I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 18342

Supported Work Leads to Lasting Labor Market Success Among TANF Recipients

Tania Barham
Brian C. Cadena
Lauren Schechter

DECEMBER 2025

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 18342

Supported Work Leads to Lasting Labor Market Success Among TANF Recipients

Tania Barham
University of Colorado Boulder

Brian C. Cadena
University of Colorado Boulder and IZA

Lauren Schechter
University of Notre Dame

DECEMBER 2025

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

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

This paper studies the effects of a supported work program that provides TANF recipients with a suite of services including a six-month subsidized internship with a local employer. We use rich administrative data and implement a stacked difference-in-differences design comparing program participants to observably similar TANF recipients to estimate effects on employment, earnings, and benefit receipt. Program enrollment led to an immediate increase in formal-sector employment and earnings, with limited post-program fadeout. The program increased employment by 10 percentage points (20 percent) and earnings by \$861 per quarter (48 percent) in the three years following program exit. Program participation also increased participants' total benefit receipt during the program, with modest decreases after program exit. The program is relatively cost-effective compared to other adult subsidized employment programs due to longer-than-average persistence of the employment and earnings gains.

JEL Classification: J24, J68, I38, H43

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

Corresponding author:
Brian C. Cadena
Economics Building
University of Colorado
256 UCB
Boulder, CO 80309
USA

E-mail: brian.cadena@colorado.edu

* We would like to acknowledge helpful comments from Patrick Turner, Chloe East, Rob Santillano, Sheena Murray, and seminar participants at the University of Notre Dame, Oregon State University, Miami University, and the Virtual Economics of Poverty and Policy Seminar. 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 and Sara Truesdale for excellent research assistance. This work would not be possible without anonymized data provided by the Linked Information Network of Colorado (LINC). The findings do not necessarily reflect the opinions of the Colorado Governor's Office of Information Technology, the Colorado Evaluation and Action Lab, or the organizations contributing data.

1 Introduction

The structure of the social safety net in the United States has changed substantially in recent decades, with the size and importance of cash welfare programs declining steadily both in absolute terms and relative to other forms of assistance (DHHS, 2024).[1] There is a broad consensus that these reforms have led to substantial earnings and consumption gains among affected families — typically poor, single parent-headed families with young children (Meyer and Sullivan, 2008, Han, Meyer and Sullivan, 2021, Kleven, 2024). However, the majority of federal safety net spending now accrues to households with at least some positive earnings and income above the poverty line (Hoynes and Schanzenbach, 2018), because programs like the Earned Income Tax Credit (EITC) or Unemployment Insurance (UI) require that applicants meet a minimum earnings threshold for eligibility. While cash welfare — now administered through the Temporary Assistance for Needy Families (TANF) program — serves families closer to the bottom of the income distribution, it is less effective as an automatic stabilizer when labor markets are weak, exhibiting no significant response to recent recessions (Bitler and Hoynes, 2016). In this modern safety net 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 can improve labor market outcomes for TANF recipients. The Colorado Works Subsidized Training and Employment Program (CW-STEP or STEP) offered individualized employment supports, including job placement assistance, work experience opportunities, and supportive services to help participants resolve barriers to employment such as unreliable transportation or

[1]The Aid to Families with Dependent Children (AFDC) program, and its predecessor, the Mothers' Pensions program, were implemented to increase the ability of poor single mothers 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. Beginning in the 1970s and growing in prevalence through the mid-1990s, a series of reforms introduced welfare-to-work initiatives that were intended to increase recipients' labor market attachment and decrease their reliance on benefits. The 1993 expansion of the Earned Income Tax Credit and the 1996 PRWORA welfare reform replacing AFDC with Temporary Assistance for Needy Families (TANF), which includes work requirements and time limits for recipients, further shifted the safety net toward assistance contingent on work.

childcare. Participants were matched to one or more activities based on their skills and interests, ranging from volunteer opportunities to paid temporary placements 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, by 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 base-

line 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 26 percentage points (52 percent) and quarterly earnings by $856 (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 three years following exit from the STEP program, STEP participation increased formal sector employment by 10 percentage points and quarterly formal sector earnings by $861.

Because STEP participants self-select into the program, we also show that the estimated program effects are qualitatively similar when using two 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 recorded by the TANF caseworker of failure to follow a case plan and take the recommended steps to find work. 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, as 27% of STEP participants have a recent sanction or re-engagement designation and 61% have a documented barrier to work—sometimes resulting in an exemption

from work requirements—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 could be positively selected on post-enrollment potential outcomes. Furthermore, we show that the results are robust to limiting the analysis to counties where STEP internships are almost always captured in UI earnings data, to dropping participants whose service receipt extends into the second year after enrollment and to changing the set of observable characteristics used to reweight the comparison group.[2]

STEP participation also led to more stable employment. Participants were 6 percentage points (40 percent) more likely than their non-participating TANF counterparts to remain continuously employed with the same employer for an entire year over the three years following program exit. They were also more likely to move into sectors with longer-than-average job duration. These 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 benefit payments 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.[3] We find

[2] 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.

[3] Grogger (2003) finds that some recipients do strategically and voluntarily exit welfare caseloads to

no impact on TANF receipt in months after participants have exited the program. By contrast, we find that the program increased participants' attachment to SNAP (likely through categorical eligibility from TANF participation) but reduced their average SNAP payment while they were participating in the program. After program exit, we find reductions in average monthly SNAP receipt, with suggestive evidence that this reduction comes in part from reduced SNAP participation rather than from reduced payment amounts alone.

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 like work requirements that focus on strengthening incentives to work through penalties for non-compliance 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 evaluating 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.[4] Another recent program with standardized placements that targeted TANF recipients was

conserve future eligibility in response to time limits. The federal lifetime limit for TANF participation is 60 months.

[4]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.

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 explicitly not intended to lead to permanent employment. Earnings impacts while participants were in their transitional jobs were smaller than our in-program estimates from STEP, and they faded out completely after participants left subsidized employment.

A limited literature examines subsidized employment programs that offered the potential for permanent job placement but with different target 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 programs, which provided similar employment services but were not bundled with TANF and therefore did not result in additional cash assistance for participants on top of the subsidized wages. 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 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 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 pro-

grams 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 for social safety net programs have been effective in offsetting disincentives to work. 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 stricter sanctions for TANF recipients failing to meet work requirements produce a small and short-lived labor supply response that is not large enough 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 by various states in the 1980s and early 1990s, before the federal 1996 Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA) instituted broad work requirements and time limits for cash welfare recipients. These programs typically combined changes in the incentive structure of cash welfare programs with some type of employment supports or job search assistance.[5] 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

[5] A similar policy experiment in Canada paired supportive services with a substantial earnings boost designed to encourage welfare recipients to leave the rolls. The results are consistent with the interpretation that the combination of incentives and supports is more effective than incentives alone (Bobonis et al., 2022).

[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).

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

[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

participants to engage in a period of supported work with a local employer. The program is administered jointly by the Colorado Department of Human Services (CDHS) and the Colorado Department of Labor and Employment (CDLE) and 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 during 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.

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, local contractors reimbursed the host site employer for wages paid during the internship.[10] 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. Private, public, and nonprofit sector employers were all eligible to be internship host sites. Most placements were structured as trial positions that could ideally

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.

[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]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.

lead to a hire into a regular position at the host site if successful. The STEP operations guide made clear that employers were under no obligation to hire successful participants but implied that, ideally, the host-site employer would consider the participant for an open role before conducting a full search.[11] 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.[12]

Importantly, any earnings from STEP internships were fully disregarded when calculating 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

[11] 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."

[12] 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).

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.[13] 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.[14]

2.2 Recruitment, Enrollment, and Service Receipt Timing

TANF case managers recruited 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 enrolled in STEP after receiving a referral from their TANF case manager and, in some cases, after meeting with a representative

[13] 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.

[14] 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.

of the STEP contractor.[15] TANF recipients who chose not to enroll in the STEP program had to find another way to satisfy work requirements in order to remain eligible for TANF benefits at subsequent six-month re-certification windows, either by finding unsubsidized employment or participating in other job training or education programs.

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. 65 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 time trends in internship placement and STEP program exit).

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 2014 to March 2024. First, the Colorado Benefits Management System (CBMS) provides data on the universe of Colorado adults 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

[15] 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.

it difficult for the client to find work, and sanctions or case plan re-engagement designations indicating failure to follow the case plan.[16] 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 and SNAP benefit payment amounts, and we analyze both binary program participation and monthly payment amounts as secondary outcomes.

Second, STEP participants are identified from the Colorado Department of Labor and 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. Less than one percent (10/1,201) 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 (W-2) employment in Colorado but do not capture earnings from self-employment, employment as an independent contractor, informal work, or 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

[16]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.

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-3), 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.[17] We then aggregate these imputed earnings to the quarterly level and add them to any observed UI earnings in that quarter. 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 STEP placement for any portion of that quarter. Although the outcome variables built from these data 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.[18]

[17] 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.

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

3.2 Analysis Sample

We limit the analysis sample to individuals whose outcomes were consistent with program eligibility—specifically, to individuals who received a TANF payment during the program's operation and who 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.[19] The final sample consists of 1,072 STEP participants and 26,906 program-eligible TANF recipients in the comparison group. See Appendix Section A.2 for details on sample restrictions.

3.3 Treatment Group Characteristics

The first column of Table 2 shows that the characteristics of STEP participants broadly mirrored the TANF recipient population from which they were drawn. Household and demographic variables measured as of the STEP enrollment month in the CBMS data show that most participants were women with children in households without another adult. The majority of participants had at least one child under 5 and at most a high school diploma or GED. UI earnings histories indicate that during the four years leading up to enrollment, participants earned on average about $6,000 to $7,000 per year. About 60 percent 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 the standard deviation of each variable for the treatment group (column 2), with unweighted means of the same characteristic for the comparison group (column 3) and corresponding standard deviations (column 4), differences in means (column 5), and t-statistics from a test of equal means (column 6). There are meaningful differences

[19] As of January 2018, 74% of adults on the Colorado TANF caseload lived in a STEP county. Non-participating counties are on average substantially smaller and more rural.

in these baseline characteristics: relative to the comparison group, treatment group members were 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. The treatment group also had higher earnings over the four years preceding their enrollment in the program. Some of these differences suggest the potential for positive bias—for example, participants without infant care needs or other work barriers may have had an easier time recovering their labor force attachment and earnings. Alternatively, because the treatment group had stronger earnings 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 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: 16 quarters prior to and 16 quarters after enrollment for earnings and employment outcomes, and 48 months prior to and 48 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. 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.

This stacking procedure allows us to limit each cohort's comparison group to individuals who never enroll in the STEP program, preventing comparisons between newly-treated and earlier-treated individuals that have been shown to produce bias in staggered difference-in-differences designs (Goodman-Bacon, 2021). We further limit each cohort's comparison group to individuals who were receiving TANF in the potential STEP enrollment month and therefore eligible to enroll in the program in that month, which would not be possible using some other staggered difference-in-difference estimators that avoid biased comparisons but do not allow for time-varying inclusion criteria for the comparison group (Callaway and Sant'Anna, 2021, Sun and Abraham, 2021, De Chaisemartin and d'Haultfoeuille, 2020).

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

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

where D_i is an indicator for STEP enrollment by individual i, m is the potential STEP enrollment 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 were receiving services including subsidized placements and other supports; and $PostProgram_{mt}$ is an indicator for relative quarters $t =$ 4–16 (months $t =$ 10–48) when the vast majority of participants had exited the program. The coefficient β_3 therefore esti-

mates 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 three years after entry.[20] 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. We also estimate the analogous event study regression to reveal more detailed dynamic effects in each individual quarter (month) relative to STEP enrollment. Balancing the dataset in event time ensures that event-study estimates are unaffected by compositional change.

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.

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

[20] 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. 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.

schooling, barriers to employment, and county.[21] 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-2 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 enrollment 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 correctly identifies the causal effect of program participation under an equal counterfactual trends assumption—non-participants' average change in outcomes must be exactly what would have occurred among the treatment group

[21] 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-2. 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

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.

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. Column 5 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 6), indicating that it is possible the equal counterfactual trends assumption may hold, as the treatment group and comparison group are balanced on a rich set of baseline characteristics that may predict post-enrollment employment and earnings trajectories.

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.[22]

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.

[22]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

Figure 1 uses the stacked dataset to present trends over time in earnings and employment 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 to each participant's STEP enrollment date. Quarter zero indicates the enrollment quarter; negative values indicate quarters prior to enrollment, and positive values indicate periods after enrollment. The black dashed line provides trends for the treatment group using only UI earnings records. The solid green line includes imputed earnings from STEP placements as described in Section 3.1 and begins in period 0 by definition because no one has earnings from an internship prior to enrollment. Both panels include two versions of the comparison group's trend lines to highlight the value of reweighting the comparison group to match baseline observables. The first (gray dotted line, main propensity weights) reweights using propensity scores estimated on the full set of available baseline characteristics (see Appendix Table A-2 for further details on this estimation). 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 prior to quarter 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

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.

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

Tables 2 and 3 demonstrate that reweighting almost completely eliminates treatment-comparison differences in observable baseline characteristics in addition to producing similar pre-trends in earnings and employment. 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. 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. 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. In that same vein, we report estimates using a "simple" propensity score estimated using only earnings and benefits histories in addition to the "main" propensity score based on the full set of covariates reported in Appendix Table A-2.

5 Results

Overall, we find that participation in the STEP program increased participants' employment and earnings both during the program and for approximately three years after exiting the program. The program also caused participants to remain more attached to the social safety net for the duration of their enrollment. However, after participants exited the program, they were no more likely than their counterparts from the same TANF cohort to receive TANF benefits. The STEP program also increased participants' SNAP participation but decreased SNAP benefit amounts in the first nine months after enrollment, with suggestive

evidence of modest decreases in SNAP receipt in month ten and beyond.

