel A and B report estimates on employment and earnings outcomes, respectively, for STEP participants who enrolled between January 2018 and March 2020. Quarter 0 represents the quarter in which an individual enrolled in STEP and is thus a different calendar quarter from person to person. Formal employment is defined as having UI-covered earnings in Colorado greater than $0 in a given quarter. When earnings from a STEP-sponsored transitional job are reported to the UI system, they are included in formal sector employment/earnings. The imputation for placement-based earnings not reported to the UI system is described in section 3.1.

The first panel indicates that, relative to changes among similar TANF participants in the comparison group, the treatment group's employment rate rose by approximately 30 percentage points more during the initial enrollment quarter. Program effects then rise to nearly 35 percentage points in quarter 1 before gradually falling through quarter 3. In the post-program period, program effects on quarterly employment stabilize at roughly 9–11 percentage points.

Similarly, the second panel demonstrates that, beginning in quarter 1 and continuing through quarter 7, the estimated treatment effects on quarterly earnings are large and positive—more than $1250 in the first quarter after enrollment and slowly decreasing to about $500-700 per quarter after program exit.

Table 4 provides estimated treatment effects on employment and earnings for the aggregate time periods while participants received services (quarters 0–3) and after participants exited (quarters 4–7) using the specification in Equation 2. Our preferred estimates in columns 1 and 3 impute missing STEP placement earnings based on the program tracking data, while columns 2 and 4 use only formal-sector UI earnings. The preferred specifications indicate that program participation increased participants' quarterly employment rate by 25 percentage points (51 percent relative to the comparison group baseline mean of 0.48) and their formal sector earnings by about $766 per quarter (44 percent of the comparison baseline mean of $1704) while enrolled. Although effects diminish after program exit, substantial gains persist through quarters 4–7: during this period, the program raised employment by about 10 percentage points (18 percent) and earnings by roughly $563 per quarter (32 percent).

Table 4: Aggregated Effects on Employment and Earnings

	Employment		Earnings	
	Imputed (Placements)	Formal Only	Imputed (Placements)	Formal Only
	(1)	(2)	(3)	(4)
In-Program (Q0-Q3)	0.250***	0.119***	766.4***	297.1***
	(0.0150)	(0.0156)	(87.79)	(87.12)
Post-Program (Q4-Q7)	0.0960***	0.0927***	562.6***	542.4***
	(0.0173)	(0.0174)	(115.0)	(114.8)
Comparison Mean (Q < 0)	0.483	0.483	1704.0	1704.0
Comparison Mean (Q0-Q3)	0.459	0.459	1640.2	1640.2
Comparison Mean (Q4-Q7)	0.431	0.431	2058.0	2058.0
N (Individuals)	21740	21740	21740	21740

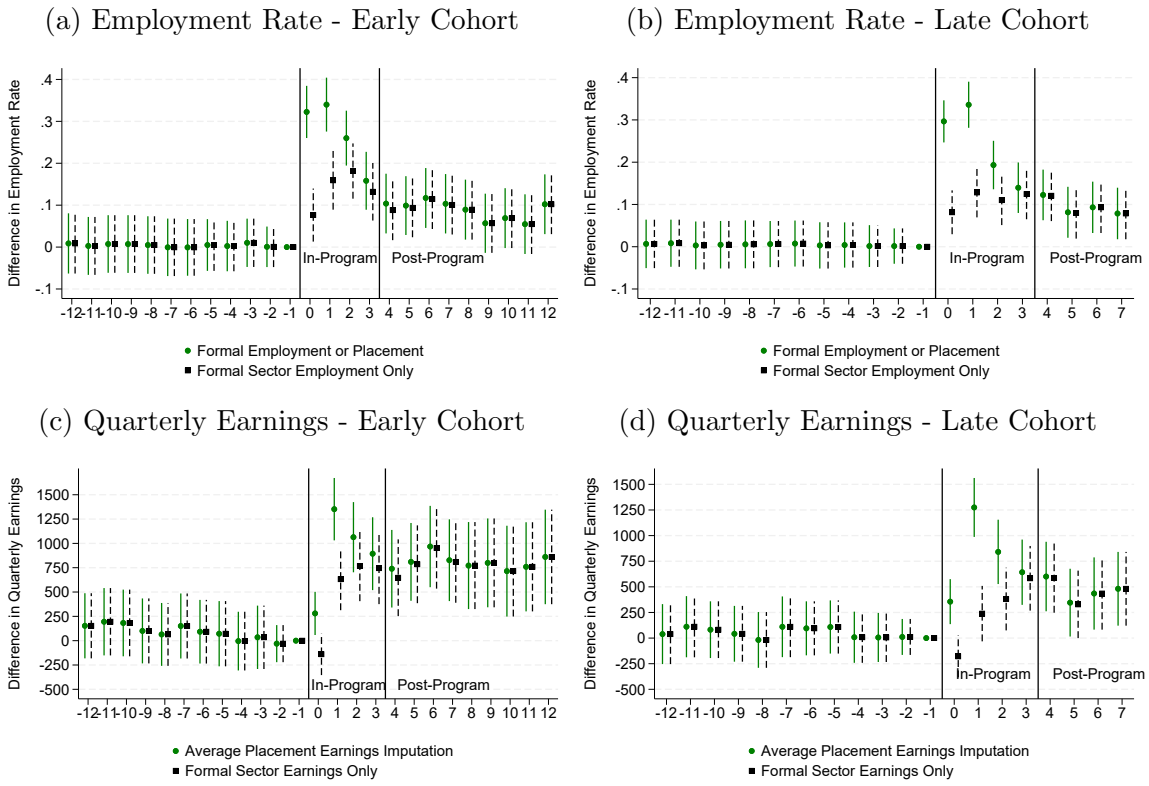
Notes: Data source is administrative UI earnings data from CDLE. Columns 1-3 and 4-6 report estimates on employment and earnings outcomes, respectively, for STEP participants who enrolled between January 2018 and March 2020. Quarter 0 represents the quarter in which an individual enrolled in STEP and is thus a different calendar quarter from person to person. Formal employment is defined as having UI-covered earnings in Colorado greater than $0 in a given quarter. When earnings from a STEP-sponsored transitional job are reported to the UI system, they are included in formal sector employment/earnings. The imputation for placement-based earnings not reported to the UI system is described in section 3.1.

5.1.2 Effects Up to 12 Quarters

To assess longer-term impacts, we use a subset of earlier enrollees with longer available follow-up data—more than two years after most participants exited the program—shown in Figure 3. This "early" cohort includes all participants who enrolled in STEP between January and December 2018 and have at least 12 quarters of post-enrollment data. The "late" cohort comprises those who enrolled between January 2019 and March 2020 and, like the full sample, have seven quarters of follow-up. Remarkably, earnings effects do not fade and may even increase slightly for the early cohort through the 12th quarter post enrollment (panels a and c). Specifically, on average across quarters 8 to 12, STEP increased employment by 7 percentage points and quarterly earnings by $697.

Participants in the late enrollment cohort (panels b and d) likely experienced COVID-19 pandemic disruptions to the labor market during or shortly after their participation in the program, potentially attenuating their post-program effects on employment and earnings. To further understand how effect sizes may differ as a result of the pandemic, we present aggregated effects by enrollment timing in Appendix Table A-6. Employment effects are qualitatively similar across the two cohorts. Earnings effects are larger in the early compared

Figure 3: Event-Study Effects on Employment and Earnings, by Cohort



Notes: This figure provides analysis comparable to the specifications in Figure 2 for subsamples of observations based on program entry dates. Panels a and c report estimates on employment and earnings outcomes, respectively, for the "Early Cohort" of STEP participants who enrolled between January 2018 and December 2018. Panels b and d report employment and earnings for the "Late Cohort" who enrolled between January 2019 and March 2020. Quarter 0 represents the quarter in which an individual enrolled in STEP and is thus a different calendar quarter from person to person. Formal employment is defined as having UI-covered earnings in Colorado greater than $0 in a given quarter. When earnings from a STEP-sponsored transitional job are reported to the UI system, they are included in formal sector employment/earnings. The imputation for placement-based earnings not reported to the UI system is described in section 3.1. Data source is administrative UI earnings data from CDLE.

to late cohort for both in-program and post-program quarters – $813 (47 percent) versus $729 (43 percent) and $752 (43 percent) versus $416 (25 percent), respectively. However, due to the relatively small number of treated observations in each cohort, these differences are not statistically different from each other (see Appendix Table A-7).

5.1.3 Robustness

While the weighted difference-in-differences design ensures that the treatment group is observably similar to the comparison group at baseline and there is evidence that the parallel trends assumption may hold, it is possible that the treatment group is still positively se-

lected on unobserved factors such as willingness or ability to find new employment. In this case, STEP participants would have experienced stronger labor market recoveries after their negative earnings shocks even without participating in the program. Fixed differences in characteristics like motivation and work-readiness are likely sufficiently correlated with earnings histories in the three years leading up to the potential enrollment month, but these factors may also vary based on time-varying shocks that could be correlated with STEP enrollment. To address this concern, we compare STEP participants to two different subsets of the non-participating TANF caseload that are likely more work-ready than the full comparison group used in the main analysis. We select these subsets based on variables recorded by the TANF caseworker during the current TANF spell up to and including the potential enrollment month so that they reflect relatively current assessments of factors that may be correlated with the individual's potential earnings trajectory. All analyses using these subsets of the comparison group use the main imputed employment and earnings measures from columns 1 and 3 of Table 4.

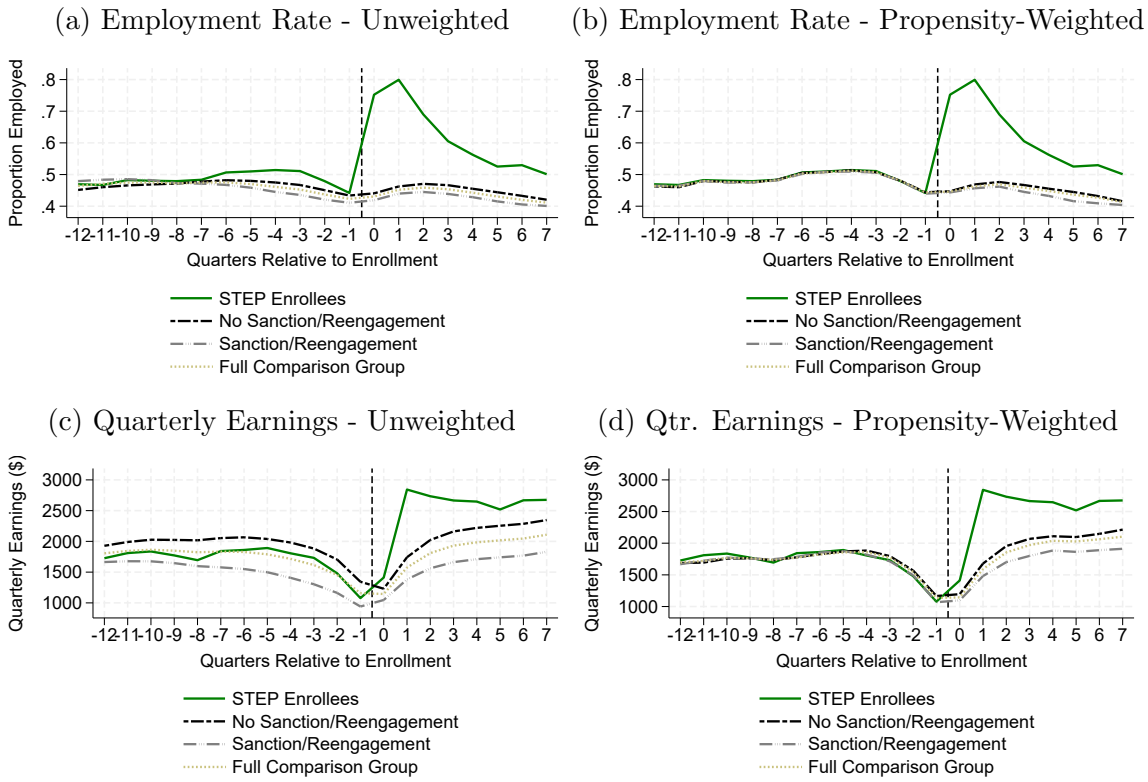
The first comparison subgroup excludes individuals who had a TANF sanction or re-engagement designation due to non-compliance with their case plan during the current TANF spell up to and including the potential STEP enrollment month. TANF sanctions result in payment reductions for failing to participate in activities set out in the case plan to satisfy work requirements, and a case plan re-engagement is typically an intermediate step that can ultimately lead to a sanction. As indicated in Appendix Table A-2, the treatment group was less likely than the comparison group to have either of these designations, but nevertheless 10% had a sanction and 37% had a re-engagement designation during the TANF spell in which they enrolled in the STEP program, and we keep such individuals in the treatment group sample. By restricting only the comparison group in this manner, we are comparing the treatment group to non-participants who are arguably more positively selected on motivation and willingness to work.

Figure 4 demonstrates that the unweighted employment rate for this new comparison

group is nearly as high as the treatment group before the enrollment month (panel a), and their unweighted average earnings are higher (panel c). Even after reweighting such that the employment and earnings trends of individuals with and without sanctions/re-engagement are similar to each other and to those of STEP participants in the pre-period, those without a sanction/re-engagement work more (panel b) and earn more (panel d) than the rest of the TANF caseload in the post-enrollment periods.[25] In other words, having no sanction/re-engagement in the TANF spell leading up to the potential enrollment month is predictive of higher post-enrollment-month employment and earnings even after conditioning on labor market histories. However, STEP participants have higher post-enrollment employment and earnings compared to even this likely more motivated group.

[25]Corresponding event-study estimates comparing STEP participants to non-participating individuals with no sanction or re-engagement designation in the current TANF spell are presented in Figure A-7.

Figure 4: Outcome Trends by TANF Sanction/Re-engagement Status



Notes: Data sources are administrative UI earnings data and STEP program tracking data from CDLE. All employment and earnings measures for STEP Enrollees include imputations for STEP placements that do not appear in the UI earnings data. Panel A and B report estimates on employment and earnings outcomes, respectively, for STEP participants who enrolled between January 2018 and March 2020. Quarter 0 represents the quarter in which an individual enrolled in STEP and is thus a different calendar quarter from person to person. Formal employment is defined as having UI-covered earnings in Colorado greater than $0 in a given quarter. When earnings from a STEP-sponsored transitional job are reported to the UI system, they are included in formal sector employment/earnings. The imputation for placement-based earnings not reported to the UI system is described in section 3.1.

In fact, aggregated difference-in-differences estimates in columns 2 and 3 of Table 5 comparing STEP participants only to those with no sanction or re-engagement are similar to the main point estimates (Column 1) in sign, magnitude, and significance, regardless of whether propensity weights are used. To the extent that unobserved motivation to work is correlated with a TANF recipient's likelihood of having a recent sanction or re-engagement designation in their case notes, these estimates suggest that there is no substantial bias in our estimates of program impacts due to differences in motivation.