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 the main propensity weights that make use of the full set of available covariates (black circles), simplified propensity weights using only earnings and benefits history variables (gray squares), and no propensity weights at all (gold triangles). Estimated effects are nearly identical across the three weighting schemes.

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 approximately 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 16, 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 $800-900 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 4–16) using the specification in Equation 2. Our preferred estimates in column 1 re-weight the comparison group using the full set of available covariates in the propensity weights; estimates in column 2 use simplified propensity weights including only

earnings and benefits histories as in Andersson et al. (2022) and Grosz (2020); and those in column 3 use only the stacking weights with no covariate adjustment. The preferred specification indicates that program participation increased participants' quarterly employment rate by 26 percentage points (52 percent relative to the comparison group baseline mean of 0.49) and their formal sector earnings by about $856 per quarter (47 percent of the comparison baseline mean of $1806) while enrolled.[23] Although employment effects diminish after program exit, substantial gains persist through quarters 4–16: during this period, the program raised employment by about 10 percentage points (20 percent). Quarterly earnings gains—with zeroes included in the outcome variable—remain stable at about $861 per quarter (47 percent) despite the fade-out in employment effects, implying growth in earnings among STEP enrollees who remain employed after exiting the program.

5.1.2 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 selected 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 four years leading up to the potential enrollment month, but their untreated outcomes may also be affected by 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 vari-

[23]The smaller estimated impacts when using only stacking weights derive in part from the gap in pre-enrollment earnings between the treatment and comparison groups. Reweighting closes this gap and aligns the pre-trends, which is why is it is our preferred specification.

ables 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 all available covariates in the propensity score, excluding those used to form the subset, as in column 1 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 Table 2, the treatment group was less likely than the comparison group to have either of these designations, but nevertheless about 27% had at least one or the other, 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 3 demonstrates that the unweighted employment rate and earnings for this new comparison group are nearly as high as the treatment group before the enrollment month (panels a and 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.[24] 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

[24]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-6.

motivated group.

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, or housing difficulties. These barriers are often noted by the caseworker, and mitigation efforts are included in the recipient's case plan. Some severe 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 Table 2, 60% of the treatment group (vs. 66% of the comparison group) had an employment barrier identified by their caseworker during the TANF spell of their STEP enrollment, and 6% (vs. 11% 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 4 shows outcome trends, with the comparison group split by the presence/absence of a work barrier or work requirements exemption. Those with no barrier or exemption are roughly 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 have similar average earnings (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 do. This pattern suggests that the absence of barriers or 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.[25]

Results in Panel A of Table 5 indicate that estimates of the STEP program's effects on both in-program and post-program employment are similar to the main estimates (column 1) in sign and magnitude even when the comparison group is limited to TANF recipients without a sanction or re-engagement designation (columns 2 and 3) or without an employment barrier or work requirements exemption (columns 4 and 5), regardless of which weighting scheme is used. Panel B shows that estimated earnings effects using these subsamples are smaller in magnitude but are still economically meaningful and statistically significant; the most conservative estimate in column 4 indicates that the program increased participants' post-program earnings by approximately $688 per quarter (39 percent, $p < 0.001$).

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 case plan compliance and barriers provides additional support for the equal counterfactual trends assumption beyond the equal pre-trends.

Robustness analyses concerning the treatment of undercounted earnings from STEP placements in the UI earnings data are reported in Appendix Section A.2. First, for completeness, Appendix Figure A-4 and Table A-3 present results with and without imputing employment and earnings outcomes for STEP placements not appearing in the UI earnings

[25]Corresponding event-study estimates are presented in Appendix Figure A-7.

data. We prefer the main estimates with the imputed earnings included, because the estimates without imputation understate the extent to which employment outcomes partially fade out as participants exit the program. Importantly, 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-4 and Figure A-5 show that post-program employment and earnings effects are similar in magnitude and significance across the two sets of counties.

Additional robustness analyses are reported in Appendix Section A.5. Figure A-8 and Table A-5 present earnings effects varying the threshold at which UI earnings data are winsorized to reduce the influence of outliers. The main specification in which we winsorize UI earnings data at the 95th percentile within each cohort-quarter cell is the most conservative, but winsorizing at the 98th and 99th percentile produce qualitatively similar results. We also show in Appendix Table A-6 that results are robust to dropping the few individuals in the treatment group that either remain enrolled in the STEP program after quarter three or have missing program exit dates, meaning that the persistence of employment and earnings effects is unlikely to be driven by individuals who continue receiving services into the "post-program" period.

Lastly, one might worry that estimated treatment effects are driven in part by adverse spillovers on comparison group members who may have otherwise held the jobs obtained by STEP participants. Because only three percent of the eligible TANF caseload is ever treated by the program, this phenomenon would not meaningfully reduce the employment rate of the comparison group, and thus cannot contribute a substantial portion of the estimated treatment effect. Regardless, we show in Appendix Figure A-9 that comparison group members in counties with a higher proportion of the TANF caseload treated are no worse off than those in counties with a lower proportion of the caseload eventually enrolled in the STEP program. Finally, we show below in Section 5.2 that the program moved STEP participants

into new industries, further reducing the likelihood they displaced their comparison group counterparts. We conclude that negative spillovers on the comparison group are unlikely to be an important channel.

5.1.3 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. First, in Appendix Table A-7, we present heterogeneity analysis based on the same measures discussed above: whether the individual had a sanction/re-engagement or barrier/exemption in the current TANF spell. Results indicate that individuals with barriers or exemptions may benefit more from the program than other STEP participants do, especially in the initial period. In contrast, there is no statistically significant difference in program effects based on whether an individual has a sanction or re-engagement designation.

We also explore heterogeneity based on the age of each individual's youngest child in the household and the number of cumulative TANF months observed for that person as of the enrollment month. Grogger (2003) shows that individuals strategically exit the TANF caseload to conserve future eligibility in light of the 60-month lifetime limit, so individuals with more months of potential future TANF eligibility may respond differently to the program requirement that STEP participants remain on TANF throughout their time in the program. Results in Appendix Table A-8 indicate no difference in estimated program effects on either of these dimensions.

5.2 Mechanisms

The STEP program, like its sister program ReHire Colorado, 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.

STEP program records do not contain information about the identities or characteristics of the internship host site employers and therefore do not allow us to observe directly whether internships are converted to permanent positions at the host site. However, those hired into regular positions at the conclusion of their internships should have more stable employment with the same employer throughout the three years we are able to observe after their exit from the program (quarters 4 through 16 after enrollment).[26] Aggregating participants' outcomes to years rather than quarters relative to STEP enrollment and using an analogous stacked, propensity-weighted difference-in-differences specification to that in Equation 3, we test whether STEP participants are more likely than non-participants to be stably employed during the post-program period, and whether this difference is driven by continuous employment with a single employer. We drop the year of program enrollment (quarters 0-3 from the main analysis) and use only the UI earnings data, since employer identities are not available in the STEP program tracking data. We instead focus on the three years before enrollment (quarters -16 through -1 in the main analysis) and the three years after (quarters 4–15), when most STEP enrollees have exited their placements and UI earnings data should contain reasonably accurate employer identity data.

Results in Table 6 indicate that STEP participants are about 7 percentage points (38 percent) more likely than non-participants to have continuous formal-sector employment in

[26]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 record why placements ended so we cannot directly measure successful transitions from that data source either.

every quarter of a given year during years 1 through 3 after enrollment. Further, they are about 6 percentage points (40 percent) more likely to be employed by the same employer in every quarter of a given year. Similarly, they are employed for roughly 0.44 more quarters on average in each year of the post-program period, and their longest match with a single employer lasts, on average, about 0.37 quarters longer. Trends in each of these outcomes are presented in Appendix Figure A-10 and show similar trends in employment stability prior to the program enrollment year on all four measures. Overall, we conclude that increased employer-employee match stability is likely a key mechanism driving the persistent increases in employment after STEP participants exited the program.

We also use 2-digit employer NAICS codes from the UI earnings data in the pre-enrollment and post-program periods to explore explore whether the STEP program moved participants into sectors they would not have otherwise worked in. Table 7 ranks 2-digit industries by average job duration among the comparison group (column 1) and reports the proportion of the comparison group ever employed in that industry (column 2). We categorize industries by average job duration as high-stability (more than 4 quarters), medium-stability (3-4 quarters), and low-stability (2-3 quarters). We then test whether the STEP program moved enrollees into differentially stable industries.[27]

Table 8 estimates the quarterly stacked difference-in-differences specification in Equation 2, excluding quarters 0-3 after enrollment, with employment in a given industry stability category in each quarter as the outcome.[28] The program made STEP enrollees about 3 percentage points (160 percent) more likely to work in the highest-stability industries in the post-program period—the public administration, education, and utilities sectors. STEP participation also leads to a roughly 5 percentage point (37 percent) increase in the likelihood of working in medium-stability industries, of which the most common is healthcare and social

[27]We are unable to report estimates for some of the individual industries in the sample due to the terms of the data sharing agreement. Any cell containing fewer than 30 individuals cannot be reported, which limits our ability to report estimates of employment rates in individual industries for the treatment group.

[28]See Appendix Figure A-11 for evidence of parallel trends in employment within each industry group prior to STEP enrollment.

assistance. Program effects on working in low-stability industries like retail, accommodation and food services, or admin support (which includes temp jobs) are smaller—about 2 percentage points (6 percent)—and only marginally statistically significant. These results show that the program differentially moves STEP participants into more stable sectors, although employment effects are positive across all industry stability categories because participants are more likely to be employed overall. We note that this measure does not capture any changes in within-industry job stability, which may also have improved, and that we cannot measure any program effects on the occupation distribution, either within or across industries.

It is worth noting that much of the movement into more stable sectors appears to be driven by the public and nonprofit sectors, which is a departure from most prior subsidized employment programs. Typically, programs seeking potential conversion to permanent employment focused exclusively on private sector positions, while most subsidized employment in the public and nonprofit sectors was explicitly temporary in nature and did not involve individualized placement efforts to find the "right match" for a specific participant (Anderson et al., 2019, Barham, Cadena and Turner, 2023, Bloom, 2010). Further, the finding that the STEP program produced lasting earnings gains in part by moving participants across sectors into higher-quality jobs is consistent with recent evidence on the mechanisms behind the success of sectoral job training programs. Katz et al. (2022) find that the long-term earnings gains from one such program were driven by movement into higher-earning sectors rather than by persistent increases in employment rates, and Narain and Noray (2024) find that the the majority of the variation in earnings gains across participants in another such program is explained by the quality of the initial job placement.

Recall from the main results in Figure 2 and Table 4 that average earnings effects remained stable after program exit while employment effects partially faded out, implying continued increases in earnings after the program period among those who remained employed. This pattern, along with the employment stability and sectoral switching results in

Tables 6 and 7, supports the interpretation that both increased employment through the formation of productive employer-employee matches and increases in job quality play a role in explaining the lasting effects of the STEP program on participants' earnings.

5.3 Effects on TANF and SNAP Receipt

In addition to affecting employment and earnings, STEP may have important effects on participants' ability and willingness to access TANF or SNAP benefits and/or their decision to exit the TANF caseload. Figure 5 shows that STEP participants had trends in TANF receipt similar to the comparison group's prior to enrolling in the program but slightly lower rates of SNAP receipt. As such, we report only aggregated difference-in-differences estimates for these outcomes to net out the small pre-existing differences between the two groups.

Results in Table 9 show that STEP participation increases participants' likelihood of TANF receipt (column 1) by about 13 percentage points (51 percent) and SNAP receipt by about 3 percentage points (5 percent) in the in-program period (months 0–9 relative to STEP enrollment). Recall that all STEP participants must remain on TANF throughout their time in the program, and all TANF recipients are categorically eligible for SNAP. Trends in Figure 5 indicate that this effect reflects disenrollment from both programs among the comparison group during this period. There are no lasting effects on TANF participation after program exit (months 10–48), as both the treatment and comparison group have low rates of TANF receipt in this period. The program may decrease participants' SNAP participation by about 2 percentage points in the post-program period, but this effect is only marginally statistically significant ($p < 0.1$) and may be sensitive to differences in pre-trends. Columns 3–5 present estimates of the effect of the STEP program on the monthly benefit amount received by the participant's household for each program and for both programs combined. Because the program increased participants' employment and earnings and Colorado has a very generous earned income disregard for TANF recipients (100% for new earned income sources and 67% for existing earned income sources among incumbent TANF recipients), SNAP benefit

amounts appear to be more sensitive to changes in earned income.

6 Cost-Effectiveness of STEP

We next perform a simple, back-of-the-envelope cost-benefit analysis 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 subsidized employment programs.

Table 10 first reports the net present value of STEP enrollees' earnings gains, discounted at a rate of 3% per year, as a result of their participation in the program. Our main estimate of $13,765 in additional earnings per participant over a four-year time horizon uses the main in-program and post-program earnings estimates reported in Column 1 of Table 4. A more conservative estimate in column 2 includes only formal-sector earnings with no imputation for placements that do not appear in the UI earnings data.

We next report the net present value of per-person program expenditures relative to participants' enrollment dates, with our main estimate in column 1 splitting program expenditures over quarters 0 and 1, discounting at 3% per year, and including all program participants who likely received full services in the denominator; we drop the 20 people who enrolled during the pandemic, as they likely received fewer services and experienced fewer program benefits. Our conservative estimate in column 2 assigns all program costs to quarter 0 with no discounting and includes only the main estimation sample in the denominator, allowing for the possibility that no earnings gains accrued to program participants who were dropped from the sample due to data quality issues.

Lastly, we report estimates of net discounted effects on TANF and SNAP benefits paid to participants throughout the four-year follow-up window. The main estimate in column 1 includes both the in-program period (in which participants' TANF receipt increased) and the post-program period (in which participants' SNAP receipt decreased). The conservative

estimate includes only the in-program period in case the post-program reduction in SNAP receipt is driven in part by differential pre-trends.

These results imply that participants' net earnings gains exceeded program-related expenditures by approximately $6,500 per participant over our four-year follow-up period according to our main estimates (column 1) or $3,600 according to our conservative estimates (column 2).

We can also think about the benefits to program participants relative to the program costs as a ratio, in the spirit of the marginal value of public funds (MVPF) framework. 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. Using directly estimated earnings gains alone as our measure of participants' willingness to pay would imply an MVPF of approximately 1.9 according to the main estimates (column 1) or 1.41 according to the conservative estimates (column 2). A full MVPF calculation would likely result in even higher estimates, as it would project participants' earnings over the life cycle. Even our current estimates imply that STEP is relatively cost-effective compared to most adult transitional jobs programs, because most do not produce substantial lasting earnings effects after the subsidies have expired (Hendren and Sprung-Keyser, 2020). For example, ReHire Colorado, which is programmatically very similar but serves a different population of low-income job seekers in Colorado, has an MVPF of 0.32 based on 16 quarters of directly estimated earnings effects and a 3% discount rate (Barham, Cadena and Turner, 2023).

7 Conclusion

This paper quantifies the effects of a supported work program available to TANF recipients on labor market outcomes and benefit 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 (10 percentage points) and quarterly earnings ($861) during the fourth through sixteenth quarter post-enrollment. Importantly, we also find improvements in job stability, with STEP participants being about 40 percent more likely to remain with the same employer for a full year in the post-program period. 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. The short-term increase in TANF receipt is offset by modest decreases in average SNAP payments both while participants are in the program and after they exit.

STEP 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 Barham, Cadena and Turner (2023). The key differences between STEP and ReHire are 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—and the fact that STEP is bundled with the TANF program, which provides time-limited cash assistance to participants who are required to meet work requirements. The difference therefore suggests 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) bundling the employment subsidy with an unconditional cash transfer may create interaction effects.

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 or increase overall income for those they served, at least in the immediate term over which program impacts were typically measured.

Overall, therefore, our findings suggest that providing job placement into subsidized positions that can become permanent is a particularly effective intervention for TANF recipients. Participants immediately increased their income after a negative earnings shock by nearly $1000 per quarter while participating in the program and by about $840 per quarter in the three years after program exit when considering combined effects on earnings and benefit receipt. Programs like STEP thus offer a promising pathway for policymakers looking to improve both the short-term well-being and longer-term self-sufficiency of low-income families.

8 Tables and Figures

Table 1: Descriptive Statistics, STEP Program Events

	Mean	Std. Dev.	Min.	Max	N
Placed	0.65	0.48	0	1	1072
Month in TANF Spell at Enrollment	6.96	8.23	1	59	1072
Days from Enrollment to Placement (If Placed)	21.04	39.05	0	555	692
Placement Length in Days (If Placed)	84.24	51.61	0	429	692
Days from Enrollment to Exit	136.25	107.58	0	771	1031

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 2: Sample Characteristics

	Mean (Trt.)	SD (Trt.)	Mean (Comp.)	SD (Comp.)	Diff. in Means	t (Diff. in Means)
Total Earnings, Year -4	7313.54	8679.87	6296.78	7952.84	1016.76	3.75
Total Earnings, Year -3	7738.23	9827.61	6629.22	9108.66	1109.02	3.61
Total Earnings, Year -2	7762.11	9560.90	6497.46	8842.08	1264.65	4.25
Total Earnings, Year -1	6226.65	7507.89	5081.71	7062.22	1144.95	4.91
Monthly SNAP Benefit Amount	447.45	172.57	444.47	190.99	2.97	0.55
Monthly TANF Payment Amount	481.29	131.93	461.86	156.69	19.43	4.68
Cumulative TANF Spells	2.01	1.31	2.19	1.54	-0.18	-4.39
Cumulative TANF Months	13.63	12.35	15.84	13.29	-2.21	-5.65
Female	0.86	0.34	0.83	0.37	0.03	2.98
Single Parent	0.85	0.36	0.81	0.39	0.03	3.06
Infant in HH	0.13	0.34	0.17	0.38	-0.04	-3.43
Child Under 5 in HH	0.62	0.49	0.63	0.48	-0.02	-1.07
Number of Children in HH	1.87	1.01	1.86	1.12	0.01	0.18
Age	33.01	8.06	31.64	8.13	1.37	5.40
Age of Youngest Child in HH	5.19	4.79	4.63	4.56	0.57	3.72
Age of Oldest Child in HH	8.20	5.16	7.73	5.14	0.47	2.88
Single Parent of Infant	0.10	0.30	0.12	0.33	-0.02	-2.34
Single Parent of Child Under 5	0.51	0.50	0.49	0.50	0.02	1.14
Employment Barrier in Current Spell	0.60	0.49	0.66	0.47	-0.06	-3.80
Sanction in Current Spell	0.09	0.28	0.18	0.38	-0.09	-10.09
Re-Engagement in Current Spell	0.24	0.43	0.34	0.47	-0.10	-7.78
Exemption in Current Spell	0.06	0.23	0.11	0.31	-0.05	-7.24
Less than High School Equivalent	0.08	0.26	0.14	0.35	-0.06	-7.34
GED	0.08	0.27	0.08	0.28	0.00	-0.39
High School Diploma	0.32	0.47	0.31	0.46	0.01	0.85

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 **??**.

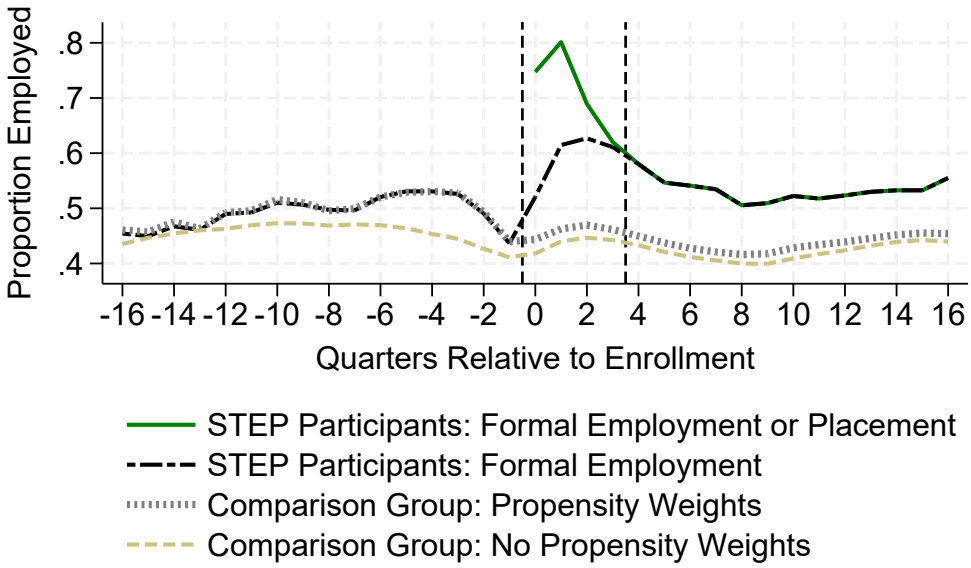
Table 3: Sample Characteristics, Reweighted Comparison Group

	(1) Mean (Trt.)	(2) SD (Trt.)	(3) Mean (Comp.)	(4) SD (Comp.)	(5) Diff. in Means	(6) t (Diff. in Means)
Total Earnings, Year -4	7313.54	8679.87	7367.89	8665.25	-54.35	-0.19
Total Earnings, Year -3	7738.23	9827.61	7757.17	9880.69	-18.94	-0.06
Total Earnings, Year -2	7762.11	9560.90	7734.87	9625.88	27.24	0.09
Total Earnings, Year -1	6226.65	7507.89	6228.57	7630.64	-1.92	-0.01
Monthly SNAP Benefit Amount	447.45	172.57	452.08	188.39	-4.63	-0.82
Monthly TANF Payment Amount	481.29	131.93	485.32	145.15	-4.03	-0.93
Cumulative TANF Spells	2.01	1.31	2.02	1.39	-0.01	-0.22
Cumulative TANF Months	13.63	12.35	13.84	12.53	-0.21	-0.53
Female	0.86	0.34	0.86	0.34	0.00	0.06
Single Parent	0.85	0.36	0.85	0.36	0.00	-0.03
Infant in HH	0.13	0.34	0.14	0.35	0.00	-0.46
Child Under 5 in HH	0.62	0.49	0.63	0.48	-0.01	-0.89
Number of Children in HH	1.87	1.01	1.91	1.03	-0.04	-1.22
Age	33.01	8.06	33.04	8.40	-0.03	-0.12
Age of Youngest Child in HH	5.19	4.79	5.18	4.75	0.02	0.11
Age of Oldest Child in HH	8.20	5.16	8.19	5.23	0.01	0.04
Single Parent of Infant	0.10	0.30	0.10	0.31	0.00	-0.40
Single Parent of Child Under 5	0.51	0.50	0.52	0.50	-0.01	-0.64
Employment Barrier in Current Spell	0.60	0.49	0.60	0.49	0.00	-0.24
Sanction in Current Spell	0.09	0.28	0.09	0.28	0.00	-0.01
Re-Engagement in Current Spell	0.24	0.43	0.24	0.43	-0.01	-0.57
Exemption in Current Spell	0.06	0.23	0.06	0.23	0.00	-0.14
Less than High School Equivalent	0.08	0.26	0.08	0.26	0.00	0.00
GED	0.08	0.27	0.08	0.27	0.00	0.14
High School Diploma	0.32	0.47	0.33	0.47	0.00	-0.19

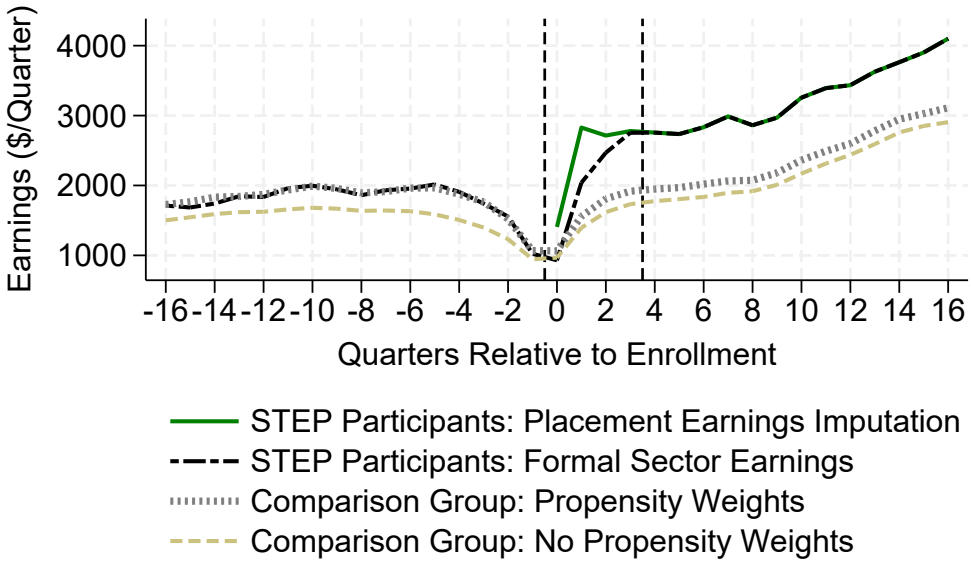
Note: Data come from administrative UI earnings data from CDLE, administrative TANF data from CDHS, and fields collected by the TANF caseworker at TANF application. The treated sample includes STEP participants who enrolled in the program between January 2018 and March 2020, and the comparison sample,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

Figure 1: Quarterly Employment and Earnings Trends

(a) Employment Rate



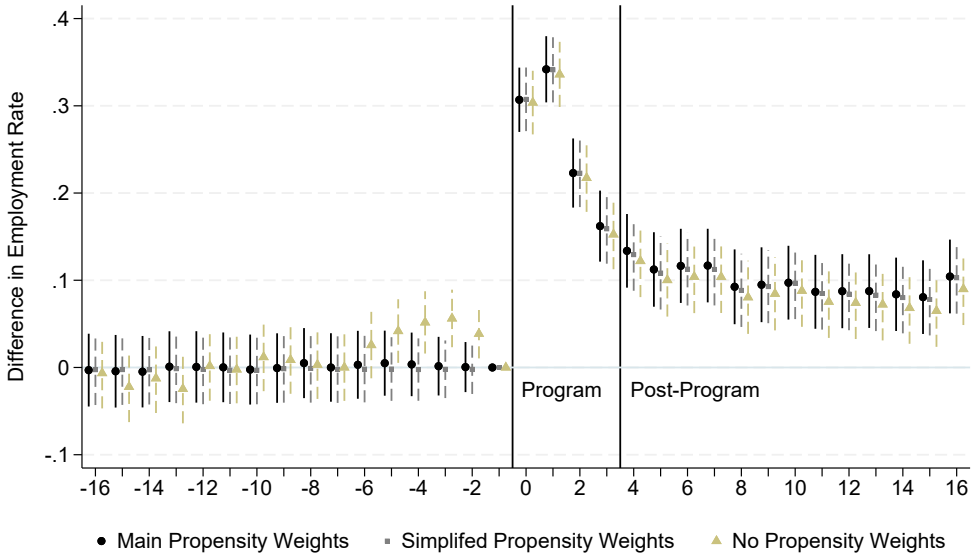
(b) Quarterly Earnings



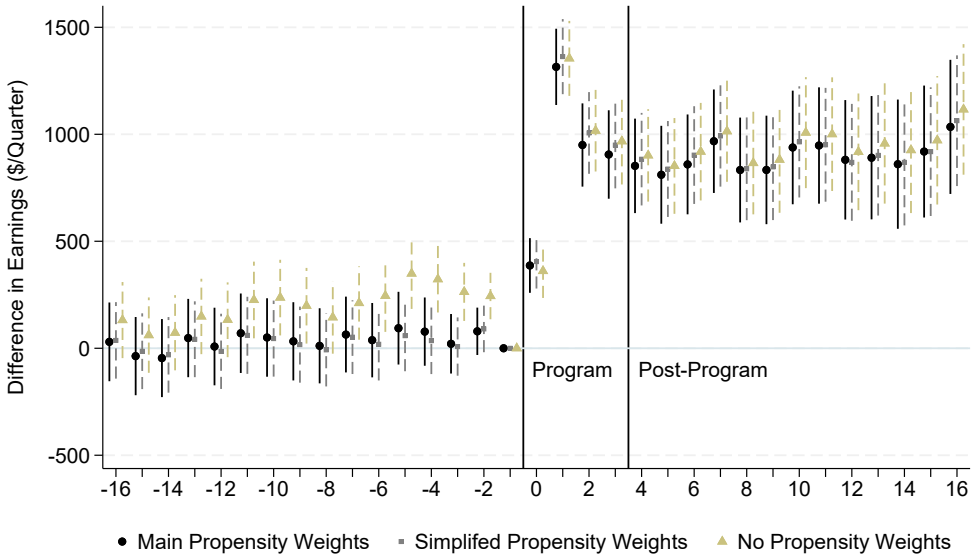
Notes: Data come from administrative UI earnings data from the CUBS data system and administrative TANF data from CDHS. Each quarterly sample includes 1072 STEP participants who applied between January 2018 and March 2020. Quarter or Month 0 represents the quarter or month in which an individual enrolled 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.