In addition to willingness or motivation to find employment, TANF recipients vary in their ability to sustain employment due to various barriers such as child care, transportation,

Table 5: Aggregated Effects on Employment and Earnings vs. Positively Selected Comparison Groups

	(1)	(2)	(3)	(4)	(5)
Panel A: Employment					
In-Program (Q0-Q3)	0.250***	0.232***	0.244***	0.214***	0.195***
	(0.0150)	(0.0141)	(0.0158)	(0.0143)	(0.0197)
Post-Program (Q4-Q7)	0.0960***	0.0714***	0.0903***	0.0753***	0.0554*
	(0.0173)	(0.0164)	(0.0183)	(0.0166)	(0.0223)
Comparison Mean (Q < 0)	0.483	0.465	0.483	0.489	0.486
Comparison Mean (Q0-Q3)	0.459	0.460	0.465	0.501	0.518
Comparison Mean (Q4-Q7)	0.431	0.438	0.437	0.458	0.475
Comparison Group	Main	No Sanction/Reeng.	No Sanction/Reeng.	No Barrier/Exempt.	No Barrier/Exempt.
Propensity Weights	Yes	No	Yes	No	Yes
N (Individuals)	21740	20813	16532	16147	12663
Panel B: Quarterly Earnings					
In-Program (Q0-Q3)	766.4***	835.3***	688.7***	769.3***	488.0***
	(87.79)	(85.15)	(90.89)	(87.14)	(107.6)
Post-Program (Q4-Q7)	562.6***	561.8***	484.4***	564.8***	296.5*
	(115.0)	(111.5)	(118.7)	(114.3)	(144.5)
Comparison Mean (Q < 0)	1704.0	1921.7	1710.4	2033.9	1758.2
Comparison Mean (Q0-Q3)	1640.2	1789.1	1724.3	1967.2	1972.8
Comparison Mean (Q4-Q7)	2058.0	2276.5	2142.5	2385.6	2378.3
Comparison Group	Main	No Sanction/Reeng.	No Sanction/Reeng.	No Barrier/Exempt.	No Barrier/Exempt.
Propensity Weights	Yes	No	Yes	No	Yes
N (Individuals)	21740	20813	16532	16147	12663

Notes: Data sources are administrative UI earnings data and STEP program tracking data from CDLE. All employment and earnings measures for STEP Enrollees include imputations for STEP placements that do not appear in the UI earnings data (see section 3.1). Panels A and B report estimates on employment and earnings outcomes, respectively, for STEP participants who enrolled between January 2018 and March 2020 relative to various comparison groups: the full (reweighted) comparison group in Column 1, those with no sanction/re-engagement designation in the current TANF spell in Columns 2-3, and those with no work barrier or work requirements exemption in the current TANF spell in Columns 4-5. Quarter 0 represents the quarter in which an individual enrolled in STEP and is thus a different calendar quarter from person to person. Formal employment is defined as having UI-covered earnings in Colorado greater than $0 in a given quarter. Sample sizes differ across propensity scores because some individuals in the comparison group are assigned missing values of the propensity score because they differ from all treated observations in their cohort on at least one discrete characteristic included in the propensity score estimation. We allow these observations to be automatically excluded from the comparison group depending on which propensity score is used.

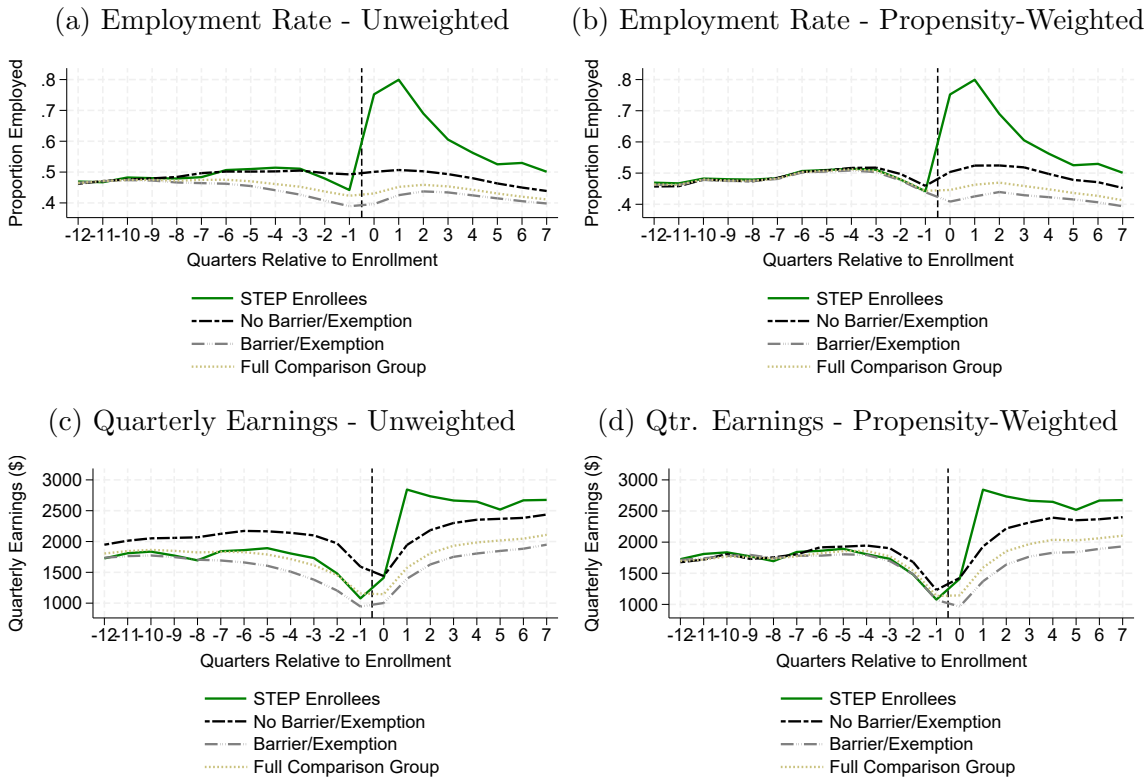
or housing difficulties. These barriers are often noted by the caseworker, and mitigation efforts are included in the recipient's case plan. Some work barriers — such as domestic violence, disability, or caregiving responsibilities for a disabled child or family member — result in the recipient receiving an exemption from TANF work requirements. As shown in Appendix Table A-2, 61% of the treatment group (vs. 65% of the comparison group) had an employment barrier identified by their caseworker during the TANF spell of their STEP enrollment, and 6% (vs. 10% of the comparison group) were exempt from work requirements. These barriers are also measured in the current TANF spell up to and including the enrollment month, providing a reasonably current summary of recent circumstances that may affect a participant's ability to find or maintain stable employment.

The second comparison subgroup uses only individuals without work barriers who do not have an exemption from work requirements. The majority of the treatment group face some type of employment barrier, so using this alternative comparison group addresses, and arguably over-corrects for, any remaining differences in work readiness between STEP participants and non-participants. Figure 5 shows outcome trends, with the comparison group split by the presence/absence of a work barrier/work requirements exemption. Those with no barrier or exemption are as likely as the treatment group (and far more likely than the rest of the non-participating TANF caseload) to be employed (panel a) up until two quarters prior to the treatment group's enrollment in STEP, and earn more on average (panel c). In the two quarters preceding STEP enrollment, they experience a dip in earnings (panel c) similar to that of the treatment group but do not experience a corresponding dip in employment (panel a). Even after reweighting to align pre-trends in employment (panel b) and earnings (panel d) across TANF recipients with and without barriers or exemptions, those without barriers/exemptions have higher post-enrollment employment rates and earnings than the remaining members of the comparison group. This pattern suggests that the absence of barriers/exemptions is a strong predictor of TANF recipients' ability to secure unsubsidized employment on their own. Nevertheless, STEP participants surpass this highly work-ready group in employment immediately after enrollment and eventually exceed their earnings as well.[26]

Column 3 of Table 5 indicates that, without reweighting on earnings histories and baseline characteristics, difference-in-differences estimates of STEP participation on employment and earnings compared to this group are similar to the main estimates (Column 1). After both restricting to this subset of the comparison group and reweighting (Columns 4 and 6), point estimates decrease in magnitude. The post-program increase in employment among the treatment group — 61% of whom have an employment barrier — is about 6 percentage points (11 percent) larger ($p < .05$) than the change among those with no employment

[26]Corresponding event-study estimates are presented in Figure A-8.

Figure 5: Outcome Trends by Work Barrier/Work Requirements Exemption Status



Notes: Data sources are administrative UI earnings data and STEP program tracking data from CDLE. All employment and earnings measures for STEP Enrollees include imputations for STEP placements that do not appear in the UI earnings data. Panel A and B report estimates on employment and earnings outcomes, respectively, for STEP participants who enrolled between January 2018 and March 2020. Quarter 0 represents the quarter in which an individual enrolled in STEP and is thus a different calendar quarter from person to person. Formal employment is defined as having UI-covered earnings in Colorado greater than $0 in a given quarter. When earnings from a STEP-sponsored transitional job are reported to the UI system, they are included in formal sector employment/earnings. The imputation for placement-based earnings not reported to the UI system is described in section 3.1.

barrier who have similar employment and earnings ex ante. Analogous earnings effects are an increase of $297 per quarter ($p < 0.05$). Again, these estimates likely represent an over-correction for potential positive selection on general work readiness, so we consider these our most conservative estimates of the program's impact.

It is important to note that the characteristics defining each of these positively selected alternative comparison groups are included as predictors of program enrollment in the propensity score used to produce the main estimates. The analysis presented in this section suggests that these characteristics are predictive of not only program participation but also post-enrollment-month employment and earnings trajectories. These two relationships therefore

suggest that reweighting the comparison group to have similar motivation and barriers provides additional support for the equal counterfactual trends assumption beyond the equal pre-trends.

Additional robustness analyses are reported in Appendix Section A.5. First, Appendix Table A-8 shows that alternate weighting schemes produce results nearly identical to the main estimates in sign and magnitude. We include results from using a parsimonious propensity score based on only employment and earnings histories as in Andersson et al. (2022) as well as results using only the stacking weights.

Second, Appendix Table A-9 shows that participants with unusually long program participation do not contribute substantially to the estimated post-program persistence of employment and earnings impacts because the effects are at least as large when excluding the relatively few STEP participants who remain enrolled in the program past quarter 3.

Finally, estimated post-program impacts are qualitatively similar regardless of whether in-program earnings appear in UI data. In some counties, earnings from STEP placement nearly always lead to UI earnings, while in others they appear inconsistently, likely due to different reimbursement practices across program contractors. Appendix Table A-5 and Figure A-6 show that post-program employment effects are nearly identical across the two sets of counties, but there is some heterogeneity in earnings effects, likely reflecting differences in prevailing wages across local labor markets.

5.1.4 Heterogeneity

To understand for whom the program may be most effective and potentially inform future targeting, we also use interaction specifications to examine potential heterogeneous treatment effects based on baseline characteristics in Appendix Table A-10. We find that participants who had no formal sector employment in the year prior to their enrollment in the STEP program experienced larger treatment effects compared to participants with at least some employment in the prior year. Estimated program effects on employment are 19 percentage

points larger during the in-program period and 10 percentage points larger in the post-program period, with both differences being statistically significant. Earnings effects are also statistically significantly larger for participants without recent work history in both the in-program ($875 larger) and post-program periods ($845 larger).

Effects on employment and earnings are slightly larger but not significantly different for participants whose TANF caseworker identified a barrier to employment, such as housing, transportation, childcare, or medical barriers. Finally, effects are slightly smaller for participants who have a child under age five in the household compared to effects for those who do not, but these differences are not statistically significant. Overall, these results suggest that the program may be especially effective for workers who have been out of work for longer, although it appears to have substantially positive and durable impacts across worker subtypes.

5.2 Mechanisms

The STEP program, like its sister program ReHire Colorado (Barham, Cadena and Turner, 2023), was intended to improve participants' long-term outcomes by helping workers gain skills and recent work experience and by creating a low-cost, low-risk way for host-site employers to learn the match quality of workers they may not otherwise hire outright. Barham, Cadena and Turner (2023) provide descriptive analysis showing that nearly all of the gains from ReHire participation accrued to participants who were eventually hired by their host-site employers. Therefore, one possible explanation for the success of the STEP program is that some STEP participants and their host sites also discovered productive matches during their time of subsidized employment, leading to permanent job offers.

While program records do not allow us to observe directly whether STEP participants are hired by their host-site employers, those converted to permanent positions at the host site at the conclusion of their internships should have more stable employment with the same employer throughout the year after their exit from the program (quarters 4 through 7

Table 6: Stability of Post-Program Employer-Employee Matches

	(1) Quarters Employed Q4-Q7	(2) Max. Quarters with Same Employer Q4-Q7	(3) Employed All of Q4-Q7	(4) Same Employer All of Q4-Q7
STEP Enrollee	0.381***	0.319***	0.0549***	0.0363*
	(0.0594)	(0.0546)	(0.0165)	(0.0143)
Comp. Mean	1.724	1.501	0.256	0.164
N (Individuals)	21740	21740	21740	21740

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Data source is administrative UI earnings data from CDLE. All results come from cross sectional regressions at the person-stacking month level of each outcome on a STEP enrollment dummy, weighted using the main set of stacked propensity weights.

after enrollment).[27] Using a stacked cross-sectional regression and the main set of stacked propensity weights defined in Equation 3, we test whether STEP participants are more likely than non-participants to be consistently employed throughout the post-program period, and whether this difference is driven by continuous employment with a single employer. This analysis focuses on the post-program period when most STEP enrollees have exited their placements and uses only the UI earnings data, as employer identities are not available in the STEP program tracking data.

Results in Table 6 indicate that STEP participants are about 5 percentage points (20 percent) more likely than non-participants to have formal-sector employment in every quarter from quarters 4 through 7 after enrollment. Further, they are about 3 percentage points (21 percent) more likely to be employed by the same employer in every quarter. Similarly, they are employed for roughly 0.4 more quarters (21 percent more) on average during this period, and their longest match with a single employer lasts, on average, about 0.3 quarters (20 percent) longer.[28] This analysis provides suggestive evidence that, as in ReHire (which was programmatically similar but served a different population), STEP participants are more likely to have found a durable employer-employee match as a result of their experience in the program.

[27]Recall that the local contractor is often the employer of record during the subsidized job and that many internships are not captured in UI earnings records at all. As a result, we cannot use the UI data to determine which participants were later hired by their host site. Unfortunately, the administrative data tracking STEP participants also does not allow us to make this determination.

[28]Columns 1 and 3 are effectively a single-difference (cross-sectional) version of the main employment results in Column 3 of Table 4. Results are of a similar magnitude, likely because outcomes prior to STEP enrollment are similar in levels as well as trends (see Figure 1).

Table 7 uses an industry-based measure of employment stability to demonstrate that STEP participants end up employed in more stable sectors relative to the comparison group in the year after their exit from the program. The table lists the seven most common Two-Digit NAICS codes for jobs held by the treatment group in quarters 4–7, and they are ranked by the average duration of an employment spell in that industry using spells from the comparison group (column 1).[29] The ranking is the same when using an alternative stability measure based on the probability of a job lasting at least one year (column 2). Columns 3 and 4 provide the share of the treatment and comparison groups, respectively, who have any employment in each of these sectors in quarters 4–7, and the final two columns provide the difference in means and the relevant t-statistic.[30] The results show that STEP participants are more likely to work in health care and social assistance, education, and the public sector, which are among the most stable sectors. They are also at least as likely as the comparison group to be employed in medium- to low-stability industries such as retail and administrative support services, but are less likely to be employed in the higher-turnover accommodation and food services industry. Therefore, at least some STEP participants appear to have experienced improved job stability through increased rates of employment in more stable industries. We note that this measure does not capture any changes in within-industry job stability, which may also have improved.

[29]Less common sectors cannot be separately identified due to the terms of the data sharing agreement.

[30]This analysis also uses only the UI earnings data, as NAICS codes are not available in the STEP program tracking data.

Table 7: Post-Program Employment by Sector, Quarters 4-7

Sectoral Job Stability Measures		**Sector (2-Digit NAICS)**	**Share with Any Job in Sector**			
Avg. Duration	Prop. Last $\geq$ 1 Year		Treat.	Comp.	$\hat{pw}_{im}$ Weighted Diff.	t (Diff.)
(1)	(2)		(3)	(4)	(5)	(6)
2.52	0.35	Public Administration (92)	0.061	0.010	0.052	6.32
2.44	0.26	Educational Services (61)	0.036	0.022	0.014	2.14
2.07	0.19	Health Care and Social Assistance (62)	0.241	0.159	0.082	5.47
2.01	0.17	Other Services except Public Admin (81)	0.038	0.032	0.007	1.01
1.85	0.14	Retail Trade (44-45)	0.160	0.142	0.018	1.35
1.76	0.11	Accommodation and Food Services (72)	0.123	0.148	-0.025	-2.07
1.65	0.08	Admin, Support and Waste Mgmt Services (56)	0.184	0.153	0.031	2.32
1.94	0.16	All Other Sectors	0.202	0.186	0.016	1.13

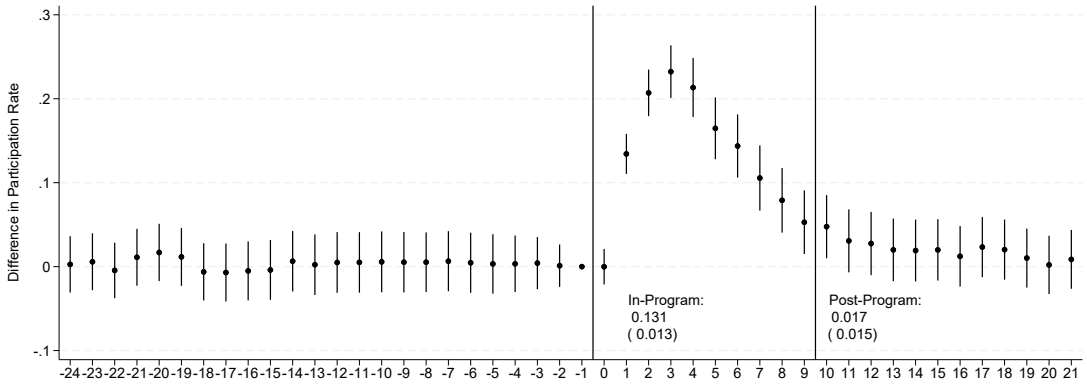
Notes: Data source is administrative UI earnings data from CDLE. Columns (1) and (2) provide descriptive estimates of job duration in quarters and proportion of jobs lasting the full year in the comparison group. Columns (3) and (4) report the proportion of each group with any employment in the relevant sector in the post-program year, and column (5) reports the difference between the two; these three columns, and the t-statistics in column (6), are estimated in cross sectional regressions at the person-stacking month level of industry employment dummies on a STEP enrollment dummy, weighted using the main (Full Covariate Set) set of stacked propensity weights. Sectors in which less than 3.3% of the treatment group is employed in Quarters 4-7 are combined in the final row, as they would create cells representing less than 30 individuals and cannot be considered de-identified.

5.3 Effects on TANF Receipt

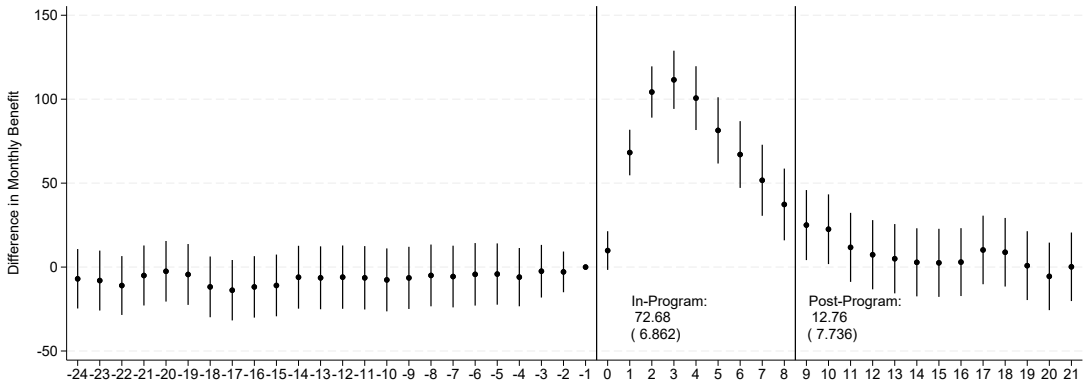
In addition to affecting employment and earnings, STEP may have important effects on participants' ability and willingness to access TANF benefits and/or their decision to exit the TANF caseload. Event studies in Figure 6 show that STEP participation increases enrollees' TANF receipt relative to non-enrollees starting in the first month after enrollment, with the difference growing over the first few months in the program and peaking in month 3 at about 25 percentage points. After month 3, this gap begins to narrow as TANF exit becomes more common among the treatment group. Like differences in rates of TANF receipt, differences in benefit amount peak in months 2–3 after enrollment and then gradually decline, becoming negligible by the time most participants have exited the STEP program.

Figure 6: Event- Study Effects on TANF and SNAP Receipt

(a) TANF Participation Rate



(b) Monthly TANF Amount



Notes: Data source is administrative TANF benefits data from the CBMS data system. Figure reports event study estimates on a binary outcome for TANF receipt, respectively, for STEP participants who enrolled between January 2018 and March 2020. Month 0 represents the month in which an individual enrolled in STEP and is thus a different calendar month from person to person. TANF receipt is defined as having any TANF payment in Colorado greater than $0 in a given month. Earnings from a STEP-sponsored transitional job do not count against the TANF eligibility threshold and in fact require the individual to remain enrolled in TANF during the transitional job.

We also report aggregated estimates for the in-program (months 0-9) and post-program (months 10-21) periods, estimated by Equation 2 and reported in the body of Figure 6. STEP participation increases participants' likelihood of TANF receipt in the in-program period by roughly 46 percent (13 percentage points), resulting in, on average, an additional $73/month in TANF cash assistance relative to the comparison group. In contrast, there is no statistically or economically significant program effect on post-program TANF receipt.

The in-program impacts on TANF receipt likely primarily reflect program take-up—as noted in Section 2, STEP program rules require participants to be enrolled in TANF until

they exit STEP, whereas a non-STEP-participant needs only to be enrolled in TANF in month 0 in order to be included in a given enrollment cohort's comparison group.

The increases in post-program employment and earnings seen in Table 4 do not seem to have resulted in declines in post-program benefit receipt relative to the comparison group. Notably, this lack of a differential is because both groups experienced a substantial decline in benefit receipt after month 0 (see Figure 1). It is possible that the declines among the treatment group are due to the increased earnings while the declines among the comparison group represent a failure to receive benefits for which they are income-qualified and/or strategic conservation of future TANF eligibility with respect to the 60-month lifetime limit.

6 Cost-Effectiveness of STEP

We next use a simplified version of the Marginal Value of Public Funds (MVPF) framework to quantify the return on investment of STEP program funds, and we find that the improvements in employment and earnings are large enough to make the program very cost-effective compared to other programs that target similar populations. The MVPF measures the shadow price of reallocating funds from one program to another in a revenue-neutral substitution (Hendren and Sprung-Keyser, 2020), and it is calculated as participants' willingness to pay for the program's benefits divided by the net costs of the program to the government. An MVPF of 1 indicates that a program is as cost-effective as a non-distortionary cash transfer. Participants' willingness to pay is calculated as the discounted after-tax increase in earnings, and the costs to the government are the total program expenditures divided by the number of participants served, including those who were not placed into a subsidized job. Rather than projecting participants' earnings over the life cycle as is typical in MVPF analysis, we conduct a simplified, back-of-the-envelope MVPF calculation assuming constant treatment effects over a much shorter time horizon. More details on this calculation and the underlying assumptions are provided in Appendix Section A.6.

Table 8: Marginal Value of Public Funds

	(1) With Placement Earnings Imputed	(2) Formal Sector Earnings Only
Panel A: All Directly Estimated Earnings Effects		
WTP	$4747	$3322
Net Costs	$5357	$5425
MVPF	0.89	0.61
Panel B: Assume Full-Sample Persistence through Q12		
WTP	$5565	$4125
Net Costs	$5310	$5378
MVPF	1.05	0.77

Note: Per-participant willingness to pay in Panel A is calculated using the in-program and post-program earnings effects in Table A-6 to allow for treatment timing heterogeneity and to directly measure earnings effects for more quarters for the earlier enrollees. Net costs per participant in Panel A use these same directly estimated earnings effects to deduct each participant's additional taxes paid from their estimated program costs. Panel B uses the same directly estimated effects for quarters 0-12 in the early cohorts, and carries forward the average post-program effects through quarter 12 for the late cohorts, whose earnings we do not observe through quarter 12. All specifications assume a 3% discount rate.

Table 8 presents the estimated MVPF under alternative lengths of program effect persistence using a 3% discount rate. Panel A uses only earnings effects estimated using quarters we observe directly—at least seven quarters post-enrollment for the latest cohorts and up to 12 quarters post-enrollment for the earliest cohorts. The MVPF based on this set of earnings gains ranges from 0.61 to 0.89, depending on whether imputed in-placement earnings are included. Because all of the costs occur and are fully measured upfront, and because there is no evidence that earnings gains fade out over time (see Appendix Table A-6), these estimates are very likely a lower bound of STEP's cost-effectiveness compared to a calculation that incorporates earnings gains over a longer time horizon.

As an example of how persistence of the program impacts can affect the estimated cost-effectiveness, Panel B imposes the assumption that earnings effects among the later-treated cohorts last through quarter 12, which aligns with the available follow-up period for earlier cohorts. Under this plausible assumption, the MVPF estimates range from 0.77 to 1.05. [31]

[31] We assume constant treatment effects based on the most recent set of earnings estimates for each cohort (quarters 4–7 for later enrollees, and quarters 8–12 for earlier enrollees). Thus, this panel does not assume that effects for the later cohorts are as large or as persistent as in the earlier cohorts. Instead, we simply assume a similar degree of persistence from quarters 4–7 to quarters 8–12 for both cohorts.

Notably, the estimates that include imputed earnings during placements lead to a willingness to pay ($5,565) that slightly exceeds the net costs ($5,310). Because the estimated effects in available data suggest that the earnings effects persist through at least quarter 12, we believe that it is appropriate to conclude that STEP is likely as efficient or more efficient than a non-distortionary cash transfer.

This set of calculations ignores any impacts on TANF participation, in part because we cannot determine whether the program simply shifts TANF participation forward in time or if it instead increases lifetime participation.[32] Because the estimated MVPF of cash assistance is near 1 (Hendren and Sprung-Keyser, 2020), this omission likely has a negligible effect on our estimate of STEP's MVPF.

Importantly, however, prior estimates of the MVPF of cash welfare programs suggest that the inter-generational effects of providing additional income to benefit-eligible households are substantial. More generally, Hendren and Sprung-Keyser (2020) find that the most cost-effective interventions for adults have positive spillovers on young children. Therefore, our inability to estimate future inter-generational effects of additional family income during childhood likely results in conservative estimates of the MVPF of the STEP program.

7 Conclusion

This paper quantifies the effects of a supported work program available to TANF recipients on labor market outcomes and TANF receipt. Developed in the presence of work requirements and time limits, this Colorado program was intended to improve low-income families' long-term self-sufficiency while maintaining their short-term access to the safety net. We find evidence that STEP met both of these goals. Using a stacked difference-in-differences design, we find relatively large increases in formal sector quarterly employment (9.6 percentage points) and quarterly earnings ($563) during the fourth through seventh quarter

[32] Given lifetime limits of 60 months of benefits, there is limited scope for increases in total benefit receipt.

post-enrollment. Importantly, we also find improvements in job stability, with STEP participants being 20 percent more likely to remain with the same employer throughout that post-program year. Finally, STEP increases TANF receipt during the program—when participants have just experienced a negative earnings shock and likely have an acute need for cash assistance—but not after they have exited the program.

With an MVPF near or exceeding 1, STEP is more cost-effective than other similar active labor market programs.[33] Specifically, it has larger and more durable effects on employment and earnings compared with other interventions that share key features of its program model—case management and access to time-limited subsidized employment.

Although not unique to STEP, a key feature is the potential for a subsidized position to become permanent if both the participant and the host-site employer agree. As demonstrated above, STEP increased the stability of employer-employee matches, which is consistent with the possibility that this program design element matters. This difference could explain why programs that placed participants in a limited set of positions that were not intended to become permanent—such as those reviewed in Bloom (2010) and the "paid work experience" model in LA County's Transitional Subsidized Employment Program (Anderson et al., 2019)—did not have enduring effects.

However, STEP has larger and more durable program impacts even when compared to other program models in which subsidized jobs could become permanent, including ReHire Colorado (Barham, Cadena and Turner, 2023) and the "on-the-job training" arm in LA County's Transitional Subsidized Employment Program (Anderson et al., 2019). ReHire Colorado showed smaller post-program earnings impacts and greater fadeout despite a very similar program model. The key difference between STEP and ReHire is the target population, with ReHire available to a much broader set of Colorado residents with household income less than 150 percent of the federal poverty level. The difference therefore suggests

[33]As an example, ReHire Colorado has an estimated MVPF of 0.32 based on 16 quarters of directly estimated earnings effects. Because MVPF is a relatively new framework, many prior studies do not explicitly calculate it. Program models with a substantial upfront investment that does not lead to lasting earnings gains, however, are very unlikely to reach an MVPF near 1.

that either 1) this program model may be especially effective for low-income single mothers with small children and limited recent work history, or 2) the work requirements of TANF may complement the effectiveness of the subsidized employment programs.

The LA "on-the-job training" intervention also served TANF recipients and included temporary placements that could become permanent, but there are some key programmatic differences. In LA, the maximum length of the subsidy was only two months rather than six months, and host-site employers were expected to begin covering the full cost of the participant's employment after that time (Anderson et al., 2019). Further, only 42 percent of participants in this arm were successfully placed in a transitional job, compared to STEP's 64 percent (see Table 1). Together these differences suggest that STEP may have led a different set of employers to be willing to host a participant, which may have increased the rate at which placements became permanent.

Finally, STEP also contrasts with early welfare-to-work programs that sought to simultaneously increase earnings and decrease benefit receipt, often resulting in some budgetary savings but negligible effects on participants' overall income (Greenberg, Deitch and Hamilton, 2010). Such programs often met their stated goals of reducing dependence on welfare but did not necessarily reduce poverty for those they served.

Overall, therefore, this paper suggests that providing job placement into subsidized positions that can become permanent is a particularly effective intervention for TANF recipients. Participants immediately increased their household income after a negative earnings shock by nearly $1000 per quarter while participating in the program and over $500 per quarter (24 percent) in the year after program exit. Programs like STEP thus offer a promising pathway for state policymakers looking to improve the short-term well-being and longer-term self-sufficiency of low-income families.

References

Aizer, Anna, Hilary Hoynes, and Adriana Lleras-Muney. 2022. "Children and the US social safety net: Balancing disincentives for adults and benefits for children." *Journal of Economic Perspectives*, 36(2): 149–174.

Aizer, Anna, Shari Eli, Joseph Ferrie, and Adriana Lleras-Muney. 2016. "The long-run impact of cash transfers to poor families." *American Economic Review*, 106(4): 935–971.

Aizer, Anna, Sungwoo Cho, Shari Eli, and Adriana Lleras-Muney. 2024. "The Impact of Cash Transfers to Poor Mothers on Family Structure and Maternal Well-Being." *American Economic Journal: Applied Economics*, 16(2): 492–529.

Altonji, Joseph G, and Charles R Pierret. 2001. "Employer Learning and Statistical Discrimination." *The Quarterly Journal of Economics*, 116(1): 313–350.

Anderson, Chloe, Mary Farrell, Asaph Glosser, and Bret Barden. 2019. "Testing Two Subsidized Employment Models for TANF Recipients: Final Impacts and Costs of the Los Angeles County Transitional Subsidized Employment Program." OPRE Report 2019-71.

Andersson, Fredrik, Harry J Holzer, Julia I Lane, David Rosenblum, and Jeffrey Smith. 2022. "Does federally-funded job training work? Nonexperimental estimates of WIA training impacts using longitudinal data on workers and firms." *Journal of Human Resources*.

Arulampalam, Wiji. 2001. "Is Unemployment Really Scarring? Effects of Unemployment Experiences on Wages." *Economic Journal*, 111(475): F585–F606.

Autor, David H, and Susan N Houseman. 2010. "Do temporary-help jobs improve labor market outcomes for low-skilled workers? Evidence from "Work First"." *American Economic Journal: Applied Economics*, 2(3): 96–128.