Figure 2: Event-Study Effects on Employment and Earnings

(a) Employment Rate



(b) Quarterly Earnings



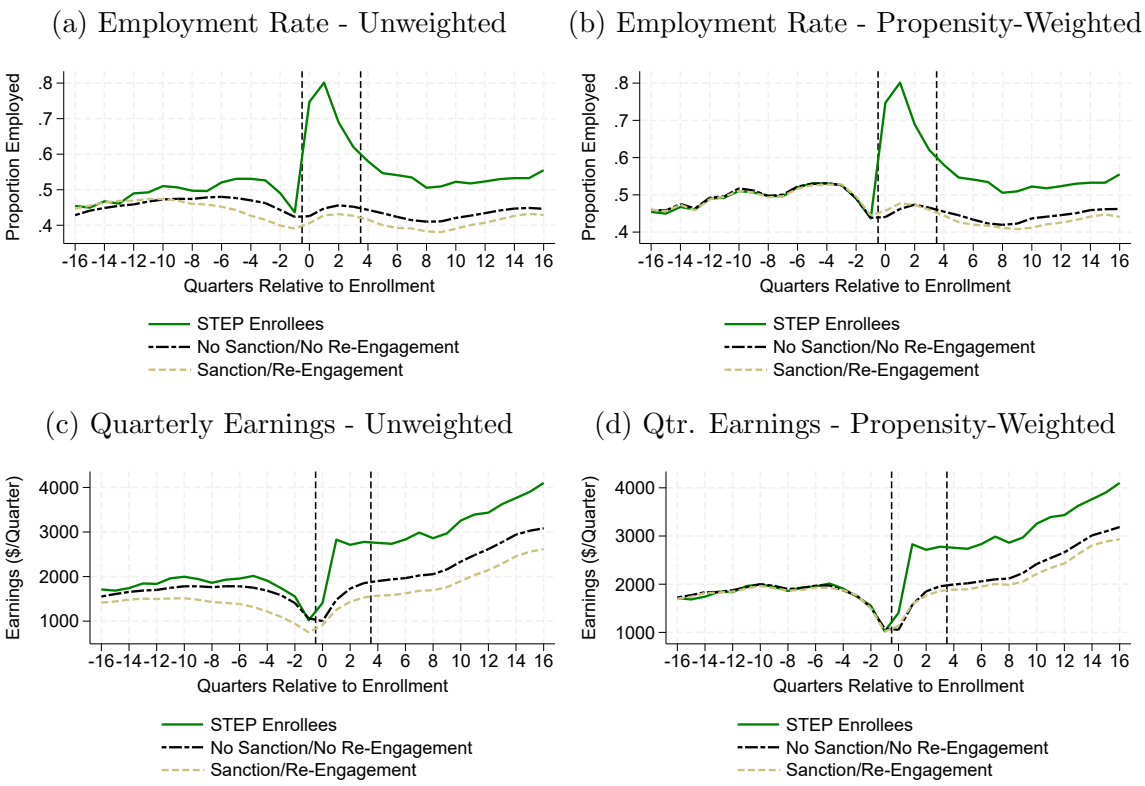
Notes: Data source is administrative UI earnings data from CDLE. Panel A and B report estimates on employment and earnings outcomes, respectively, for STEP participants who enrolled between January 2018 and March 2020. Quarter 0 represents the quarter in which an individual enrolled in STEP and is thus 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.

Table 4: Aggregated Effects on Employment and Earnings

	(1)	(2)	(3)
Panel A: Employment			
In-Program (Q0-Q3)	0.258***	0.260***	0.242***
	(0.0132)	(0.0127)	(0.0127)
Post-Program (Q4-Q16)	0.0993***	0.0985***	0.0760***
	(0.0138)	(0.0133)	(0.0133)
Comparison Mean (Q < 0)	0.494	0.492	0.455
Comparison Mean (Q0-Q3)	0.459	0.455	0.437
Comparison Mean (Q4-Q16)	0.437	0.435	0.421
Panel B: Quarterly Earnings			
In-Program (Q0-Q3)	855.6***	906.2***	738.2***
	(74.90)	(72.57)	(72.16)
Post-Program (Q4-Q16)	860.7***	886.0***	761.8***
	(102.6)	(99.52)	(99.25)
Comparison Mean (Q < 0)	1805.6	1807.1	1529.0
Comparison Mean (Q0-Q3)	1587.4	1538.3	1428.1
Comparison Mean (Q4-Q16)	2429.0	2405.2	2251.2
N (Individuals)	22613	27978	27978
Propensity Weights	Main	Simple	None

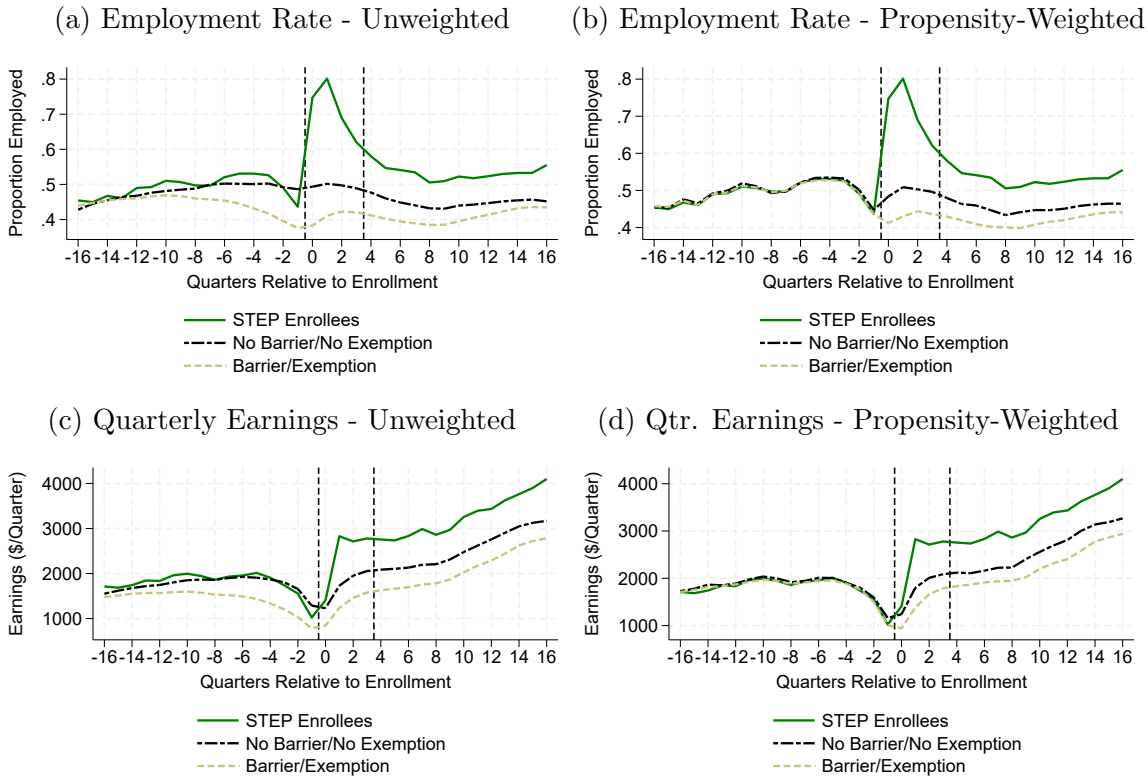
Notes: Data source is administrative UI earnings data from CDLE. Columns 1-3 and 4-6 report estimates on employment and earnings outcomes, respectively, for STEP participants who enrolled between January 2018 and March 2020. Quarter 0 represents the quarter in which an individual enrolled in STEP and 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. 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. Standard errors are clustered at the person level (*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$).

Figure 3: 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.

Figure 4: 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.

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)	0.257***	0.238***	0.257***	0.208***	0.221***
	(0.0132)	(0.0128)	(0.0135)	(0.0131)	(0.0150)
Post-Program (Q4-Q16)	0.0990***	0.0706***	0.0934***	0.0730***	0.0825***
	(0.0138)	(0.0134)	(0.0142)	(0.0137)	(0.0155)
Comparison Mean (Mth. < 0)	0.494	0.460	0.494	0.480	0.496
Comparison Mean (Mth. 0-9)	0.460	0.445	0.460	0.495	0.498
Comparison Mean (Mth. 10-48)	0.437	0.431	0.443	0.449	0.455
N (Individuals)	22613	24853	19161	17285	12878
Comparison Group	Full	No Sanc./No Re-eng.	No Sanc./No Re-eng.	No Barrier/No Exemption	No Barrier/No Exemption
Propensity Weights	Main	None	Main	None	Main
Prop. Treat. w/ Sanc./Re-eng.		0.266	0.266		
Prop. Treat. w/ Barrier/Exemption				0.614	0.614
Panel B: Quarterly Earnings					
In-Program (Q0-Q3)	850.1***	766.6***	837.4***	648.7***	694.8***
	(74.94)	(72.86)	(76.29)	(74.37)	(83.37)
Post-Program (Q4-Q16)	861.9***	721.8***	809.9***	687.7***	717.6***
	(102.7)	(100.2)	(104.8)	(102.3)	(114.6)
Comparison Mean (Mth. < 0)	1805.1	1645.7	1809.0	1753.2	1842.0
Comparison Mean (Mth. 0-9)	1592.4	1516.5	1609.0	1741.9	1784.7
Comparison Mean (Mth. 10-48)	2427.3	2407.9	2483.1	2549.6	2608.5
N (Individuals)	22613	24853	19161	17285	12878
Comparison Group	Full	No Sanc./No Re-eng.	No Sanc./No Re-eng.	No Barrier/No Exemption	No Barrier/No Exemption
Propensity Weights	Main	None	Main	None	Main
Prop. Treat. w/ Sanc./Re-eng.		0.266	0.266		
Prop. Treat. w/ Barrier/Exemption				0.614	0.614

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 weighting schemes 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 whether the propensity weights are used. Standard errors are clustered at the person level (*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$).

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

	(1) Employed Full Year	(2) Same Employer Full Year	(3) Qtrs. Employed in Year	(4) Qtrs. Same Employer
Post-Program (Years 1-3)	0.0698***	0.0639***	0.440***	0.367***
	(0.0116)	(0.0116)	(0.0746)	(0.0499)
Comparison Mean (Years -3 to -1)	0.184	0.161	2.487	1.649
Comparison Mean (Years 1-3)	0.180	0.178	2.131	1.514
N (Individuals)	22613	22613	22613	22613

Notes: Data source is administrative UI earnings data from CDLE. All results come from person-year stacked difference-in-differences regressions excluding the year of program enrollment and using the main set of stacked propensity weights. Standard errors are clustered at the person level (*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$).