Barden, Bret, Randall Juras, Cindy Redcross, Mary Farrell, and Dan Bloom. 2018. "New Perspectives on Creating Jobs: Final Impacts of the Next Generation of Subsidized Employment Programs." New York: MDRC.

Barham, Tania, Brian C Cadena, and Patrick S Turner. 2023. "Taking a Chance on Workers: Evidence on the Effects and Mechanisms of Subsidized Employment from an RCT." *IZA Discussion Paper No. 16221*.

Bitler, Marianne, and Hilary Hoynes. 2016. "The more things change, the more they stay the same? The safety net and poverty in the Great Recession." *Journal of Labor Economics*, 34(S1): S403–S444.

Bitler, Marianne, and Hilary W Hoynes. 2010. "The state of the safety net in the post-welfare reform era." *Brookings Papers on Economic Activity*, Fall: 71–142.

Bloom, Dan. 2010. "Transitional Jobs: Background, Program Models, and Evaluation Evidence." New York: MDRC.

Cook, Jason B, and Chloe N East. 2024. "Work Requirements with No Teeth Still Bite: Disenrollment and Labor Supply Effects of SNAP General Work Requirements." *NBER Working Paper No. 32441.*

DHHS. 2024. "TANF 13th Report to Congress." U.S. Department of Health and Human Services, Administration for Children and Families, Office of Family Assistance.

Freedman, Stephen. 2000. "Four-Year Impacts of Ten Programs on Employment Stability and Earnings Growth. The National Evaluation of Welfare-to-Work Strategies." New York: MDRC, ERIC.

Gangl, Markus. 2006. "Scar Effects of Enemployment: An Assessment of Institutional Complementarities." *American Sociological Review*, 71(6): 986–1013.

Goodman-Bacon, Andrew. 2021. "Difference-in-differences with variation in treatment timing." *Journal of econometrics*, 225(2): 254–277.

Gray, Colin, Adam Leive, Elena Prager, Kelsey Pukelis, and Mary Zaki. 2023. "Employed in a SNAP? The impact of work requirements on program participation and labor supply." *American Economic Journal: Economic Policy*, 15(1): 306–341.

Greenberg, David H, Victoria Deitch, and Gayle Hamilton. 2010. "A synthesis of random assignment benefit-cost studies of welfare-to-work programs." *Journal of Benefit-Cost Analysis*, 1(1): 1–30.

Grogger, Jeffrey. 2003. "The effects of time limits, the EITC, and other policy changes on welfare use, work, and income among female-headed families." *Review of Economics and statistics*, 85(2): 394–408.

Grosz, Michel. 2020. "The returns to a large community college program: Evidence from admissions lotteries." *American Economic Journal: Economic Policy*, 12(1): 226–253.

Guldi, Melanie, and Lucie Schmidt. 2017. "Taxes, Transfers, and Women's Labor Supply in the United States." In *The Oxford Handbook of Women and the Economy*. 453.

Han, Jeehoon, Bruce D Meyer, and James X Sullivan. 2021. "The consumption, income, and well-being of single mother–headed families 25 years after welfare reform." *National Tax Journal*, 74(3): 791–824.

Hendren, Nathaniel, and Ben Sprung-Keyser. 2020. "A unified welfare analysis of government policies." *The Quarterly Journal of Economics*, 135(3): 1209–1318.

Hoynes, Hilary W, and Diane Whitmore Schanzenbach. 2018. "Safety Net Investments in Children." *Brookings Papers on Economic Activity*, Spring: 89–138.

Jacobson, Louis S., Robert J. LaLonde, and Daniel G. Sullivan. 1993. "Earnings Losses of Displaced Workers." *The American Economic Review*, 83(4): 685–709.

Krolikowski, Pawel. 2018. "Choosing a Control Group for Displaced Workers." *ILR Review*, 71(5): 1232–1254.

Macartney, Suzanne, and Robin Ghertner. 2023. "How Many People Participate in the Social Safety Net?" Office of the Assistant Secretary for Planning and Evaluation, United States Department of Health and Human Services.

Meyer, Bruce D, and James X Sullivan. 2008. "Changes in the consumption, income, and well-being of single mother headed families." *American Economic Review*, 98(5): 2221–2241.

Pries, Michael, and Richard Rogerson. 2005. "Hiring Policies, Labor Market Institutions, and Labor Market Flows." *Journal of Political Economy*, 14(1): 260–300.

Richard, Katherine, and Lea Bart. 2024. "Penalties in the Safety Net: Effects of Work Requirement Enforcement on Program Participation and Labor Supply."

Rose, Evan K., and Yotam Shem-Tov. 2023. "How Replaceable Is a Low-Wage Job?" *NBER Working Paper 31447*.

Ruhm, Christopher J. 1991. "Are Workers Permanently Scarred by Job Displacements?" *The American Economic Review*, 81(1): 319–324.

Stevens, Ann Huff. 1997. "Persistent Effects of Job Displacement: The Importance of Multiple Job Losses." *Journal of Labor Economics*, 15(1): 165–188.

Wing, Coady, Seth M Freedman, and Alex Hollingsworth. 2024. "Stacked Difference-in-Differences." *NBER Working Paper No. 32054*.

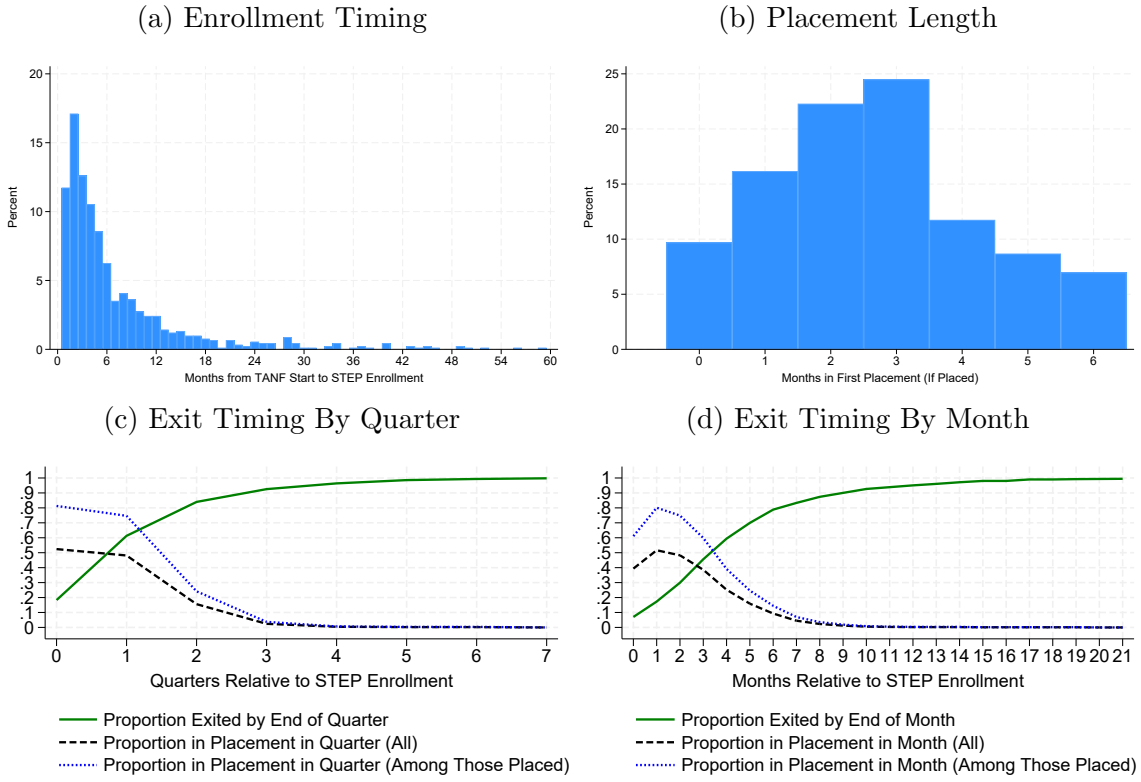
A Appendix – For Online Publication

A.1 Key Program Events

This section provides additional descriptive analysis of the timing of key program events, including enrollment, placement, and program exit for the treatment group. Panel (a) of Figure A-1 shows the distribution of enrollment times for STEP participants, measured as months since the start of the TANF spell. Panel (b) shows the distribution of placement lengths. Panel (c) shows the proportion of STEP participants who are in a placement, both unconditionally (black dashed line) and among those who were placed (blue dotted line) in a given quarter relative to STEP enrollment. The solid green line shows the proportion of those with a placement who have exited by the listed quarter. Panel (d) provides similar analysis but at the monthly level. These lower panels motivate our treatment of outcomes observed during quarters 4–7 and months 10–21 as post-program outcomes.

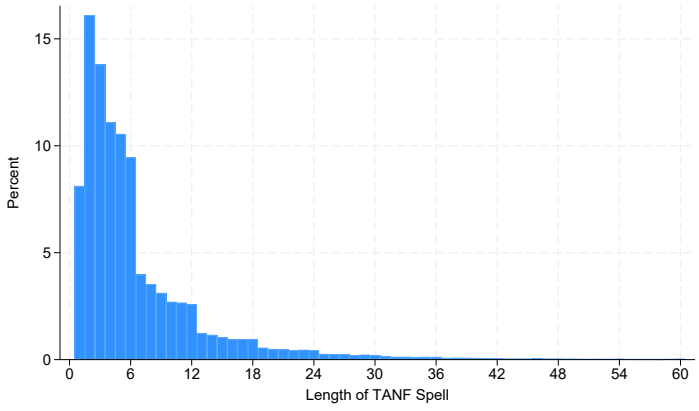
Figure A-2 provides a distribution of the duration of TANF spells, measured in months. The sharp drops at multiples of 6 reflect the recertification requirements, which may provide quasi-random variation in STEP program interest (see section 4.3).

Figure A-1: STEP Enrollment, Placement, and Exit Timing



Note: Shows proportion of treatment group that has exited the program in each quarter or month relative to enrollment. Quarters 0-3 (months 0-9) largely represent outcomes during individuals' enrollment in the program and quarters 4-7 (months 10-21) largely represent outcomes after they have exited the program.

Figure A-2: TANF Spell Duration



Note: Data source is administrative TANF benefits data from the CBMS data system. Figure shows a histogram of TANF spell duration in months.

A.2 Sample Description

As discussed in Section 3.2, we remove the three percent (35/1,201) of treatment group observations who do not match to the CBMS and earnings data using social security number, name, and/or birthdate. We also drop treatment group members who did not receive a TANF payment in their enrollment month as the program requires (24/1,201) or who enrolled in STEP after the first quarter of 2020 (19/1,201) because their program experience was affected by the COVID-19 pandemic and the impending end of the program's initial funding period.

Finally, we remove treatment group (211/1,201) and comparison group individuals from our analysis if their earnings records are inconsistent with qualifying for TANF participation. CDHS personnel identified a common issue with the CDLE records where earnings amounts were entered incorrectly by the employer. To mitigate the impact of this data error, we drop individuals who had person-level median quarterly earnings during quarters in which they received TANF benefits in the two years prior to their focal TANF spell that exceeded $10,000 per quarter, or the equivalent of $40,000 per year, were the only adult in their household according to TANF case data, and had fewer than 4 children. We assume that such earnings are employer data entry errors, since they far exceed the maximum income limit for a family with 4 children to qualify for TANF benefits in Colorado, which was $806 per month (the equivalent of $2,418 per quarter) during the program's operation. Additionally, any STEP participants or other TANF recipients who were ever employed by an employer with greater than than 50% of their entries for TANF recipients below $100 or exceeding $10,000 per quarter are removed from the sample as these employers were consistently reporting amounts inconsistent with TANF participation in a sample consisting solely of TANF beneficiaries. We also remove any individual who reported working more than 5 jobs in a given quarter since these are likely instances of sharing a social security number with other individuals who are ineligible to work in the U.S. Finally, the CDLE records included payments related to settlements, unemployment insurance, and housing allowances, and we remove these entries as they were not earned income by an individual. We also correct quarterly earnings entries

that are probable within-person outliers. An outlier is flagged when a participant's earnings are in the top one percent of both percentage and absolute deviations from the individual's median earnings when employed. For these observations, we impute participants' earnings using lags, leads, and within-person medians (split before and after STEP enrollment to allow for the possibility of large program effects).

Table A-1 provides sample sizes and stacking weights for each of the monthly program entry cohorts.

Table A-2 provides descriptive statistics on the variables used to estimate the propensity score. There are many notable differences in means between the treatment and comparison groups.

Table A-1: Sample Sizes and Stacking Weights by Enrollment Month

(1) m	(2) $N^m_{D=1}$	(3) $N^m_{D=0}$	(4) $w_{im,D=0}$	(5) $w_{im,D=0} * N^m_{D=0}$
1/2018	4	7451	0.11	835.9079
2/2018	34	7430	0.96	7105.22
3/2018	30	7402	0.85	6269.31
4/2018	39	7177	1.14	8150.10
5/2018	38	7316	1.09	7941.12
6/2018	32	7205	0.93	6687.26
7/2018	22	7293	0.63	4597.49
8/2018	36	7229	1.04	7523.17
9/2018	28	7185	0.81	5851.36
10/2018	38	7388	1.07	7941.13
11/2018	44	7322	1.26	9194.99
12/2018	37	7422	1.04	7732.15
1/2019	43	7208	1.25	8986.01
2/2019	55	7011	1.64	11493.73
3/2019	46	6907	1.39	9612.94
4/2019	47	6826	1.44	9821.92
5/2019	33	7102	0.97	6896.24
6/2019	32	6980	0.96	6687.26
7/2019	39	7125	1.14	8150.10
8/2019	40	7132	1.17	8359.08
9/2019	32	6813	0.98	6687.26
10/2019	32	6847	0.98	6687.26
11/2019	29	6613	0.92	6060.33
12/2019	28	6592	0.89	5851.36
1/2020	31	6497	1.00	6478.29
2/2020	21	6553	0.67	4388.52
3/2020	22	6561	0.70	4597.49
	Treatment	Comparison		
N (Unique)	912	24991		
N (Total)	912	190587	$\sum w_{im,D=0}$	190587

Note: Data come from administrative UI earnings data from CDLE, administrative TANF data from CDHS, and CW STEP program data from CDLE. The sample includes STEP participants who applied between January 2018 and March 2020 and other individuals who received TANF during the same period.

Table A-2: Descriptive Statistics - No Propensity Weights

	(1) Mean (Trt.)	(2) SD (Trt.)	(3) Mean (Comp.)	(4) SD (Comp.)	(5) Diff. Means	(6) SE (Diff.)	(7) t (Diff.)	(8) Diff. / SD (All)	(9) Diff. / SD (Comp.)
Age	33.21	8.10	31.96	8.07	1.25	0.28	4.52	0.15	0.15
Female	0.87	0.33	0.83	0.38	0.05	0.01	4.01	0.14	0.12
Only Adult in Household	0.85	0.35	0.80	0.40	0.05	0.01	4.42	0.15	0.13
Number of Children in Household	1.94	1.13	1.94	1.24	0.00	0.04	0.08	0.00	0.00
Infant in Household	0.15	0.36	0.19	0.39	-0.04	0.01	-3.13	-0.11	-0.10
Child Under 5 in Household	0.58	0.49	0.60	0.49	-0.02	0.02	-0.99	-0.03	-0.03
Single Parent of Infant	0.12	0.32	0.14	0.34	-0.02	0.01	-1.85	-0.06	-0.06
Single Parent of Child Under 5	0.49	0.50	0.45	0.50	0.04	0.02	2.15	0.07	0.07
Less than High School Equivalent	0.13	0.33	0.22	0.41	-0.09	0.01	-7.98	-0.28	-0.22
GED	0.13	0.34	0.12	0.33	0.01	0.01	0.59	0.02	0.02
High School Diploma	0.41	0.49	0.40	0.49	0.01	0.02	0.55	0.02	0.02
More than High School Equivalent	0.28	0.45	0.20	0.40	0.07	0.02	4.74	0.16	0.18
Education Data Missing	0.05	0.21	0.05	0.22	-0.01	0.01	-0.77	-0.03	-0.02
TANF Employment Barrier Identified	0.61	0.49	0.65	0.48	-0.04	0.02	-2.62	-0.09	-0.09
TANF Work Requirements Exemption	0.06	0.24	0.10	0.30	-0.04	0.01	-4.52	-0.16	-0.13
TANF Sanction	0.10	0.30	0.18	0.38	-0.08	0.01	-7.78	-0.27	-0.21
TANF Case Plan Re-engagement	0.37	0.48	0.46	0.50	-0.08	0.02	-5.10	-0.17	-0.17
Total Earnings, 3 Years Prior (Thousands	20.53	24.36	20.58	30.49	-0.05	0.84	-0.06	0.00	0.00
Change in Earnings, Q-4 to Q-3 (Thousand	0.07	2.13	0.10	2.23	-0.03	0.07	-0.38	-0.01	-0.01
Change in Earnings, Q-3 to Q-2 (Thousand	0.25	2.35	0.16	2.26	0.09	0.08	1.16	0.04	0.04
Change in Earnings, Q-2 to Q-1 (Thousand	0.40	1.85	0.30	2.28	0.11	0.06	1.72	0.06	0.05
Unemployed Full Year Prior to T=0	0.29	0.45	0.32	0.47	-0.03	0.02	-1.90	-0.06	-0.06
Employed Q-1	0.44	0.50	0.42	0.49	0.02	0.02	1.11	0.04	0.04
Employed Q-2	0.48	0.50	0.44	0.50	0.04	0.02	2.51	0.08	0.09
Employed Q-3	0.51	0.50	0.45	0.50	0.06	0.02	3.48	0.12	0.12
Employed Q-4	0.51	0.50	0.46	0.50	0.05	0.02	3.16	0.11	0.11
Employed Q-5	0.51	0.50	0.47	0.50	0.04	0.02	2.36	0.08	0.08
Employed Q-6	0.51	0.50	0.48	0.50	0.03	0.02	1.87	0.06	0.06
Employed Q-7	0.48	0.50	0.47	0.50	0.01	0.02	0.53	0.02	0.02
Employed Q-8	0.48	0.50	0.47	0.50	0.01	0.02	0.42	0.01	0.01
Employed Q-9	0.48	0.50	0.47	0.50	0.01	0.02	0.32	0.01	0.01
Employed Q-10	0.48	0.50	0.47	0.50	0.01	0.02	0.47	0.02	0.02
Employed Q-11	0.47	0.50	0.47	0.50	0.00	0.02	-0.22	-0.01	-0.01
Employed Q-12	0.47	0.50	0.46	0.50	0.01	0.02	0.30	0.01	0.01
Total TANF Benefits, 3 Years Prior	4599.33	4540.91	5287.70	4955.49	-688.36	155.96	-4.41	-0.15	-0.14
TANF Receipt M-1	0.88	0.32	0.86	0.35	0.03	0.01	2.46	0.08	0.07
TANF Receipt M-2	0.71	0.45	0.73	0.44	-0.02	0.02	-1.33	-0.04	-0.05
TANF Receipt M-3	0.60	0.49	0.64	0.48	-0.04	0.02	-2.63	-0.09	-0.09
TANF Receipt M-4	0.51	0.50	0.57	0.49	-0.06	0.02	-3.72	-0.12	-0.13
TANF Receipt M-5	0.45	0.50	0.52	0.50	-0.07	0.02	-4.03	-0.13	-0.13
TANF Receipt M-6	0.42	0.49	0.48	0.50	-0.06	0.02	-3.68	-0.12	-0.12
TANF Receipt M-7	0.40	0.49	0.46	0.50	-0.05	0.02	-3.30	-0.11	-0.11
TANF Receipt M-8	0.37	0.48	0.43	0.50	-0.06	0.02	-3.50	-0.12	-0.12
TANF Receipt M-9	0.35	0.48	0.41	0.49	-0.05	0.02	-3.23	-0.11	-0.11
TANF Receipt M-10	0.34	0.47	0.38	0.49	-0.05	0.02	-2.91	-0.10	-0.10
TANF Receipt M-11	0.33	0.47	0.36	0.48	-0.04	0.02	-2.37	-0.08	-0.08
TANF Receipt M-12	0.31	0.46	0.35	0.48	-0.04	0.02	-2.45	-0.08	-0.08
Total TANF Months, M-36 to M-13	4.76	6.86	5.58	7.21	-0.82	0.23	-3.49	-0.12	-0.11
N (Treatment)	912								
N (Comparison)	190587								
N (All)	191499								

Note: Data come from administrative UI earnings data from CDLE, administrative TANF data from CDHS, and fields collected by the TANF caseworker at TANF application. The sample includes TANF recipients who applied between January 2018 and March 2020.

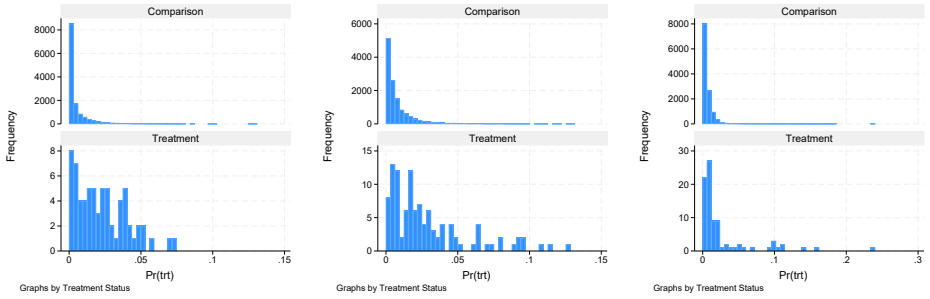
A.3 Propensity Score Weighting

This section provides additional details on the propensity score estimation and provides results, including coefficient estimates and common support graphs.

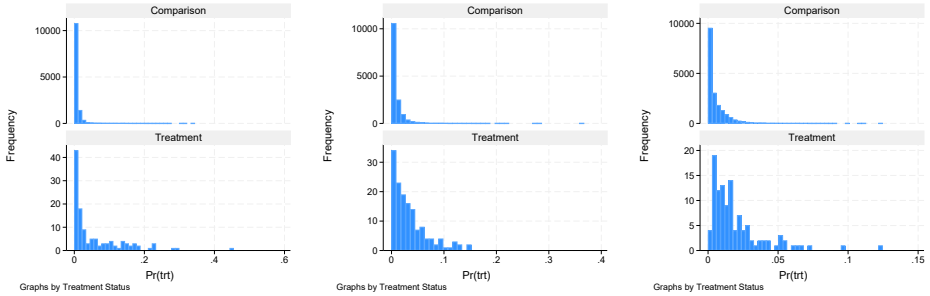
Figure A-3 demonstrates common support by showing the distribution of estimated treatment probabilities for each of nine program entry cohorts. Importantly, across all cohorts, there are a substantial number of comparison group members with estimated propensity scores similar to the treatment group's.

Figure A-3: Common Support - Enrollment Propensity by Quarter

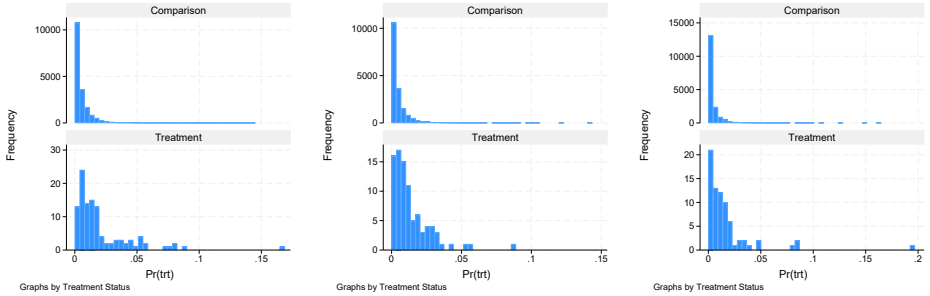
(a) Q1 2018 (b) Q2 2018 (c) Q3 2018



(d) Q4 2018 (e) Q1 2019 (f) Q2 2019



(g) Q3 2019 (h) Q4 2019 (i) Q1 2020



Note: Vertical axes show total number of individuals in each bin in the treatment and comparison groups. Horizontal axis shows predicted probability of STEP enrollment for each quarter's enrollees and the corresponding comparison observations. Individuals in the comparison group can have up to three distinct observations in a given quarter if they remain on TANF for the entire quarter, with covariates observed relative to each of the treatment group's different enrollment months in that quarter. Propensity score estimation is pooled within each enrollment quarter to avoid perfectly predicting treatment status due to small treated sample sizes in some individual enrollment months; see Table A-1 for details.

Table A-3 provides the same descriptive statistics as Table A-2, but after reweighting. The differences in means between the treatment and comparison group are much smaller and none is statistically significant, suggesting that the reweighting procedure worked as intended.

Table A-3: Descriptive Statistics - Propensity-Weighted

	(1) Mean (Trt.)	(2) SD (Trt.)	(3) Mean (Comp.)	(4) SD (Comp.)	(5) Diff. Means	(6) SE (Diff.)	(7) t (Diff.)	(8) Diff. / SD (All)	(9) Diff. / SD (Comp.)
Age	33.21	8.10	33.25	8.08	-0.03	0.30	-0.11	0.00	0.00
Female	0.87	0.33	0.87	0.38	0.00	0.01	0.27	0.01	0.01
Only Adult in Household	0.85	0.35	0.85	0.41	0.00	0.01	0.10	0.00	0.00
Number of Children in Household	1.94	1.13	1.94	1.21	0.00	0.04	-0.07	0.00	0.00
Infant in Household	0.15	0.36	0.16	0.39	0.00	0.01	-0.33	-0.01	-0.01
Child Under 5 in Household	0.58	0.49	0.58	0.49	0.00	0.02	-0.21	-0.01	-0.01
Single Parent of Infant	0.12	0.32	0.12	0.34	0.00	0.01	-0.35	-0.01	-0.01
Single Parent of Child Under 5	0.49	0.50	0.49	0.50	0.00	0.02	-0.10	0.00	0.00
Less than High School Equivalent	0.13	0.33	0.12	0.41	0.00	0.01	0.11	0.00	0.00
GED	0.13	0.34	0.13	0.33	0.00	0.01	0.23	0.01	0.01
High School Diploma	0.41	0.49	0.40	0.49	0.00	0.02	0.10	0.00	0.00
More than High School Equivalent	0.28	0.45	0.29	0.41	-0.01	0.02	-0.64	-0.02	-0.03
Education Data Missing	0.05	0.21	0.05	0.22	0.00	0.01	-0.13	0.00	0.00
TANF Employment Barrier Identified	0.61	0.49	0.61	0.48	0.00	0.02	0.15	0.01	0.01
TANF Work Requirements Exemption	0.06	0.24	0.06	0.27	0.00	0.01	0.50	0.02	0.02
TANF Sanction	0.10	0.30	0.10	0.37	0.00	0.01	0.24	0.01	0.01
TANF Case Plan Re-engagement	0.37	0.48	0.37	0.50	0.00	0.02	0.08	0.00	0.00
Total Earnings, 3 Years Prior (Thousands	20.53	24.36	20.45	30.76	0.08	0.87	0.09	0.00	0.00
Change in Earnings, Q-4 to Q-3 (Thousand	0.07	2.13	0.09	2.26	-0.01	0.07	-0.18	-0.01	-0.01
Change in Earnings, Q-3 to Q-2 (Thousand	0.25	2.35	0.23	2.32	0.02	0.08	0.31	0.01	0.01
Change in Earnings, Q-2 to Q-1 (Thousand	0.40	1.85	0.41	2.34	-0.01	0.07	-0.10	0.00	0.00
Unemployed Full Year Prior to T=0	0.29	0.45	0.29	0.47	0.00	0.02	0.10	0.00	0.00
Employed Q-1	0.44	0.50	0.44	0.49	0.00	0.02	-0.09	0.00	0.00
Employed Q-2	0.48	0.50	0.48	0.50	0.00	0.02	-0.06	0.00	0.00
Employed Q-3	0.51	0.50	0.51	0.50	0.00	0.02	0.18	0.01	0.01
Employed Q-4	0.51	0.50	0.51	0.50	0.00	0.02	0.11	0.00	0.00
Employed Q-5	0.51	0.50	0.51	0.50	0.00	0.02	0.14	0.01	0.01
Employed Q-6	0.51	0.50	0.50	0.50	0.00	0.02	0.16	0.01	0.01
Employed Q-7	0.48	0.50	0.48	0.50	0.00	0.02	0.05	0.00	0.00
Employed Q-8	0.48	0.50	0.48	0.50	0.00	0.02	0.21	0.01	0.01
Employed Q-9	0.48	0.50	0.48	0.50	0.00	0.02	0.24	0.01	0.01
Employed Q-10	0.48	0.50	0.48	0.50	0.00	0.02	0.16	0.01	0.01
Employed Q-11	0.47	0.50	0.46	0.50	0.00	0.02	0.27	0.01	0.01
Employed Q-12	0.47	0.50	0.46	0.50	0.01	0.02	0.32	0.01	0.01
Total TANF Benefits, 3 Years Prior	4599.33	4540.91	4552.65	4885.08	46.68	161.93	0.29	0.01	0.01
TANF Receipt M-1	0.88	0.32	0.88	0.35	0.00	0.01	0.01	0.00	0.00
TANF Receipt M-2	0.71	0.45	0.71	0.44	0.00	0.02	0.08	0.00	0.00
TANF Receipt M-3	0.60	0.49	0.60	0.48	0.00	0.02	0.26	0.01	0.01
TANF Receipt M-4	0.51	0.50	0.51	0.50	0.00	0.02	0.20	0.01	0.01
TANF Receipt M-5	0.45	0.50	0.45	0.50	0.00	0.02	0.19	0.01	0.01
TANF Receipt M-6	0.42	0.49	0.42	0.50	0.00	0.02	0.28	0.01	0.01
TANF Receipt M-7	0.40	0.49	0.40	0.50	0.01	0.02	0.38	0.01	0.01
TANF Receipt M-8	0.37	0.48	0.37	0.49	0.01	0.02	0.32	0.01	0.01
TANF Receipt M-9	0.35	0.48	0.35	0.49	0.01	0.02	0.32	0.01	0.01
TANF Receipt M-10	0.34	0.47	0.33	0.48	0.01	0.02	0.35	0.01	0.01
TANF Receipt M-11	0.33	0.47	0.32	0.48	0.01	0.02	0.31	0.01	0.01
TANF Receipt M-12	0.31	0.46	0.30	0.47	0.01	0.02	0.31	0.01	0.01
Total TANF Months, M-36 to M-13	4.76	6.86	4.71	7.04	0.05	0.25	0.21	0.01	0.01
N (Treatment)	912								
N (Comparison)	138554								
N (All)	139466								

Note: Data come from administrative UI earnings data from CDLE, administrative SNAP and TANF data from CDHS, and fields collected by the TANF caseworker at TANF application. The sample includes individuals who received TANF payments in Colorado between January 2018 and March 2020.

Table A-4 provides coefficient estimates from the probit regressions used to construct the propensity scores. Each column represents a separate regression for the listed cohort.