Table 7: Job Duration by Industry, Non-Enrollees

Industry (2-Digit NAICS)	(1) Average Job Duration (Quarters)	(2) Proportion Ever Employed in Industry
Panel A: High-Stability Industries		
Utilities (22)	4.9	0.001
Public Administration (92)	4.8	0.040
Educational Services (61)	4.7	0.084
Panel B: Medium-Stability Industries		
Information (51)	3.6	0.054
Manufacturing (31-33)	3.5	0.149
Finance and Insurance (52)	3.4	0.106
Health Care and Social Assistance (62)	3.2	0.402
Management of Companies and Enterprises (55)	3.2	0.029
Mining (21)	3.1	0.010
Panel C: Low-Stability Industries		
Wholesale Trade (42)	2.9	0.136
Other Services (81)	2.8	0.159
Construction (23)	2.8	0.128
Professional Services (54)	2.7	0.152
Retail Trade (44-45)	2.6	0.545
Real Estate Rental and Leasing (53)	2.6	0.162
Arts, Entertainment, and Recreation (71)	2.5	0.073
Transportation and Warehousing (48-49)	2.5	0.186
Accommodation and Food Services (72)	2.3	0.543
Admin/Support/Waste Services (56)	2.2	0.605
Industry Code Missing	2.2	0.021
Agriculture/Forestry/Fishing/Hunting (11)	2.1	0.039

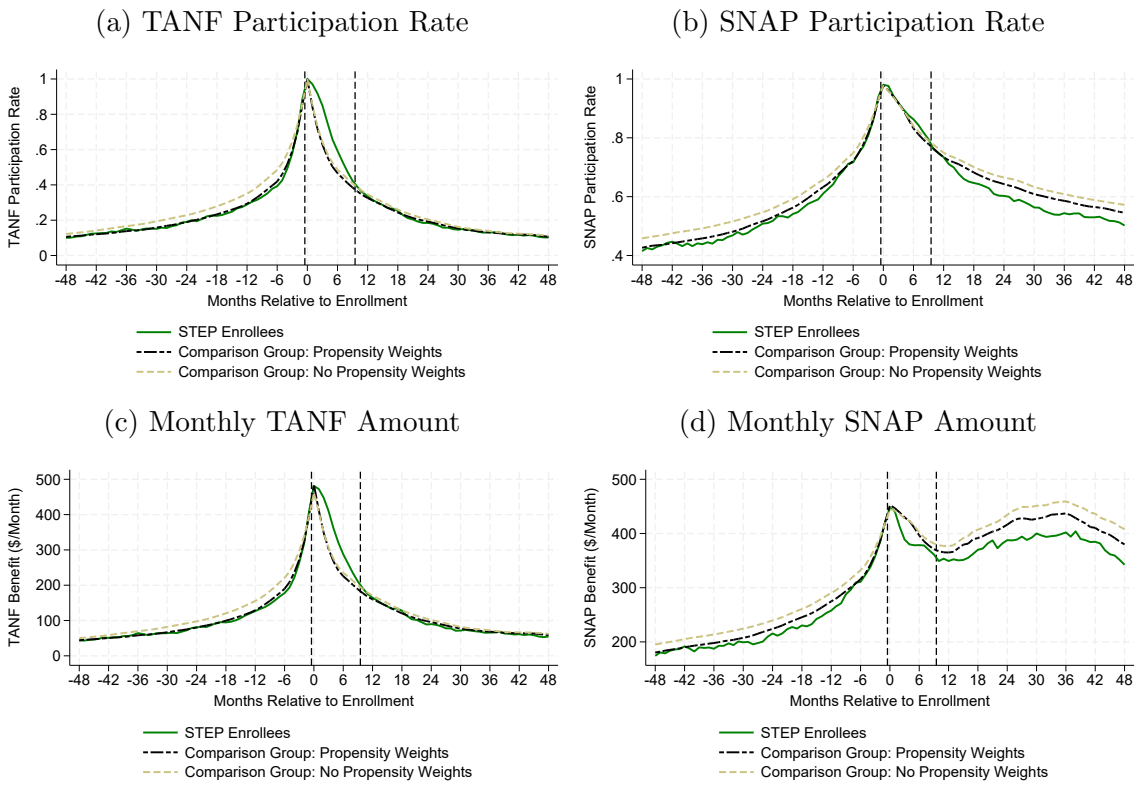
Notes: Data source is administrative UI earnings data from CDLE. Job durations and proportions of the TANF caseload employed represent only the comparison group.

Table 8: Sorting by Industry Stability

	(1) High-Stability Industries	(2) Medium-Stability Industries	(3) Low-Stability Industries
Post-Program (Q4-Q16)	0.0320***	0.0521***	0.0235+
	(0.00696)	(0.0110)	(0.0126)
Comparison Mean (Q < 0)	0.0196	0.142	0.362
Comparison Mean (Q4-Q16)	0.0238	0.148	0.292
N (Individuals)	22613	22613	22613
Average Job Duration	4-5 Quarters	3-4 Quarters	2-3 Quarters
NAICS Codes	22, 92, 61	51, 31-33, 52, 62, 55, 21	All Others

Notes: Data source is administrative UI earnings data from CDLE. All results come from person-quarter stacked difference-in-differences regressions excluding quarters 0-3 relative to potential STEP program enrollment and using the main set of stacked propensity weights. High-stability industries are defined as those with average job durations of 4-5 quarters, medium-stability industries are those with average job durations of 3-4 quarters, and low-stability industries are those with average job durations of 2-3 quarters. See Table 7 for details. Standard errors are clustered at the person level (*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$).

Figure 5: Trends in TANF and SNAP Receipt



Notes: Data source is administrative TANF and SNAP 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 and SNAP receipt is defined as having any 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.

Table 9: Aggregated Effects on TANF and SNAP Receipt

	(1) TANF Receipt	(2) SNAP Receipt	(3) TANF Amt.	(4) SNAP Amt.	(5) Combined Amt.
In-Program (Mths. 0-9)	0.126***	0.0265*	62.72***	-10.77+	51.95***
	(0.0109)	(0.0114)	(5.757)	(6.464)	(10.23)
Post-Program (Mths. 10-48)	0.00221	-0.0233+	0.450	-20.27*	-19.82
	(0.0108)	(0.0129)	(5.286)	(9.038)	(12.42)
Comparison Mean (Mth. < 0)	0.243	0.556	106.8	241.5	348.3
Comparison Mean (Mth. 0-9)	0.595	0.877	288.6	413.1	701.7
Comparison Mean (Mth. 10-48)	0.185	0.628	93.71	406.6	500.3
N (Individuals)	22613	22613	22613	22613	22613

Notes: Data source is administrative TANF and SNAP payments data from the CBMS data system. All results come from person-month stacked difference-in-differences regressions using the main set of stacked propensity weights. Standard errors are clustered at the person level (*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$).

Table 10: Cost-Effectiveness of the STEP Program

	(1) Main Estimate	(2) Conservative Estimate
Earnings Gains per Enrollee (Q0-Q16)	\$13764.81	\$12246.87
Direct Program Costs per Enrollee	\$7368.67	\$8148.23
Additional TANF/SNAP Benefits per Enrollee	\$-147.07	\$527.81
Discount Rate	3%	3%
Prog. Cost Denominator	1181	1072
Prog. Expenditure Timing	Q0-Q1	Q0
Earnings	Imputed	UI Earnings Only
Benefits Effects	M0-M48	M0-M9

Notes: 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. Main estimates of quarterly earnings effects are the coefficients from column 1 of Table 4 and conservative estimates are the coefficients from column 4 of Table A-3. Estimates of monthly benefits effects are the coefficients from column 5 of Table 9. The main estimate uses imputed earnings to calculate earnings gains where placements are not covered in UI earnings data, splits program expenditures across quarters 0 and 1, and removes only the 20 program participants dropped for enrolling during the pandemic from the program costs denominator. The conservative estimate assigns all program costs to quarter 0 (no discounting), does not impute earnings for placements not covered in the UI earnings data, sets post-program reductions in SNAP and TANF receipt to zero, and assigns all program costs to individuals in the main estimation sample.

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.

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.

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

Bobonis, Gustavo J, Aneta Bonikowska, Philip Oreopoulos, W Craig Riddell, and Steven P Ryan. 2022. "A helping hand goes a long way: Long-term effects of counselling and support to workfare program participants." National Bureau of Economic Research.

Callaway, Brantly, and Pedro HC Sant'Anna. 2021. "Difference-in-differences with multiple time periods." *Journal of Econometrics*, 225(2): 200–230.

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*.

De Chaisemartin, Clément, and Xavier d'Haultfoeuille. 2020. "Two-way fixed effects estimators with heterogeneous treatment effects." *American economic review*, 110(9): 2964–2996.

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.

Katz, Lawrence F, Jonathan Roth, Richard Hendra, and Kelsey Schaberg. 2022. "Why do sectoral employment programs work? Lessons from WorkAdvance." *Journal of Labor Economics*, 40(S1): S249–S291.

Kleven, Henrik. 2024. "The EITC and the extensive margin: A reappraisal." *Journal of Public Economics*, 236: 105135.

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

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.

Narain, Namrata, and Kadeem Noray. 2024. "Whose Bridge to Opportunity and Why? Unpacking the Impacts of Sectoral Job Training." *Three Essays on Talent Selection and Allocation in the Economy*, 108.

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.

Sun, Liyang, and Sarah Abraham. 2021. "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects." *Journal of econometrics*, 225(2): 175–199.

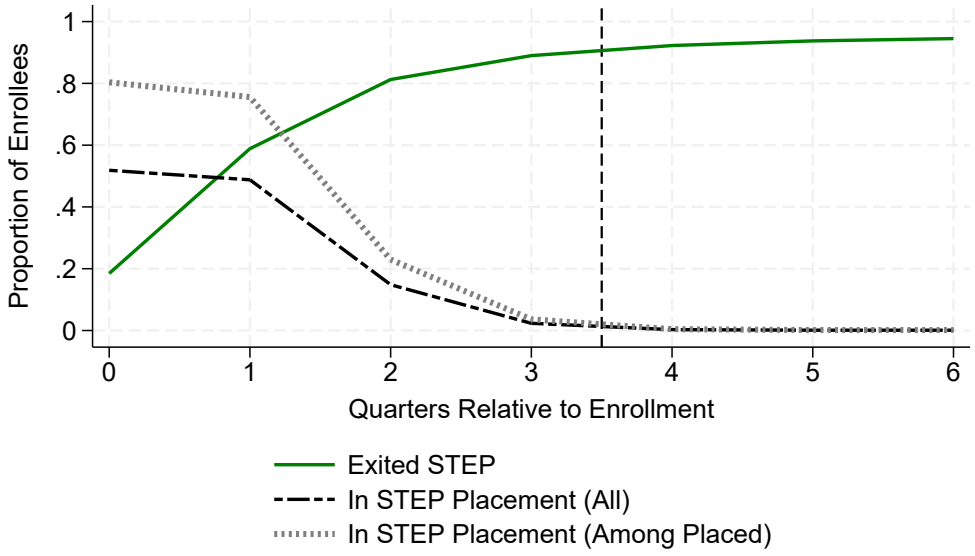
Wing, Coady, Seth M Freedman, and Alex Hollingsworth. 2024. "Stacked Difference-in-Differences." *NBER Working Paper No. 32054*.

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-1: STEP Placement and Exit Timing



Note: Shows proportion of treatment group that is in a STEP placement or has exited the program in each quarter relative to enrollment. Quarters 0–3 largely represent outcomes during individuals' enrollment in the program and quarters 4 and beyond largely represent outcomes after they have exited the program.

A.2 Sample Description

As discussed in Section 3.2, drop treatment group members who enrolled in STEP after the first quarter of 2020 (20/1,201) because their program experience was affected by the COVID-19 pandemic and the impending end of the program's initial funding period. We also remove treatment group observations who do not match to the CBMS data using social security number, name, and/or birthdate (10/1,201) or who matched to some CBMS history but did not receive a TANF payment in their enrollment month as the program requires (24/1,201) or who

Finally, we remove treatment group (34/1,201) and comparison group individuals from our analysis if their earnings records are 1) inconsistent with qualifying for TANF participation or 2) otherwise indicate a likely incorrect SSN match. First, we drop individuals who had person-level median quarterly earnings during quarters in which they received TANF that exceeded $10,000 per quarter, or the equivalent of $40,000 per year, and had fewer than 4 children according to their CBMS records. We assume that such earnings are employer data entry errors or bad identity matches, 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. 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. We further correct for outliers in quarterly earnings entries by winsorizing earnings at the 95th percentile in each quarter-by-possible-enrollment-month cell.

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

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

m	$N^m_{D=1}$	$N^m_{D=0}$m	$w^{D=0}_{im}$
1/2018	4	8062	0.10
2/2018	40	8054	0.96
3/2018	35	8042	0.84
4/2018	53	7798	1.31
5/2018	42	7982	1.02
6/2018	36	7843	0.89
7/2018	29	7941	0.71
8/2018	40	7839	0.99
9/2018	30	7810	0.74
10/2018	44	8001	1.06
11/2018	46	7934	1.12
12/2018	40	8003	0.97
1/2019	56	7793	1.39
2/2019	68	7604	1.73
3/2019	53	7499	1.37
4/2019	54	7437	1.40
5/2019	43	7742	1.07
6/2019	38	7625	0.96
7/2019	47	7778	1.17
8/2019	43	7781	1.07
9/2019	37	7450	0.96
10/2019	34	7492	0.88
11/2019	35	7162	0.94
12/2019	36	7171	0.97
1/2020	39	7086	1.06
2/2020	29	7147	0.78
3/2020	21	7195	0.56
	Treatment	Comparison	
N_i	1072	26906	
N_{im}	1072	207271	
		$\sum w^{D=0}_{im}$	207271

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.

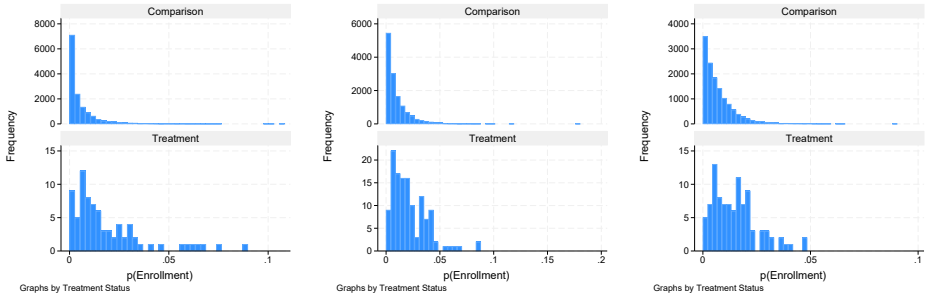
A.3 Propensity Score Weighting

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

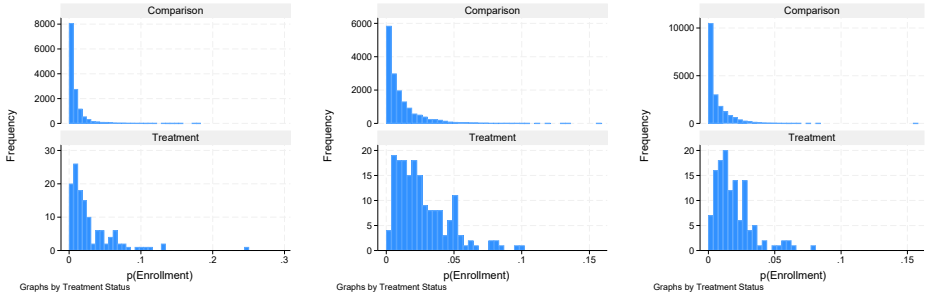
Figure A-2 demonstrates common support by showing the distribution of estimated treatment probabilities for each of nine program entry cohorts. Importantly, across all cohorts, there are a substantial number of comparison group members with estimated propensity scores similar to the treatment group's.

Figure A-2: Common Support - Enrollment Propensity by Quarter

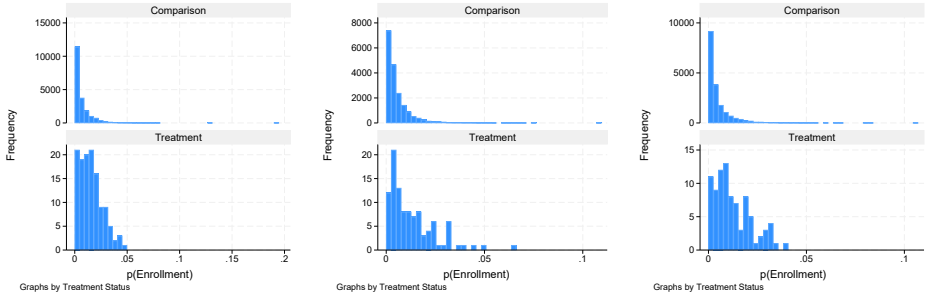
(a) Q1 2018 (b) Q2 2018 (c) Q3 2018



(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.

Table A-2 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-2: Probit Results

	2018Q1	2018Q2	2018Q3	2018Q4	2019Q1	2019Q2	2019Q3	2019Q4	2020Q1
	-0.000	0.000	0.000	-0.000	0.000	-0.000	0.000	0.000	0.000
Monthly SNAP Amount									
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Monthly TANF Amount	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)
Age	-0.000	0.000	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)	(0.000)	(0.000)	(0.000)
Number of Children in HH	-0.001	-0.003	-0.004**	-0.003*	-0.004*	-0.002	-0.002	0.000	-0.002*
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Infant in HH	-0.006	-0.009	0.002	0.008	0.001	-0.001	-0.007	-0.001	-0.002
	(0.005)	(0.006)	(0.004)	(0.005)	(0.005)	(0.003)	(0.004)	(0.004)	(0.003)
Child Under 5 in HH	-0.002	0.009	0.003	-0.002	-0.005	0.001	-0.001	0.003	0.003
	(0.004)	(0.007)	(0.006)	(0.006)	(0.005)	(0.004)	(0.003)	(0.005)	(0.004)
Age of Youngest Child in HH	0.000	0.001	-0.000	0.000	0.000	-0.000	0.000	0.000	-0.000

Continued on next page

Table A-2 – *Continued from previous page*

	2018Q1	2018Q2	2018Q3	2018Q4	2019Q1	2019Q2	2019Q3	2019Q4	2020Q1
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age of Oldest Child in HH	-0.000	-0.001	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)
Adams County	0.000	0.000	0.000	0.000	-0.039***	0.000	0.000	0.000	0.000
	(.)	(.)	(.)	(.)	(0.008)	(.)	(.)	(.)	(.)
Arapahoe County	-0.001	-0.003	-0.007*	-0.020***	-0.010**	0.001	-0.006*	-0.001	-0.004
	(0.002)	(0.004)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Denver County	0.000	0.000	0.000	0.000	0.000	-0.023***	-0.014***	-0.005*	-0.003
	(.)	(.)	(.)	(.)	(.)	(0.005)	(0.003)	(0.002)	(0.002)
El Paso County	-0.008**	-0.008*	-0.010***	-0.021***	-0.018***	-0.010***	-0.010***	-0.007**	-0.006**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Jefferson County	-0.000	-0.000	-0.005	-0.014***	-0.004	-0.002	-0.002	0.005*	0.004*
	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Larimer County	0.004	0.002	-0.005	-0.023***	0.000	-0.002	-0.001	0.001	0.003

Continued on next page

Table A-2 – *Continued from previous page*

	2018Q1	2018Q2	2018Q3	2018Q4	2019Q1	2019Q2	2019Q3	2019Q4	2020Q1
	(0.002)	(0.004)	(0.003)	(0.004)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Mesa County	0.002	0.007	-0.002	-0.009**	-0.006	-0.001	-0.004	0.001	0.002
	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.002)
Weld County	0.001	0.006	-0.004	-0.013***	-0.014**	0.000	-0.003	0.000	0.002
	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.002)	(0.003)	(0.003)	(0.002)
Cumulative TANF Spells	-0.001**	-0.001	-0.001	-0.000	-0.001	-0.000	-0.000	0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Cumulative TANF Months	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)
Less than High School Equivalent	-0.006*	-0.012***	-0.010**	-0.000	-0.009**	0.001	-0.005*	0.001	-0.001
	(0.003)	(0.004)	(0.004)	(0.003)	(0.004)	(0.002)	(0.002)	(0.002)	(0.002)
GED	-0.002	0.001	0.001	0.004	-0.005	0.001	-0.002	-0.001	-0.002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)

Continued on next page

Table A-2 – *Continued from previous page*

	2018Q1	2018Q2	2018Q3	2018Q4	2019Q1	2019Q2	2019Q3	2019Q4	2020Q1
High School Diploma	0.000	-0.002	-0.001	0.003	-0.003	-0.000	-0.000	0.002	0.000
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Employment Barrier in Current Spell	-0.001	-0.002	-0.001	-0.002	-0.006**	-0.002	-0.001	-0.001	-0.001
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Sanction in Current Spell	-0.000	-0.001	-0.005	-0.003	-0.003	-0.002	-0.004	-0.000	-0.005*
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Re-Engagement in Current Spell	-0.004*	-0.004*	-0.004	-0.004*	-0.006*	-0.003	-0.002	0.001	0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Exemption in Current Spell	-0.005	-0.007*	-0.009*	-0.006	-0.009*	-0.007*	-0.004	-0.000	0.002
	(0.002)	(0.004)	(0.004)	(0.003)	(0.004)	(0.003)	(0.002)	(0.002)	(0.002)
Female	0.007**	0.003	0.003	0.004	0.003	0.005**	0.002	0.003	-0.001
	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)

Continued on next page

Table A-2 – *Continued from previous page*

	2018Q1	2018Q2	2018Q3	2018Q4	2019Q1	2019Q2	2019Q3	2019Q4	2020Q1
Single Parent	-0.002	0.017**	0.010*	0.004	0.006	0.007	-0.000	0.004	0.004
	(0.003)	(0.006)	(0.005)	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)
Single Parent of Infant	0.006	0.006	-0.006	-0.011	-0.002	-0.002	0.005	-0.001	-0.000
	(0.005)	(0.007)	(0.005)	(0.006)	(0.006)	(0.004)	(0.005)	(0.004)	(0.003)
Single Parent of Child Under 5	0.003	-0.006	-0.002	0.007	0.006	-0.004	0.003	-0.002	-0.001
	(0.004)	(0.007)	(0.006)	(0.005)	(0.005)	(0.004)	(0.003)	(0.004)	(0.004)
Employed, Quarter -1	0.002	-0.002	-0.004*	0.003	-0.000	-0.001	-0.001	0.000	0.001
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)
Employed, Quarter -2	-0.002	0.002	0.001	-0.002	0.001	0.004*	0.003	-0.001	0.003*
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
Employed, Quarter -3	0.003	0.001	0.001	0.000	0.004	0.000	-0.000	0.005**	0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)
Employed, Quarter -4	0.000	0.001	0.004	0.005	-0.002	0.001	0.003	-0.001	-0.000
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)

Continued on next page

Table A-2 – *Continued from previous page*

	2018Q1	2018Q2	2018Q3	2018Q4	2019Q1	2019Q2	2019Q3	2019Q4	2020Q1
Employed, Quarter -5	0.002	-0.000	0.001	0.001	-0.001	0.001	0.003	0.000	-0.001
	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Employed, Quarter -6	0.001	0.008**	-0.001	0.004	0.001	-0.001	-0.003	-0.002	0.002
	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Employed, Quarter -7	-0.003	-0.002	-0.001	-0.006*	0.002	0.000	0.002	0.001	-0.002
	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Employed, Quarter -8	0.001	-0.001	0.000	-0.000	0.002	-0.001	-0.001	-0.002	-0.002
	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.001)	(0.002)
Employed, Quarter -9	-0.000	-0.000	0.001	-0.000	-0.003	0.000	-0.003	0.003	0.003
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)
Employed, Quarter -10	0.001	-0.003	-0.001	-0.001	-0.001	0.004*	0.002	0.001	0.001
	(0.002)	(0.003)	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Employed, Quarter -11	-0.002	0.004	0.003	0.002	0.001	-0.004	-0.000	-0.001	-0.003*
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)

Continued on next page

Table A-2 – *Continued from previous page*

	2018Q1	2018Q2	2018Q3	2018Q4	2019Q1	2019Q2	2019Q3	2019Q4	2020Q1
Employed, Quarter -12	0.000	0.002	-0.003	0.001	0.003	0.003	-0.002	0.001	0.002
	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Employed, Quarter -13	-0.000	-0.005	-0.005	0.002	-0.002	-0.004*	-0.003	-0.001	-0.001
	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)
Employed, Quarter -14	-0.004	-0.001	0.002	-0.001	0.000	0.002	0.003	0.004	-0.003
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Employed, Quarter -15	0.004	0.001	0.000	-0.004	0.001	-0.002	-0.003	-0.004*	-0.000
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Employed, Quarter -16	-0.004*	-0.000	-0.002	-0.001	0.000	0.001	0.002	0.002	0.002
	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Total Earnings, Year -4	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 Earnings, Year -3	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)

Continued on next page

Table A-2 – *Continued from previous page*

	2018Q1	2018Q2	2018Q3	2018Q4	2019Q1	2019Q2	2019Q3	2019Q4	2020Q1
Total Earnings, Year -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)
Total Earnings, Year -1	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)
Observations	13925	13456	13179	13736	15960	19431	19769	18611	18273

Notes: Marginal 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

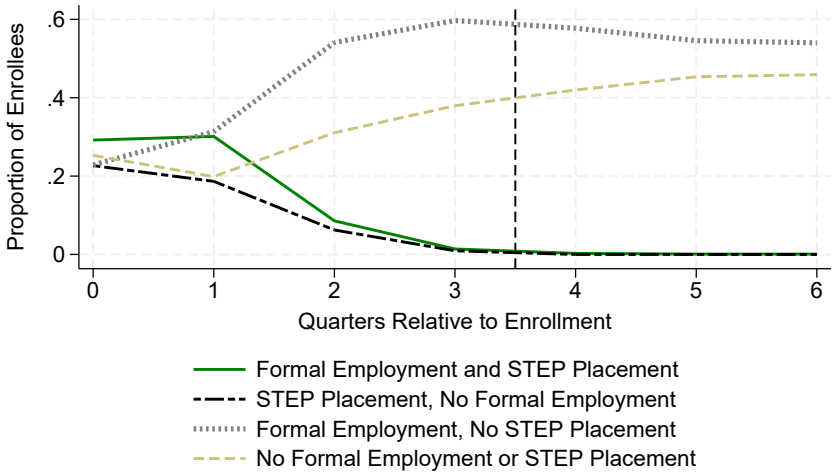
As discussed in Section 2, STEP program contractors had multiple options for how to pay participants during their placements. They could serve as the employer of record and pay the participant directly, use a staffing agency as the employer of record, or reimburse internship host site employers for subsidized wages paid to participants.

Figure A-3 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. Figure A-4 prevents event-study estimates with and without imputing employment and earnings for placements that did not appear to result in formal-sector earnings in the UI data. Note that this imputation affects only the in-program quarters and not the post-program quarters. Our preferred estimates include the imputation so as not to understate the fade-out of the program's effects on employment and earnings after participants exit the program.