Table A-4: Probit Results

	2018Q1	2018Q2	2018Q3	2018Q4	2019Q1	2019Q2	2019Q3	2019Q4	2020Q4
Employed Q-1	0.001	0.001	-0.004*	0.004	0.001	-0.001	-0.001	-0.001	0.001
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)
Employed Q-2	-0.002	-0.000	0.002	-0.003	0.001	0.003	0.002	-0.001	0.002
	(0.001)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Employed Q-3	0.002	0.001	0.001	0.000	0.001	0.000	-0.001	0.003*	-0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)
Employed Q-4	0.002	0.001	0.001	0.002	0.001	0.000	0.001	0.000	0.001
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
Employed Q-5	0.002	0.000	0.002	-0.001	-0.003	0.000	0.002	0.001	-0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
Employed Q-6	0.002	0.006**	-0.003	0.006**	0.001	-0.000	-0.002	-0.002	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)
Employed Q-7	-0.002	-0.004	-0.000	-0.005*	0.003	0.001	0.003	0.002	-0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)
Employed Q-8	0.002	-0.001	-0.000	0.001	0.001	0.000	0.000	-0.002	-0.000
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Employed Q-9	-0.001	0.002	0.002	0.002	-0.001	-0.000	-0.002	0.002	0.000
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)
Employed Q-10	0.001	-0.003	-0.001	-0.000	-0.001	0.002	0.001	0.002	0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Employed Q-11	0.000	0.005*	0.002	-0.001	0.000	-0.005*	0.001	-0.000	-0.004**
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)
Employed Q-12	-0.001	0.001	-0.002	-0.001	0.003	0.002	-0.002	0.000	0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Total Earnings, 3 Years Prior	-0.000	-0.000*	-0.000	-0.000*	-0.000*	-0.000	-0.000	-0.000**	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Change in Earnings, Q-4 to Q-3	-0.000	0.000	0.000	0.001	0.001*	0.000	0.000	-0.000	-0.000

Continued on next page

Table A-4 – *Continued from previous page*

	2018Q1	2018Q2	2018Q3	2018Q4	2019Q1	2019Q2	2019Q3	2019Q4	2020Q4
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Change in Earnings, Q-3 to Q-2	-0.000	0.001	0.000	0.001	0.001*	-0.000	0.000	-0.001*	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Change in Earnings, Q-2 to Q-1	-0.000	0.001	-0.000	0.001	0.001	-0.000	0.000	-0.000	0.000
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age	0.000	0.001	0.000	0.001	-0.000	0.001	0.000	0.001	0.000
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Female	0.005*	0.004	0.001	0.005*	0.002	0.003	0.002	0.001	-0.001
	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)
Number of Children in Household	0.000	-0.005	0.002	-0.000	-0.001	-0.001	-0.001	0.002*	-0.001
	(0.001)	(0.003)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Adams County	0.000	0.000	0.000	0.000	-0.045***	0.000	0.000	0.000	0.000
	(.)	(.)	(.)	(.)	(0.008)	(.)	(.)	(.)	(.)
Arapahoe County	-0.009***	-0.023***	-0.022***	-0.032***	-0.022***	-0.007**	-0.014***	-0.008***	-0.011***
	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.002)	(0.003)	(0.002)	(0.003)
Denver County	0.000	0.000	0.000	0.000	0.000	-0.029***	-0.022***	-0.013***	-0.009***
	(.)	(.)	(.)	(.)	(.)	(0.005)	(0.003)	(0.002)	(0.002)
El Paso County	-0.015***	-0.023***	-0.021***	-0.031***	-0.028***	-0.016***	-0.018***	-0.014***	-0.011***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.002)
Jefferson County	-0.009**	-0.018***	-0.020***	-0.028***	-0.019***	-0.010***	-0.012***	-0.005*	-0.004
	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
Larimer County	-0.003	-0.016***	-0.018***	-0.033***	-0.017***	-0.010***	-0.011***	-0.009***	-0.005*
	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)	(0.003)	(0.002)	(0.003)	(0.002)
Mesa County	-0.005	-0.011**	-0.018***	-0.025***	-0.019***	-0.011***	-0.011***	-0.007**	-0.007**
	(0.003)	(0.004)	(0.004)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.002)
Weld County	-0.006	-0.011**	-0.017***	-0.025***	-0.024***	-0.009**	-0.013***	-0.008**	-0.006**
	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.002)
High School Diploma	-0.001	-0.003	-0.003	-0.001	-0.005**	0.000	-0.003*	0.000	-0.004**
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)

Continued on next page

Table A-4 – *Continued from previous page*

	2018Q1	2018Q2	2018Q3	2018Q4	2019Q1	2019Q2	2019Q3	2019Q4	2020Q4
GED	-0.011**	0.003	-0.003	0.002	-0.005	-0.000	-0.002	0.000	-0.005**
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
t0_edmissing	-0.003	-0.004	-0.002	-0.004	-0.006	-0.000	0.000	-0.005	-0.005
	(0.002)	(0.004)	(0.003)	(0.004)	(0.004)	(0.003)	(.)	(0.003)	(0.002)
Only Adult in Household	-0.004	-0.005	0.007	-0.001	0.001	-0.002	-0.005	0.002	0.000
	(0.003)	(0.006)	(0.005)	(0.005)	(0.005)	(0.003)	(0.003)	(0.004)	(0.003)
Age^2	-0.000	-0.000	-0.000	-0.000	0.000	-0.000	0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Infant in Household	0.001	-0.008	-0.001	0.003	0.002	0.000	-0.001	0.003	-0.002
	(0.004)	(0.006)	(0.004)	(0.004)	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)
Child Under 5 in Household	-0.006	0.001	0.001	-0.002	-0.002	-0.001	0.000	-0.005	0.002
	(0.004)	(0.005)	(0.004)	(0.004)	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)
Single Parent of Infant	-0.003	0.006	0.000	-0.007	-0.004	-0.001	0.000	-0.003	0.000
	(0.005)	(0.007)	(0.005)	(0.005)	(0.006)	(0.004)	(0.004)	(0.004)	(0.003)
Single Parent of Child Under 5	0.006	0.002	-0.002	0.003	0.006	0.003	0.003	0.007*	0.000
	(0.004)	(0.006)	(0.005)	(0.004)	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)
(Single Parent) X (Num. Children)	0.001	0.005*	-0.001	0.002	0.001	0.001	0.002	-0.001	0.001
	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Currently in School	-0.003	0.009*	0.006	0.004	-0.003	0.005*	0.002	0.003	0.003
	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)	(0.002)	(0.003)	(0.002)	(0.002)
Any Employment Barrier Identified	-0.003*	-0.008***	-0.003*	-0.004*	-0.007***	-0.002	-0.003*	-0.002*	-0.002*
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Less than High School Equivalent	-0.009**	-0.009**	-0.011***	-0.004	-0.007**	-0.001	-0.007***	-0.001	-0.003*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)
TANF Receipt M-1	0.005*	0.003	0.003	0.007**	0.010**	0.002	0.003	-0.001	0.002
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
TANF Receipt M-2	-0.002	0.003	0.004	-0.003	-0.005*	-0.001	-0.001	0.002	0.001
	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)
TANF Receipt M-3	0.005*	-0.002	-0.002	0.001	-0.000	0.000	0.000	-0.001	-0.002

Continued on next page

Table A-4 – *Continued from previous page*

	2018Q1	2018Q2	2018Q3	2018Q4	2019Q1	2019Q2	2019Q3	2019Q4	2020Q4
	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
TANF Receipt M-4	-0.002	0.003	-0.002	-0.005	-0.002	-0.001	0.003	-0.001	0.001
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)
TANF Receipt M-5	0.002	-0.003	-0.001	-0.006	0.002	-0.000	-0.002	0.002	-0.003*
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
TANF Receipt M-6	0.002	0.000	-0.003	0.005	-0.005	0.000	-0.001	-0.001	0.001
	(0.002)	(0.003)	(0.003)	(0.004)	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)
TANF Receipt M-7	-0.006*	0.000	0.000	0.005	0.004	-0.005*	0.005*	0.001	0.002
	(0.003)	(0.004)	(0.002)	(0.004)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
TANF Receipt M-8	0.002	-0.003	0.007*	0.001	-0.008**	0.005*	-0.002	0.001	-0.006**
	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
TANF Receipt M-9	-0.001	-0.004	-0.001	0.001	0.009**	-0.001	-0.003	-0.005*	0.006**
	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
TANF Receipt M-10	0.003	-0.001	0.001	-0.005	-0.006	0.000	0.003	0.002	-0.002
	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
TANF Receipt M-11	-0.003	0.002	0.003	0.001	0.007*	0.001	0.002	0.000	-0.004*
	(0.003)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
TANF Receipt M-12	0.002	-0.001	-0.001	0.003	-0.001	0.002	-0.006**	0.001	0.005**
	(0.002)	(0.003)	(0.003)	(0.004)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Total TANF Benefits, 3 Years Prior	-0.000	0.000*	-0.000	-0.000	0.000	0.000	-0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Total TANF Months, M-36 to M-13	0.000	-0.001	0.000	0.000	-0.000	-0.000	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	13254	12655	12401	13111	15199	18697	18433	17990	17583

Marignal effects reported; Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A.4 Undercounting of In-Program Earnings in UI Data

As discussed in Section 2, STEP program contractors had multiple options for how to pay participants during their placements. They could serve as the employer of record and pay the participant directly, use a staffing agency as the employer of record, or reimburse internship host site employers for subsidized wages paid to participants. Panel a of Figure A-4 indicates that a substantial proportion of STEP participants had at least one in-program quarter in which they were in a STEP placement but had no formal-sector earnings in the UI data, likely due to participants being paid by host site employers in a manner other than standard W-2 employment. Each program contractor serves a distinct geographic area, and in some counties the program contractors appear to have paid participants in ways that typically generated UI earnings. We treat counties where at least 75 percent of participants have UI earnings in the quarter that contains the start date of their placement as "good coverage" counties. Panel c of Figure A-4 shows that, in these counties, the UI-based quarterly employment measure is never substantially lower than the share who are in a placement in that quarter. In contrast, when using data from all counties (panel b), the formal-sector employment rate is meaningfully lower than the share who are in a placement in quarters 0 and 1.

Figure A-5 presents employment and earnings trends splitting the sample into counties with good coverage of STEP placements in the UI earnings data and those with inconsistent coverage. As expected, in-program outcomes among the treatment group in good coverage counties (panels a and b) are reasonably insensitive to the use of imputation. In contrast, UI-based outcomes in poor coverage counties (panels c and d) meaningfully understate in-program improvements in employment and earnings.

Figure A-6 and Table A-5 present estimates of the effects of STEP participation for the same county split, using both the imputed measures of employment and earnings and the measures relying on formal-sector earnings data only. Post-program employment effects are nearly identical across the two groups of counties. Earnings effects differ somewhat, likely due

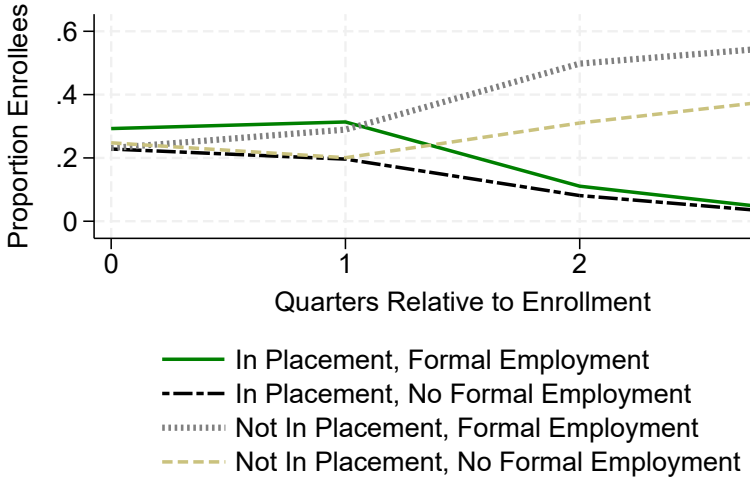
Table A-5: Aggregated Effects on Employment and Earnings, Split by Placement Earnings Coverage in UI Data

	Employment				Earnings			
	Imputed	Formal Only	Imputed	Formal Only	Imputed	Formal Only	Imputed	Formal Only
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
In-Program (Q0-Q3)	0.220***	0.189***	0.272***	0.0719***	639.0***	499.3***	869.2***	174.8
	(0.0236)	(0.0233)	(0.0175)	(0.0188)	(142.7)	(141.4)	(102.1)	(101.3)
Post-Program (Q4-Q7)	0.0961***	0.0961***	0.0966***	0.0910***	398.7*	398.7*	673.9***	640.0***
	(0.0249)	(0.0249)	(0.0217)	(0.0218)	(172.5)	(172.5)	(143.3)	(142.9)
UI Coverage of Placements	Good	Good	Poor	Poor	Good	Good	Poor	Poor
Comparison Mean (Q < 0)	0.472	0.472	0.492	0.492	1732.4	1732.4	1666.9	1666.9
Comparison Mean (Q0-Q3)	0.455	0.455	0.462	0.462	1630.3	1630.3	1613.4	1613.4
Comparison Mean (Q4-Q7)	0.421	0.421	0.439	0.439	2050.7	2050.7	2045.9	2045.9
N (Individuals)	9046	9046	18209	18209	9046	9046	18209	18209

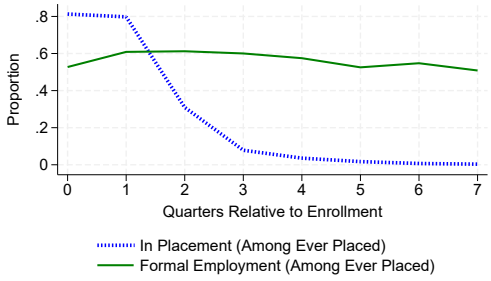
to differences in the prevailing wage across local labor markets. Importantly, the qualitative finding of meaningful but incomplete fadeout in program effects can be seen using only the non-imputed data in the good coverage counties.

Figure A-4: Formal-Sector Employment During STEP Placements

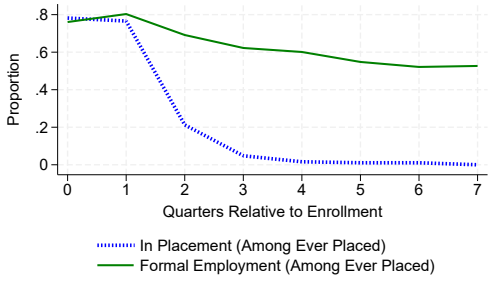
(a) All STEP Enrollees



(b) Enrollees with Placements, All STEP Counties



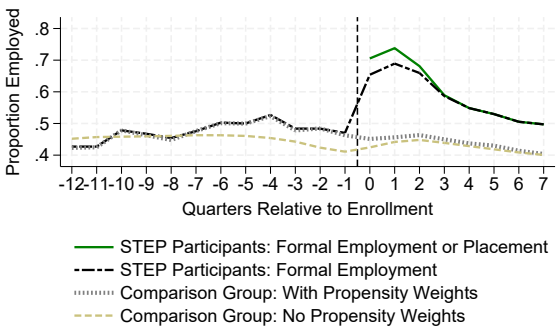
(c) Enrollees with Placements, Counties with Good Coverage of STEP Placements in UI Earnings Data



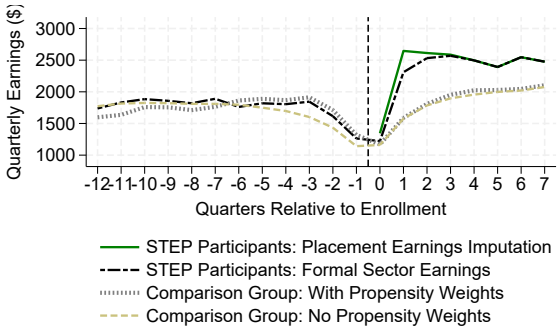
Notes: Data source is administrative UI earnings data and STEP program tracking data from CDLE. Panel a provides a decomposition of formal sector employment and placement rates in each quarter for all STEP enrollees who enrolled from January 2018 to March 2020. Panel b provides placement and formal-sector employment rates for all such enrollees who had a placement, and Panel c provides the same for enrollees in counties where placements nearly always resulted in formal-sector earnings.