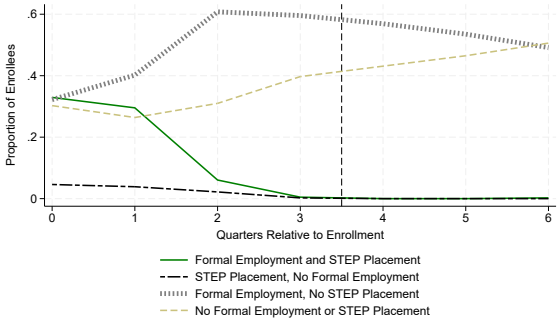
Each program contractor serves a distinct geographic area, and in some counties the program contractors appear to have paid participants in ways that typically generated UI earnings. We treat counties where at least 75 percent of participants have UI earnings in the quarter that contains the start date of their placement as "good coverage" counties. Panel b of Figure A-3 shows that, in these counties, the vast majority of placements are covered in the UI earnings data, and panel c shows that coverage is much more inconsistent in the other counties. Figure A-5 and Table A-4 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 and earnings effects are qualitatively similar across the two groups of counties, with small differences in magnitudes likely reflecting differences in 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-3: In-Program Employment Among STEP Enrollees

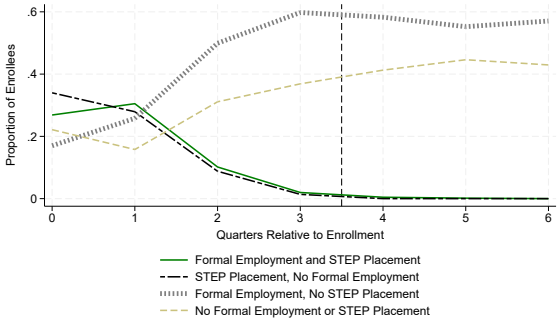
(a) All STEP Enrollees



(b) Good UI Placement Coverage Counties



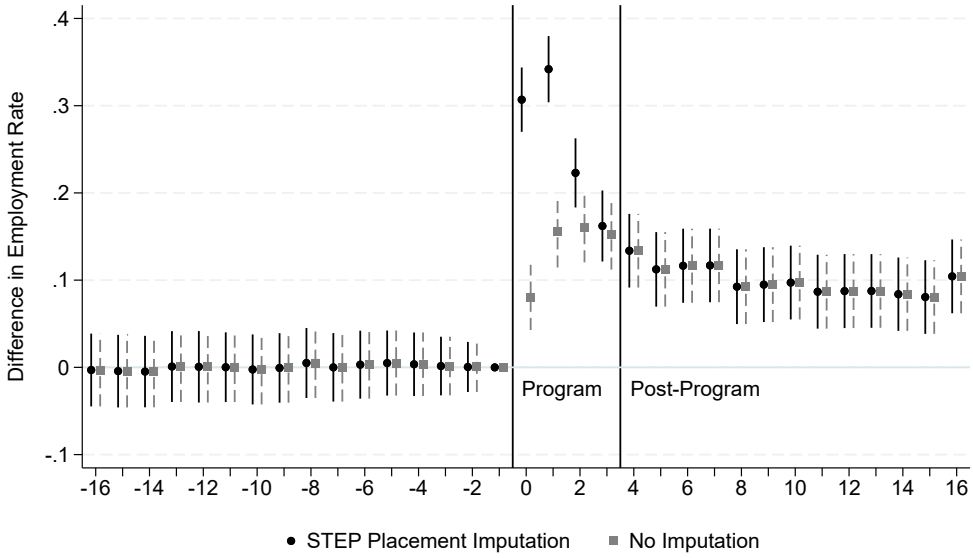
(c) Poor UI Placement Coverage Counties



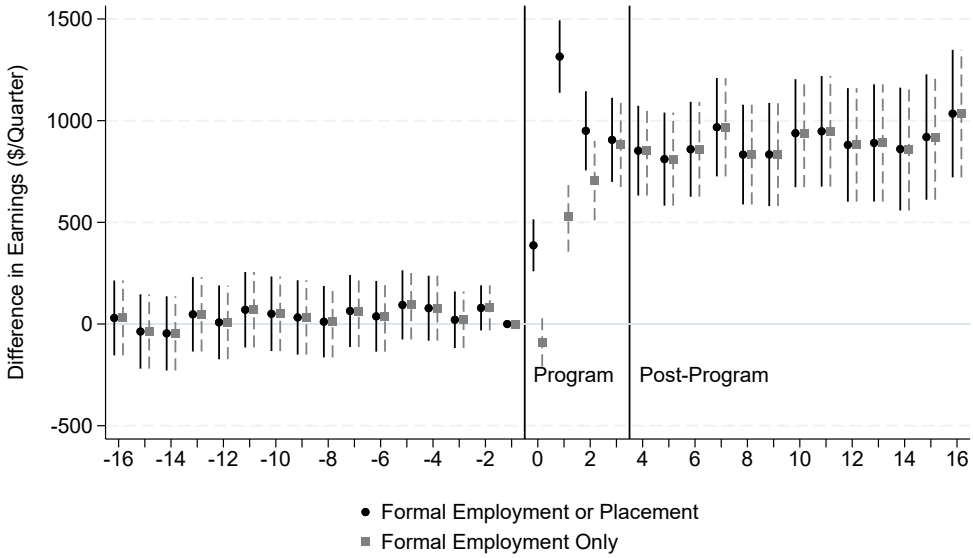
Notes: Data source is administrative UI earnings data and STEP program tracking data from CDLE. Panel a provides a decomposition of formal sector employment and placement rates in each quarter for all STEP enrollees who enrolled from January 2018 to March 2020. Panel b provides placement and formal-sector 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. Standard errors are clustered at the person level (*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$).

Figure A-4: Event-Study Effects on Employment and Earnings, With and Without Imputation

(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-3: Aggregated Effects on Employment and Earnings, With and Without Imputation

	(1) Employment	(2) Employment	(3) Earnings	(4) Earnings
In-Program (Q0-Q3)	0.258***	0.137***	855.6***	471.9***
	(0.0132)	(0.0139)	(74.90)	(75.23)
Post-Program (Q4-Q16)	0.0993***	0.0993***	860.7***	860.7***
	(0.0138)	(0.0138)	(102.6)	(102.6)
Comparison Mean (Q< 0)	0.494	0.494	1805.6	1805.6
Comparison Mean (Q0-Q3)	0.459	0.459	1587.4	1587.4
Comparison Mean (Q4-Q16)	0.437	0.437	2429.0	2429.0
N (Individuals)	22613	22613	22613	22613
Placement Imputation	Yes	No	Yes	No

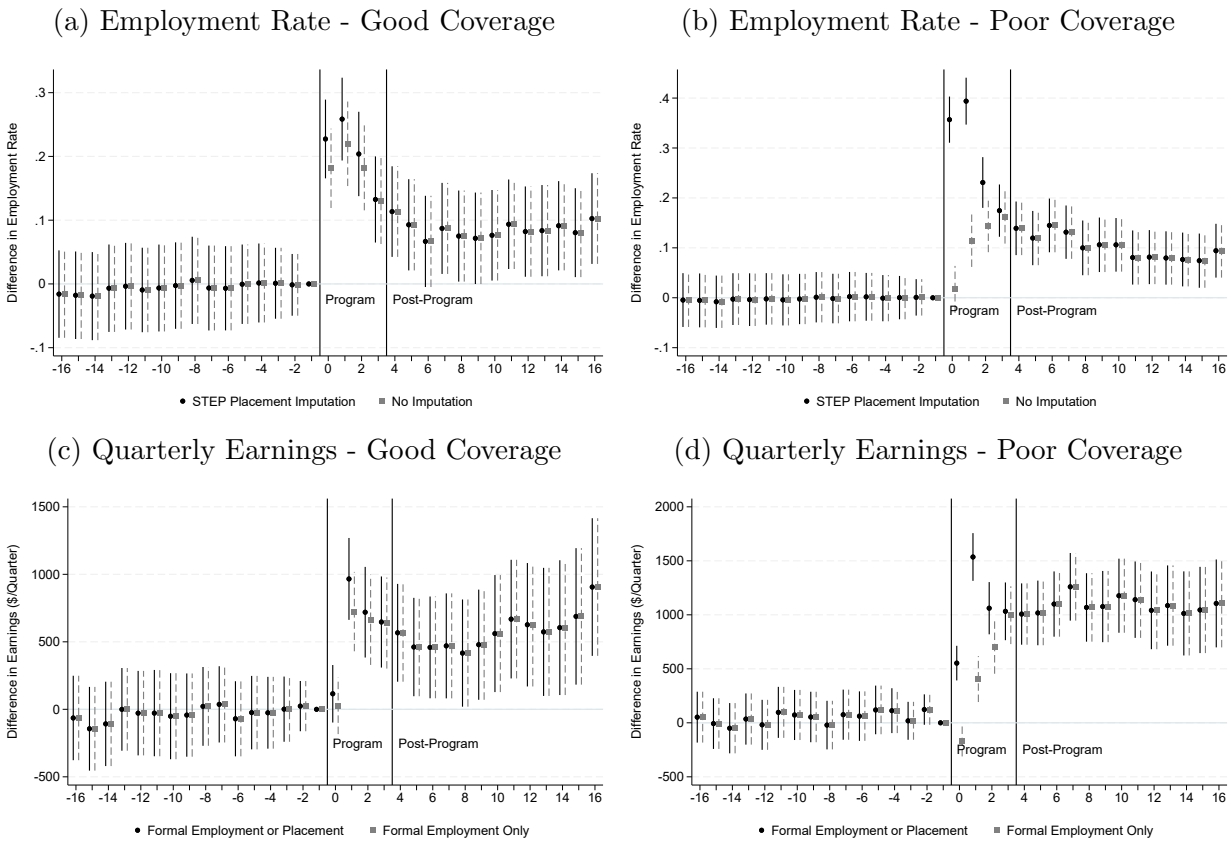
Notes: Data source is administrative UI earnings data from CDLE. Panel A and B report estimates on employment and earnings outcomes, respectively, for STEP participants who enrolled between January 2018 and March 2020. Quarter 0 represents the quarter in which an individual enrolled in STEP and is thus 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. Standard errors are clustered at the person level (*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$).

Table A-4: Aggregated Effects on Employment and Earnings, Split by Placement Earnings Coverage in UI Data

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Employment						
In-Program (Q0-Q3)	0.257***	0.136***	0.211***	0.184***	0.291***	0.111***
	(0.0132)	(0.0139)	(0.0244)	(0.0241)	(0.0162)	(0.0175)
Post-Program (Q4-Q16)	0.0990***	0.0990***	0.0914***	0.0914***	0.104***	0.104***
	(0.0138)	(0.0138)	(0.0230)	(0.0230)	(0.0182)	(0.0182)
Comparison Mean (Q< 0)	0.494	0.494	0.474	0.474	0.504	0.504
Comparison Mean (Q0-Q3)	0.460	0.460	0.474	0.474	0.444	0.444
Comparison Mean (Q4-Q16)	0.437	0.437	0.431	0.431	0.437	0.437
N (Individuals)	22613	22613	9058	9058	13869	13869
Sample	Main	Main	Good UI Coverage	Good UI Coverage	Poor UI Coverage	Poor UI Coverage
Imputation	Yes	No	Yes	No	Yes	No
Panel A: Employment						
In-Program (Q0-Q3)	850.1***	466.3***	643.3***	544.5***	1000.3***	438.0***
	(74.94)	(75.27)	(127.9)	(126.8)	(97.06)	(97.84)
Post-Program (Q4-Q16)	861.9***	861.9***	606.9***	606.9***	1042.1***	1042.1***
	(102.7)	(102.7)	(167.2)	(167.2)	(136.8)	(136.8)
Comparison Mean (Q< 0)	1805.1	1805.1	1728.9	1728.9	1842.4	1842.4
Comparison Mean (Q0-Q3)	1592.4	1592.4	1675.1	1675.1	1509.6	1509.6
Comparison Mean (Q4-Q16)	2427.3	2427.3	2429.8	2429.8	2394.9	2394.9
N (Individuals)	22613	22613	9058	9058	13869	13869
Sample	Main	Main	Good UI Coverage	Good UI Coverage	Poor UI Coverage	Poor UI Coverage
Imputation	Yes	No	Yes	No	Yes	No

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. Standard errors are clustered at the person level (*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$).

Figure A-5: 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

Figures A-6 and A-7 provide event study estimates based on the trends as shown in Figures 3 and 4 in the main text, respectively. The primary description of this analysis appears in the main text in section 5.1.2.

Figure A-8 and Table A-5 present earnings effects varying the threshold at which UI earnings data are winsorized to reduce the influence of outliers. In the main specification (column 1), we winsorize UI earnings data at the 95th percentile in each cohort-by-quarter cell. Columns 2 and 3 winsorize UI earnings at the 98th and 99th percentile, respectively. Table A-6 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 similar im magnitude 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. Figure A-9 shows that trends in the comparison group's employment rates are similar across counties with higher and lower proportions of the TANF caseload treated. These analyses also discussed in Section 5.1.2.

Table A-7 shows effect heterogeneity based on whether individuals had a sanction/re-engagement designation or a work barrier/work requirements exemption in the TANF spell of their (potential) program enrollment month. Individuals with barriers appear to benefit more from the program, with no substantial heterogeneity based on sanction/re-engagement status. Additionally, Table A-8 provides heterogeneity analysis of program effects based on the age of the participant's youngest child in the household and the participant's cumulative TANF months observed in the data at the time of enrollment. Program effects do not vary substantially along these dimensions. These analyses are discussed in section 5.1.3 in the main text.

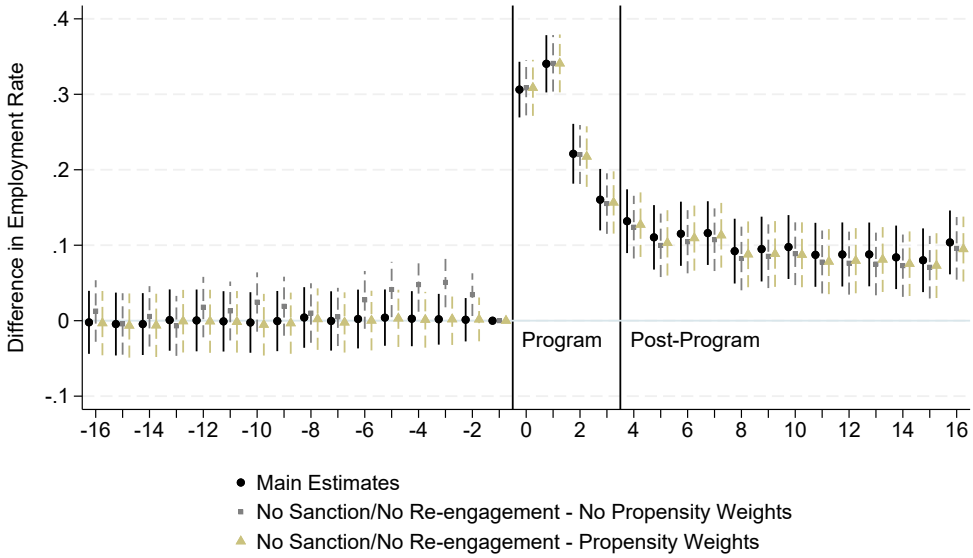
Figure A-10 presents trends before and after STEP enrollment in the person-year-level employment stability measures used in the analysis reported in Table 6. These trends indicate

that STEP enrollees and non-enrollees exhibited similar trends in both general employment stability and stability of employer-employee matches in the three years prior to enrollment in the program, and that the treatment group was more likely to maintain stable employment and form stable matches in the three years after participating in the program. The program year (quarters 0-3) is excluded from this analysis due to the fact that many STEP placements do not appear in UI earnings data at all and even fewer appear under the identity of the host site employer. Similarly, Figure A-11 presents trends before and after STEP enrollment in rates of employment in high-stability, medium-stability, and low-stability industries, as defined in Tables 7 and 8, showing parallel trends prior to enrollment and increases in STEP enrollees' employment in higher-stability industries after the program. These potential mechanisms for STEP's lasting employment and earnings effects are discussed in detail in Section 5.2.