Figure A-5: Split by Coverage of STEP Placements in UI Earnings Data

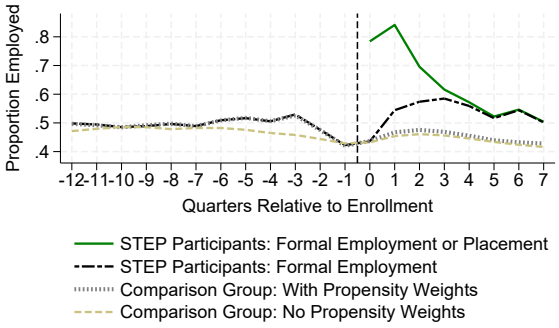
(a) Employment: Counties with Good Coverage of Placement Earnings



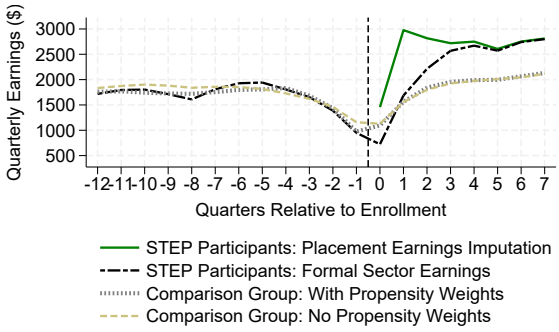
(b) Earnings: Counties with Good Coverage of Placement Earnings



(c) Employment: Counties with Poor Coverage of Placement Earnings

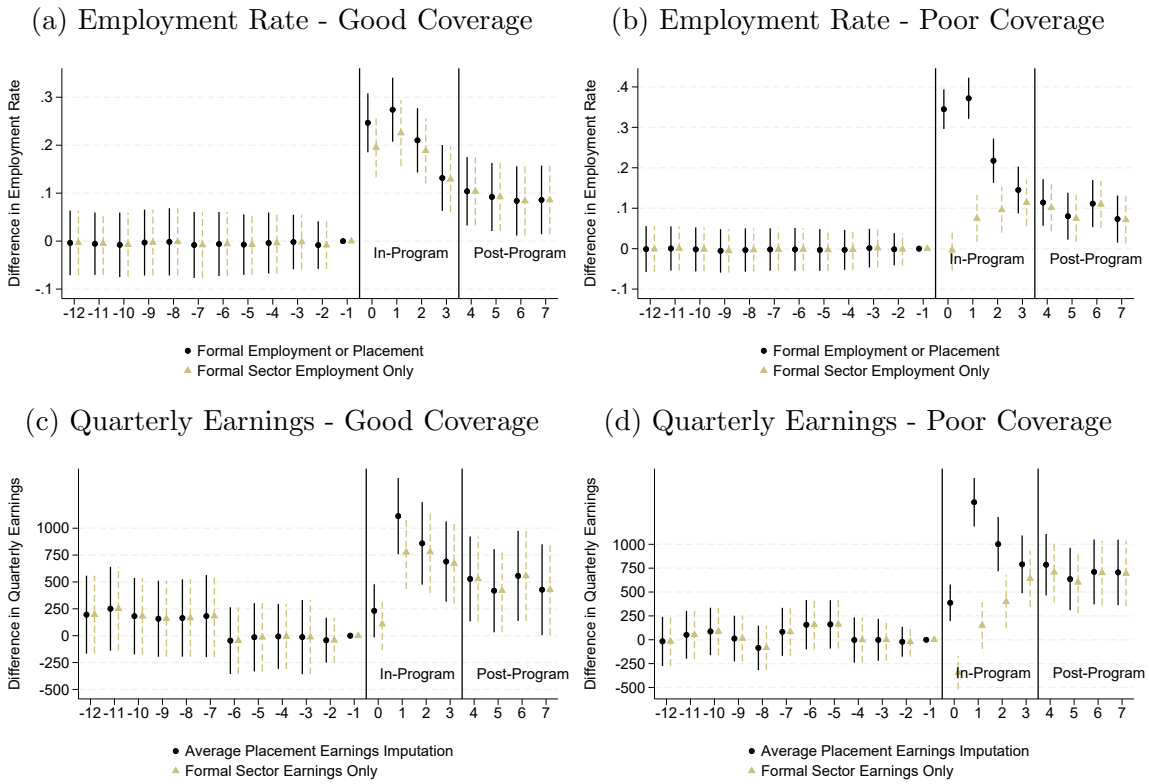


(d) Earnings: Counties with Poor Coverage of Placement Earnings



Notes: Data source is administrative UI earnings data from CDLE. Quarter 0 represents the quarter in which an individual enrolled in STEP and is thus a different calendar quarter from person to person. Formal employment is defined as having UI-covered earnings in Colorado greater than $0 in a given quarter. When earnings from a STEP-sponsored transitional job are reported to the UI system, they are included in formal sector employment/earnings. The imputation for placement-based earnings not reported to the UI system is described in section 3.1.

Figure A-6: Event-Study Effects on Employment and Earnings, by Counties' UI Coverage of Placement Earnings



Notes: This figure provides analysis comparable to the specifications in Figure 2 for subsamples of observations based on counties' coverage of in-placement earnings in the UI data. Panels a and c report estimates on employment and earnings outcomes, respectively, for counties where the vast majority of STEP participants have UI earnings during the quarters when they are listed as being in a placement. Panels b and d report employment and earnings effects for counties where a substantial portion of STEP participants have quarters for which they are listed as being in a subsidized placement but have no UI earnings. Quarter 0 represents the quarter in which an individual enrolled in STEP and is thus a different calendar quarter from person to person. Formal employment is defined as having UI-covered earnings in Colorado greater than $0 in a given quarter. When earnings from a STEP-sponsored transitional job are reported to the UI system, they are included in formal sector employment/earnings. The imputation for placement-based earnings not reported to the UI system is described in section 3.1. Data source is administrative UI earnings data from CDLE.

A.5 Additional Results and Robustness

Table A-6 provides aggregated versions of the results shown in the event study graphs in Figure 3 in the main text.

Table A-7 provides estimates from interaction specifications to test whether the differences in subgroup estimates in Table A-6 are statistically significant. The results demonstrate that the differences are not statistically different.

Figures A-7 and A-8 provide event study estimates based on the trends as shown in Figures 4 and 5 in the main text, respectively. The primary description of this analysis appears in the main text in section 5.1.3.

Table A-8 and Figure A-9 present aggregated and event-study estimates of STEP participation using the main earnings and employment outcome variables (with imputation) and varying the set of propensity weights used. Columns 1 and 3 reproduce the main estimates, in which the full covariate set is included in the propensity weights in addition to earnings and TANF receipt histories. Columns 2 and 4 use a "minimal" set of propensity weights using only earnings and TANF receipt history, in the spirit of Andersson et al. (2022). Lastly, Columns 3 and 6 use only the stacking weights with no propensity score adjustment. Results are qualitatively similar using all three weighting schemes.

Table A-9 reproduces the main employment and earnings estimates in columns 1 and 3 and compares them to a version in columns 2 and 4 dropping individuals who remain enrolled in the STEP program for more than three quarters (i.e., into the "post-program" period). These estimates are nearly identical to the main estimates, suggesting that the observed persistence of program effects after quarter 3 is not driven by these individuals with unusually long STEP enrollments.

Finally, Table A-10 provides heterogeneity analysis of program effects based on three key baseline characteristics: a lack of formal-sector employment in the year prior to program application (Panel A), having an employment barrier identified by the TANF caseworker (Panel B), and having a child under age 5 in the household (Panel C). Panel A shows that

Table A-6: Aggregated Effects on Employment and Earnings: Early vs. Late Cohorts

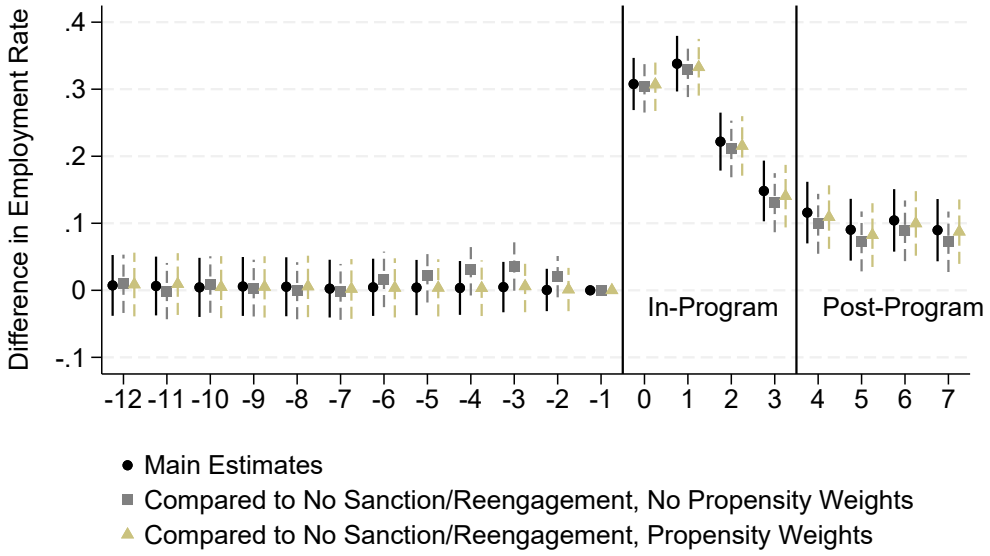
	Employment	Employment	Employment	Earnings	Earnings	Earnings
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: With Placement Employment and Earnings Imputation						
In-Program (Q0-Q3)	0.250***	0.266***	0.237***	766.4***	813.3***	729.1***
	(0.0150)	(0.0237)	(0.0185)	(87.79)	(139.1)	(109.9)
Post-Program (Q4-Q7)	0.0960***	0.102***	0.0896***	562.6***	752.0***	416.1**
	(0.0173)	(0.0270)	(0.0216)	(115.0)	(181.7)	(142.9)
Long-Term (Q8-Q12)		0.0705**			697.2***	
		(0.0265)			(194.3)	
Comparison Mean (Q < 0)	0.483	0.483	0.483	1704.0	1741.4	1678.6
Comparison Mean (Q0-Q3)	0.459	0.488	0.440	1640.2	1762.1	1557.5
Comparison Mean (Q4-Q7)	0.431	0.483	0.396	2058.0	2331.8	1872.1
Cohort	Full Sample	Early	Late	Full Sample	Early	Late
N (Individuals)	21740	10635	16400	21740	10635	16400
	(1)	(2)	(3)	(4)	(5)	(6)
Panel B: Formal Sector Employment and Earnings Only						
In-Program (Q0-Q3)	0.119***	0.133***	0.107***	297.1***	417.6**	206.7
	(0.0156)	(0.0248)	(0.0193)	(87.12)	(139.5)	(107.7)
Post-Program (Q4-Q7)	0.0927***	0.0953***	0.0887***	542.4***	714.7***	408.3**
	(0.0174)	(0.0270)	(0.0216)	(114.8)	(181.5)	(142.6)
Long-Term (Q8-Q12)		0.0705**			697.2***	
		(0.0265)			(194.3)	
Comparison Mean (Q < 0)	0.483	0.483	0.483	1704.0	1741.4	1678.6
Comparison Mean (Q0-Q3)	0.459	0.488	0.440	1640.2	1762.1	1557.5
Comparison Mean (Q4-Q7)	0.431	0.483	0.396	2058.0	2331.8	1872.1
Cohort	Full Sample	Early	Late	Full Sample	Early	Late
N (Individuals)	21740	10635	16400	21740	10635	16400

Notes: Data source is administrative UI earnings data from CDLE. Panels A and B report estimates on employment and earnings outcomes, respectively, for STEP participants who enrolled between January 2018 and March 2020. Quarter 0 represents the quarter in which an individual enrolled in STEP and is thus a different calendar quarter from person to person. Formal employment is defined as having UI-covered earnings in Colorado greater than $0 in a given quarter. When earnings from a STEP-sponsored transitional job are reported to the UI system, they are included in formal sector employment/earnings. The imputation for placement-based earnings not reported to the UI system is described in section 3.1. Estimates use the main set of stacked propensity weights.

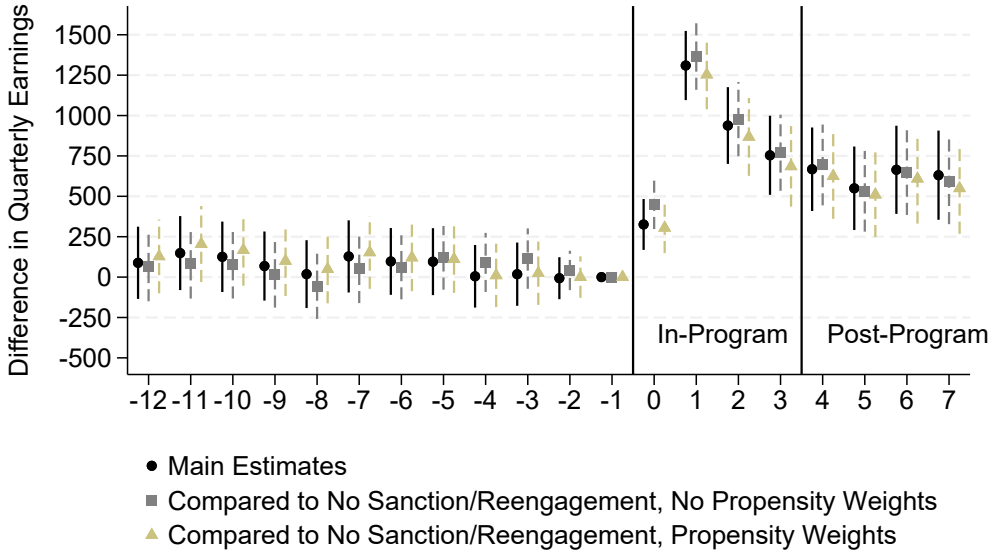
both employment and earnings effects are statistically significantly larger for participants without work experience in the year prior to application. The estimated differences in treatment effects in Panels B and C are not statistically significantly different from zero.

Figure A-7: Event-Study Effects on Employment and Earnings vs. No Sanction/Reengagement Comparison Group

(a) Employment Rate



(b) Quarterly Earnings



Notes: Data source is administrative UI earnings data from CDLE. Panel A and B report estimates on employment and earnings outcomes, respectively, for STEP participants who enrolled between January 2018 and March 2020. Quarter 0 represents the quarter in which an individual enrolled in STEP and is thus a different calendar quarter from person to person. Formal employment is defined as having UI-covered earnings in Colorado greater than $0 in a given quarter. When earnings from a STEP-sponsored transitional job are reported to the UI system, they are included in formal sector employment/earnings. The imputation for placement-based earnings not reported to the UI system is described in section 3.1.

Table A-7: Difference in Employment and Earnings Effects: Early vs. Late Cohorts

	(1) Employment	(2) Earnings
(Early Cohort) X (In-Program)	0.0277	80.12
	(0.0306)	(197.2)
(Early Cohort X (Post-Program)	0.00982	366.7
	(0.0352)	(259.9)
N (Individuals)	21673	21673

Notes: Data source is administrative UI earnings data from CDLE. Panels A and B report estimates on employment and earnings outcomes, respectively, for STEP participants who enrolled between January 2018 and March 2020. Estimates reported are interaction terms from an analysis analogous to the estimates above in Table A-6, but with the early and late cohorts pooled in the same regression to test for significant differences between the two groups. Estimates use the main set of stacked propensity weights.

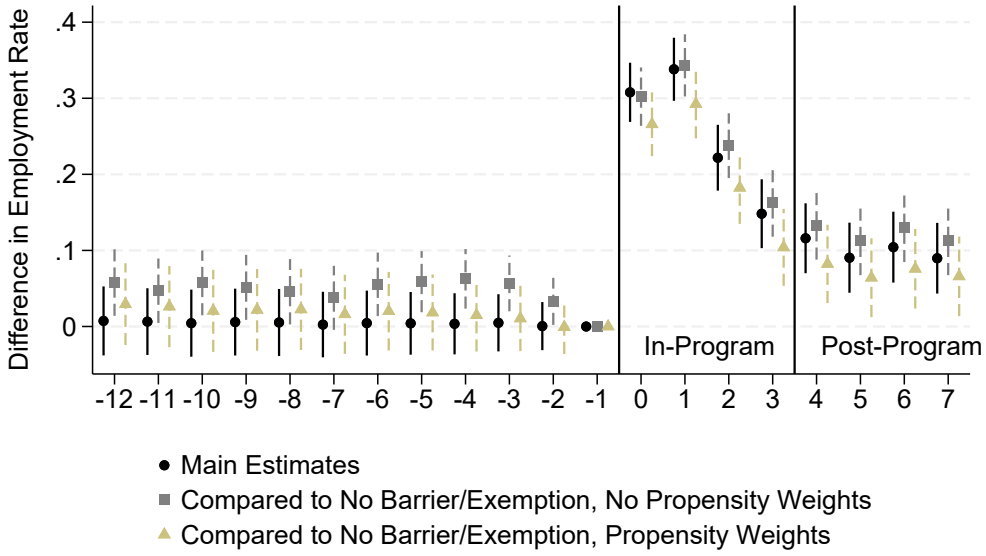
Table A-8: Aggregated Effects on Employment and Earnings, Varying Propensity Score

	Employment (1)	Employment (2)	Employment (3)	Earnings (4)	Earnings (5)	Earnings (6)
In-Program (Q0-Q3)	0.250***	0.255***	0.240***	766.4***	815.2***	802.6***
	(0.0150)	(0.0139)	(0.0139)	(87.79)	(82.89)	(83.12)
Post-Program (Q4-Q7)	0.0960***	0.0991***	0.0806***	562.6***	613.1***	592.6***
	(0.0173)	(0.0162)	(0.0162)	(115.0)	(109.1)	(109.4)
Comparison Mean (Q < 0)	0.483	0.486	0.463	1704.0	1711.4	1715.2
Comparison Mean (Q0-Q3)	0.459	0.458	0.449	1640.2	1598.8	1615.2
Comparison Mean (Q4-Q7)	0.431	0.431	0.426	2058.0	2014.9	2039.2
Propensity Weights	Main	Minimal	None	Main	Minimal	None
N (Individuals)	21740	25903	25903	21740	25903	25903

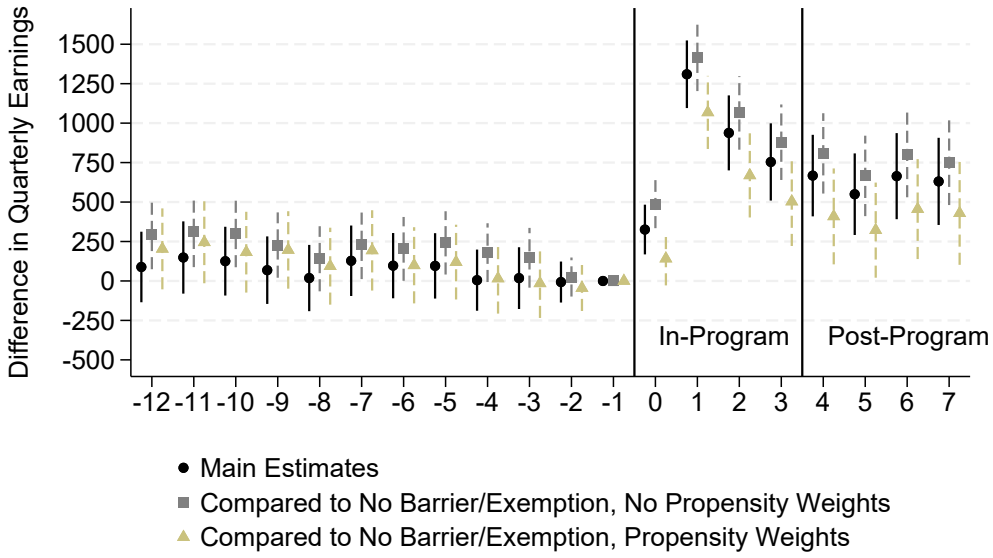
Notes: Data source is administrative UI earnings data from CDLE. Quarter 0 represents the quarter in which an individual enrolled in STEP and is thus a different calendar quarter from person to person. Formal employment is defined as having UI-covered earnings in Colorado greater than $0 in a given quarter. When earnings from a STEP-sponsored transitional job are reported to the UI system, they are included in formal sector employment/earnings. The imputation for placement-based earnings not reported to the UI system is described in section 3.1.

Figure A-8: Event-Study Effects on Employment and Earnings vs. No Barrier/Exemption Comparison Group

(a) Employment Rate



(b) Quarterly Earnings



Notes: Data source is administrative UI earnings data from CDLE. Panel A and B report estimates on employment and earnings outcomes, respectively, for STEP participants who enrolled between January 2018 and March 2020. Quarter 0 represents the quarter in which an individual enrolled in STEP and is thus a different calendar quarter from person to person. Formal employment is defined as having UI-covered earnings in Colorado greater than $0 in a given quarter. When earnings from a STEP-sponsored transitional job are reported to the UI system, they are included in formal sector employment/earnings. The imputation for placement-based earnings not reported to the UI system is described in section 3.1.

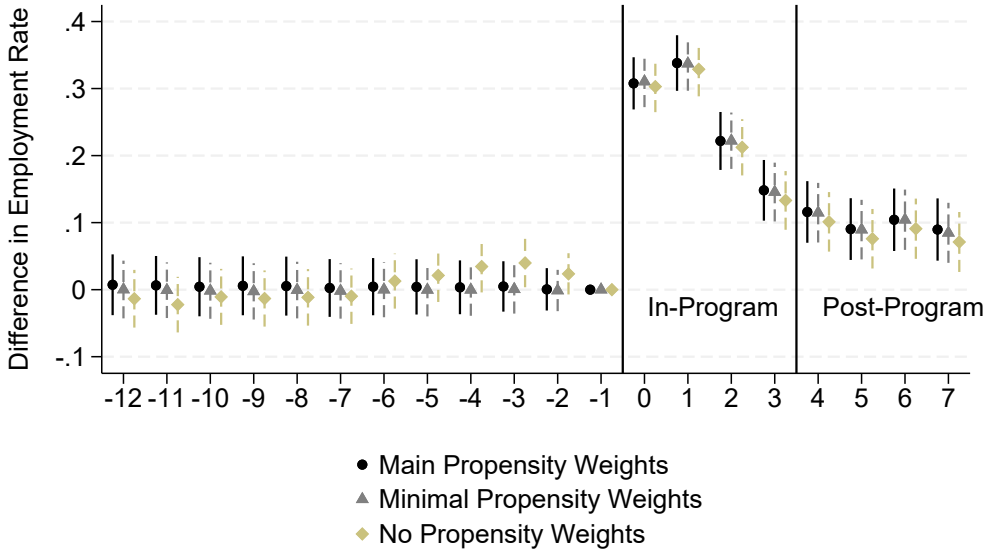
Table A-9: Aggregated Effects on Employment and Earnings, Dropping Individuals Enrolled Past Quarter 3

	Employment	Employment	Earnings	Earnings
	(1)	(2)	(3)	(4)
In-Program (Q0-Q3)	0.250***	0.250***	766.4***	856.5***
	(0.0150)	(0.0154)	(87.79)	(95.00)
Post-Program (Q4-Q7)	0.0960***	0.100***	562.6***	711.6***
	(0.0173)	(0.0179)	(115.0)	(130.5)
Comparison Mean (Q < 0)	0.483	0.483	1704.0	1701.9
Comparison Mean (Q0-Q3)	0.459	0.458	1640.2	1632.8
Comparison Mean (Q4-Q7)	0.431	0.430	2058.0	2049.9
Sample	Full	Drop Q>3 Enrolled	Full	Drop Q>3 Enrolled
N (Individuals)	21740	21673	21740	21673

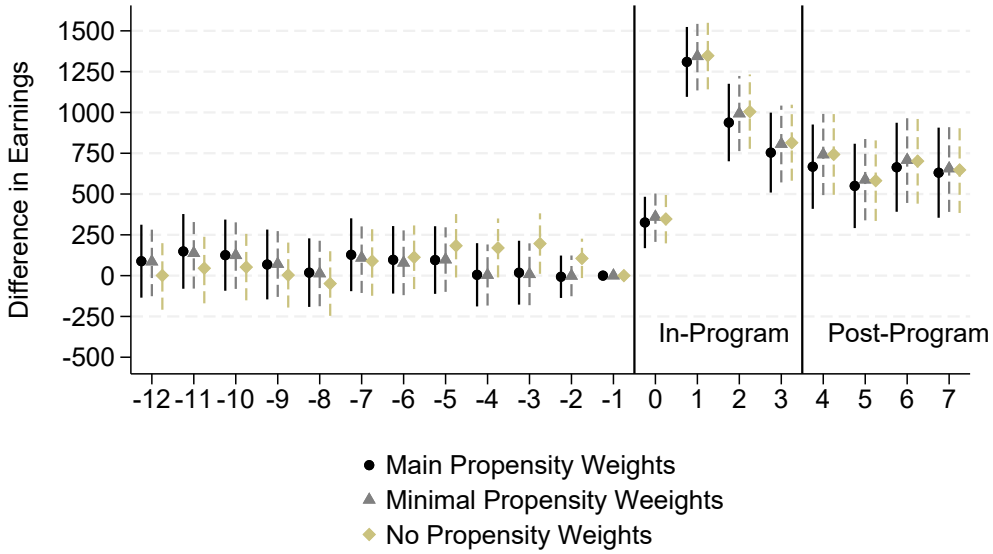
Notes: Data source is administrative UI earnings data from CDLE. Quarter 0 represents the quarter in which an individual enrolled in STEP and is thus a different calendar quarter from person to person. Formal employment is defined as having UI-covered earnings in Colorado greater than $0 in a given quarter. When earnings from a STEP-sponsored transitional job are reported to the UI system, they are included in formal sector employment/earnings. The imputation for placement-based earnings not reported to the UI system is described in section 3.1.

Figure A-9: Event-Study Effects on Employment and Earnings, Varying Propensity Score

(a) Employment Rate



(b) Quarterly Earnings



Notes: Data source is administrative UI earnings data from CDLE. Quarter 0 represents the quarter in which an individual enrolled in STEP and is thus a different calendar quarter from person to person. Formal employment is defined as having UI-covered earnings in Colorado greater than $0 in a given quarter. When earnings from a STEP-sponsored transitional job are reported to the UI system, they are included in formal sector employment/earnings. The imputation for placement-based earnings not reported to the UI system is described in section 3.1.

Table A-10: Employment and Earnings Heterogeneity

	(1) Employment	(2) Earnings
Panel A: No Employment in Prior Year		
In-Program (Q0-Q3)	0.194***	510.9***
	(0.0161)	(101.8)
Post-Program (Q4-Q7)	0.0653***	315.4*
	(0.0189)	(128.4)
(Prior Unemp.) X (In-Program)	0.193***	875.4***
	(0.0303)	(176.7)
(Prior Unemp.) X (Post-Program)	0.104**	845.9***
	(0.0356)	(252.4)
Panel B: Employment Barrier Identified by Caseworker		
In-Program (Q0-Q3)	0.230***	655.8***
	(0.0248)	(155.1)
Post-Program (Q4-Q7)	0.0642*	297.5
	(0.0274)	(190.8)
(Barrier) X (In-Program)	0.0325	182.6
	(0.0309)	(185.8)
(Barrier) X (Post-Program)	0.0523	436.5
	(0.0351)	(237.4)
Panel C: Child Under 5 in Household		
In-Program (Q0-Q3)	0.265***	805.2***
	(0.0235)	(146.8)
Post-Program (Q4-Q7)	0.122***	619.0**
	(0.0274)	(190.4)
(Child Under 5) X (In-Program)	-0.0258	-64.07
	(0.0300)	(180.1)
(Child Under 5) X (Post-Program)	-0.0440	-95.37
	(0.0349)	(236.4)
N (Individuals)	21740	21740

Notes: Data source is administrative UI earnings data from CDLE. Panels A and B report estimates on employment and earnings outcomes, respectively, for STEP participants who enrolled between January 2018 and March 2020. Quarter 0 represents the quarter in which an individual enrolled in STEP and is thus a different calendar quarter from person to person. Formal employment is defined as having UI-covered earnings in Colorado greater than $0 in a given quarter. When earnings from a STEP-sponsored transitional job are reported to the UI system, they are included in formal sector employment/earnings. The imputation for placement-based earnings not reported to the UI system is described in section 3.1.

Table A-11: Aggregated Effects of STEP Participation on TANF Receipt, by Cohort

	(1) Participation	(2) Participation	(3) Participation	(4) Monthly Benefit	(5) Monthly Benefit	(6) Monthly Benefit
In-Program (Mth. 0-9)	0.131***	0.120***	0.138***	72.68***	68.76***	75.69***
	(0.0127)	(0.0192)	(0.0168)	(6.862)	(10.34)	(9.053)
Post-Program (Mth. 10-21)	0.0174	0.000289	0.0298	12.76	5.901	17.64
	(0.0151)	(0.0221)	(0.0202)	(7.736)	(11.30)	(10.37)
Long-Term (Mth. 22-36)		0.000275			-0.986	
		(0.0208)			(10.40)	
Comparison Mean (Mth. < 0)	0.287	0.276	0.295	126.5	119.2	131.4
Comparison Mean (Mth. 0-9)	0.587	0.576	0.595	281.2	269.5	289.2
Comparison Mean (Mth. 10-21)	0.272	0.260	0.280	133.7	128.0	137.5
Comparison Mean (Mth. 22-36)		0.153			77.35	
Cohort	Full Sample	Early	Late	Full Sample	Early	Late
N (Individuals)	21740	10635	16400	21740	10635	16400

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Data source is administrative TANF benefits data from the CBMS CUBS data system. Figure reports event study estimates on a binary outcome for TANF receipt, respectively, for STEP participants who enrolled between January 2018 and March 2020. Month 0 represents the month in which an individual enrolled in STEP and is thus a different calendar month from person to person. TANF receipt is defined as having any TANF payment in Colorado greater than $0 in a given month. Earnings from a STEP-sponsored transitional job do not count against the TANF eligibility threshold and in fact require the individual to remain enrolled in TANF during the transitional job. Sample sizes differ across propensity scores because some individuals in the comparison group are assigned missing values of the propensity score because they differ from all treated observations in their cohort on some discrete characteristic included in the propensity score estimation. We allow these observations to be automatically excluded from the comparison group depending on which propensity score is used.

A.6 Marginal Value of Public Funds Calculation

Participants' willingness to pay for the program is calculated as discounted after-tax earnings gains, using the estimated tax rates from Hendren and Sprung-Keyser (2020) based on each participant's earnings as a percentage of the federal poverty level. Panel A of Table 8 uses the estimated cohort-specific in-program and post-program earnings effects in columns 5 and 6 of Table A-6, directly observing effects through quarter 12 for the early cohort and through quarter 7 for the late cohort. For simplicity, we assume constant treatment effects over time within quarters 0-3, 4-7, and 8-12. In Panel B, we use the same directly estimated effects and assume that the later cohort of enrollees' earnings effects from quarters 4-7 remain constant through quarters 8-12, because only 7 quarters of data are observable for this group.

To calculate gross program costs we take the total STEP program expenditures of $8,734,906 and divide by the total number of STEP participants to get the cost per participant, and we then distribute that cost evenly across quarters 0-3 for each STEP participant. We drop from the denominator all participants who enrolled in STEP after March 2020, assuming that they received minimal benefits from the program due to the pandemic. We keep enrollees who were dropped from the main analysis sample for data quality reasons (see Appendix Section A.2), assuming that their program benefits are similar to the main sample. These restrictions result in a denominator of 1,182 participants. Gross costs are set to zero from quarter 4 onward, and we subtract participants' estimated additional taxes paid from those gross costs to calculate the net program cost for each participant in each quarter. We sum each participant's willingness to pay and net costs across all quarters and divide to get the MVPF.