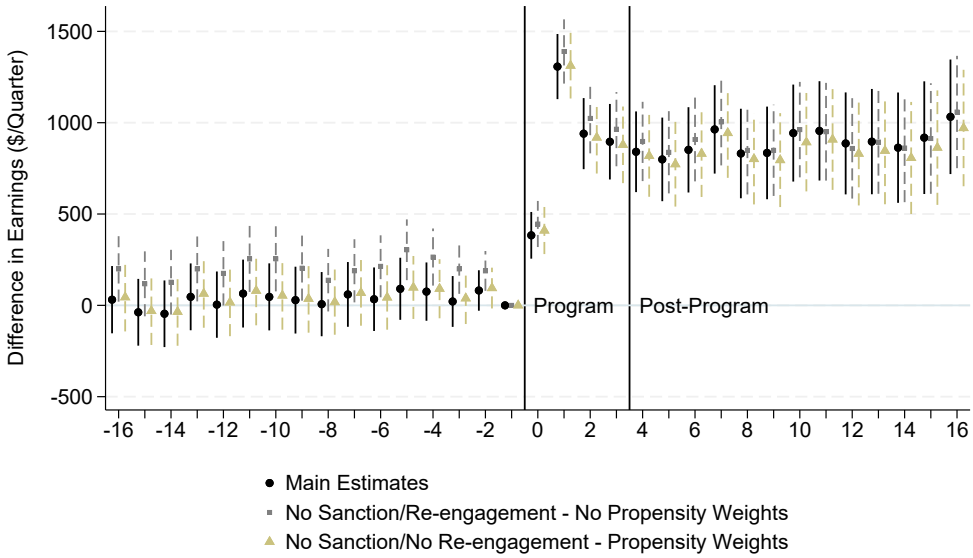
Finally, Figure A-12 shows additional trends in benefit receipt outcomes, corresponding to the analysis in Table 9 in Section 5.3.

Figure A-6: Event-Study Effects on Employment and Earnings vs. No Sanction/No Re-engagement Comparison Group

(a) Employment Rate



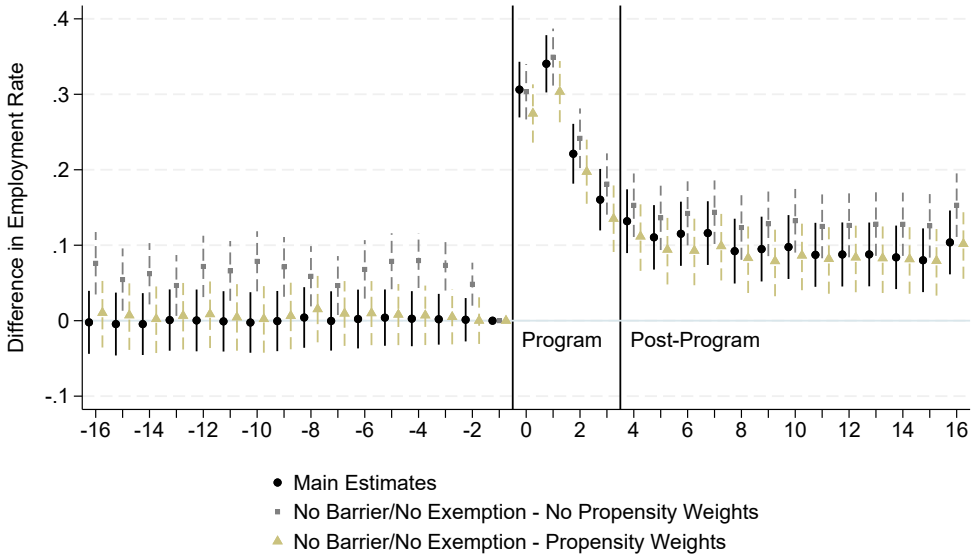
(b) Quarterly Earnings



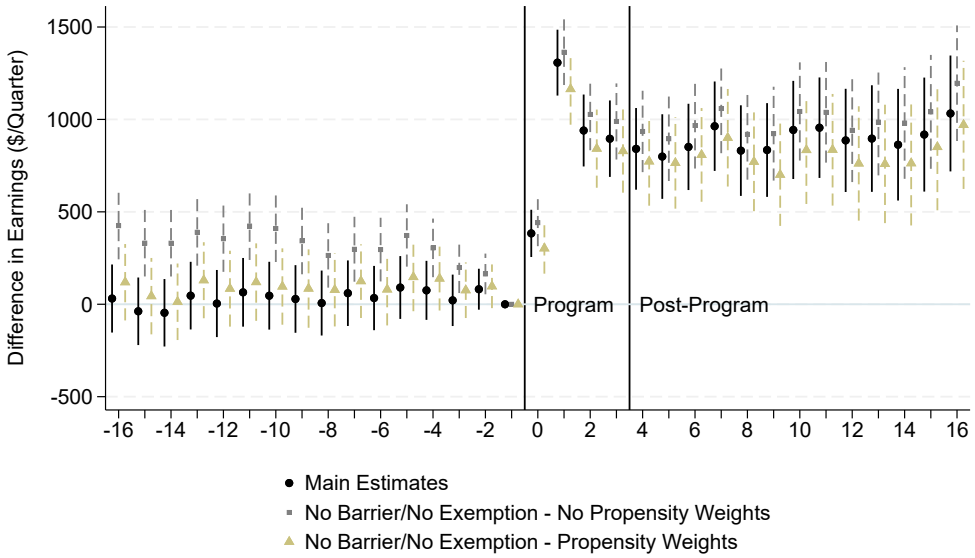
Notes: Data source is administrative UI earnings data from CDLE. Panel A and B report estimates on employment and earnings outcomes, respectively, for STEP participants who enrolled between January 2018 and March 2020. Quarter 0 represents the quarter in which an individual enrolled in STEP and is thus 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.

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

(a) Employment Rate

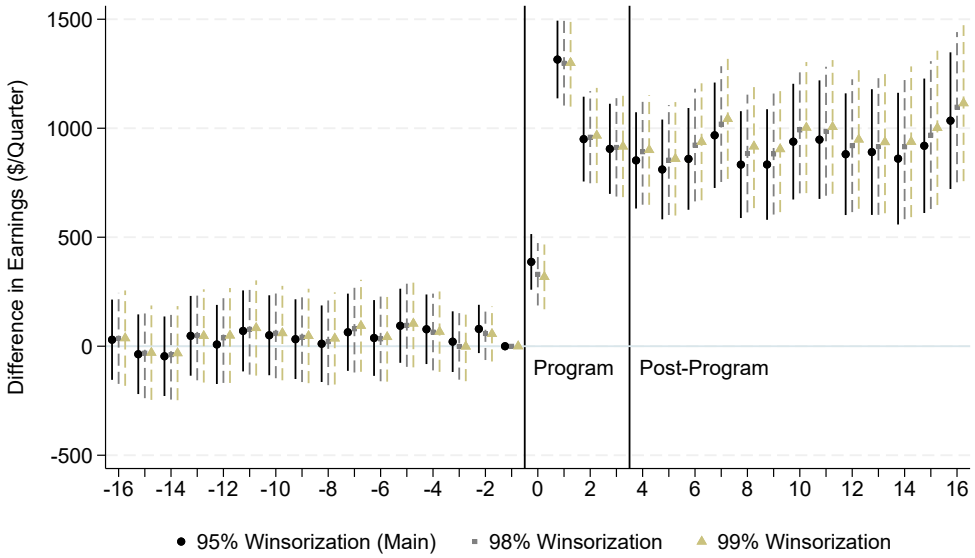


(b) Quarterly Earnings



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.

Figure A-8: Event-Study Effects on Quarterly Earnings, Varying Earnings Winsorization



Notes: Data source is administrative UI earnings data and STEP program 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. Earnings are winsorized within cohort-by-quarter cells as described in Section 3.1.

Table A-5: Aggregated Quarterly Earnings Effects, Varying Earnings Winsorization

	(1) Earnings	(2) Earnings	(3) Earnings
In-Program (Q0-Q3)	855.6***	838.3***	833.9***
	(74.90)	(82.37)	(86.05)
Post-Program (Q4-Q16)	860.7***	905.6***	921.0***
	(102.6)	(112.5)	(117.3)
Comparison Mean (Q < 0)	1805.6	1924.9	1974.0
Comparison Mean (Q0-Q3)	1587.4	1678.0	1710.1
Comparison Mean (Q4-Q16)	2429.0	2554.2	2606.4
N (Individuals)	22613	22613	22613
UI Earnings Winsorization	95%	98%	99%

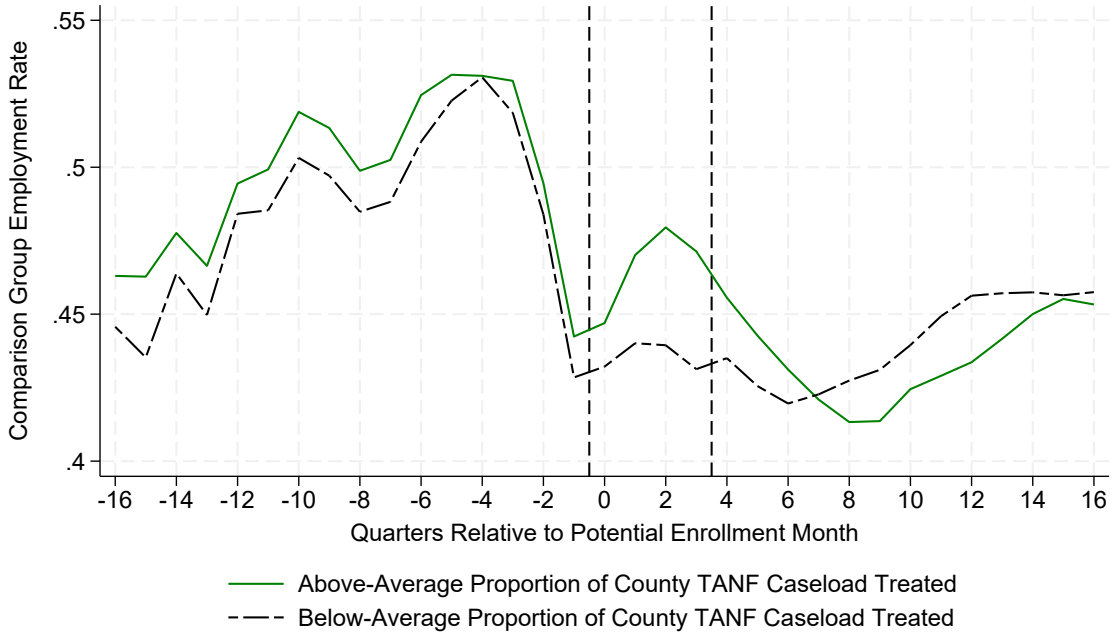
Notes: Data source is administrative UI earnings data and STEP program data from CDLE. Quarter 0 represents the quarter in which an individual enrolled in STEP and is thus a different calendar quarter from person to person. Formal employment is defined as having UI-covered earnings in Colorado greater than $0 in a given quarter. Formal-sector earnings are winsorized within cohort-by-quarter cells as described in Section 3.1.

Table A-6: Aggregated Effects on Employment and Earnings, Dropping Individuals Enrolled Past Quarter 3

	(1) Employment	(2) Employment	(3) Earnings	(4) Earnings
In-Program (Q0-Q3)	0.257***	0.212***	850.1***	729.2***
	(0.0132)	(0.0136)	(74.94)	(77.37)
Post-Program (Q4-Q16)	0.0990***	0.0876***	861.9***	797.1***
	(0.0138)	(0.0142)	(102.7)	(105.5)
Comparison Mean ($Q < 0$)	0.494	0.491	1805.1	1790.2
Comparison Mean (Q0-Q3)	0.460	0.506	1592.4	1753.8
Comparison Mean (Q4-Q16)	0.437	0.450	2427.3	2523.0
N (Individuals)	22613	22054	22613	22054
Sample	Main	Drop Enrolled Past Q3	Main	Drop Enrolled Past Q3

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: Trends in Comparison Group Employment by County Program Saturation



Notes: Data come from administrative UI earnings data from the CUBS data system and administrative TANF data from CDHS. Each quarterly sample includes 1072 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. High-stability industries are those with average job durations of 4-5 quarters among the comparison group; medium-stability have durations of 3-4 quarters and low-stability have durations of 2-3 quarters. See industry details in Table 7.

Table A-7: Employment and Earnings Heterogeneity by Sanction/Re-Engagement or Barrier/Exemption Status

	(1) Employment	(2) Earnings
Panel A: Sanction/Re-Engagement in Current TANF Spell		
In-Program (Q0-Q3)	0.254***	879.7***
	(0.0155)	(89.31)
Post-Program (Q4-Q16)	0.0927***	880.2***
	(0.0162)	(122.4)
(Sanc./Re-eng.)X(In-Program)	0.00910	-111.2
	(0.0286)	(159.0)
(Sanc./Re-eng.)X(Post-Program)	0.0238	-68.95
	(0.0303)	(217.3)
N (Individuals)	22613	22613
Panel B: Barrier/Exemption in Current TANF Spell		
In-Program (Q0-Q3)	0.225***	724.4***
	(0.0218)	(129.3)
Post-Program (Q4-Q16)	0.0789***	769.4***
	(0.0216)	(166.6)
(Barrier/Exempt.)X(In-Program)	0.0517+	204.4
	(0.0272)	(157.3)
(Barrier/Exempt.)X(Post-Program)	0.0327	150.4
	(0.0279)	(210.2)
N (Individuals)	22613	22613

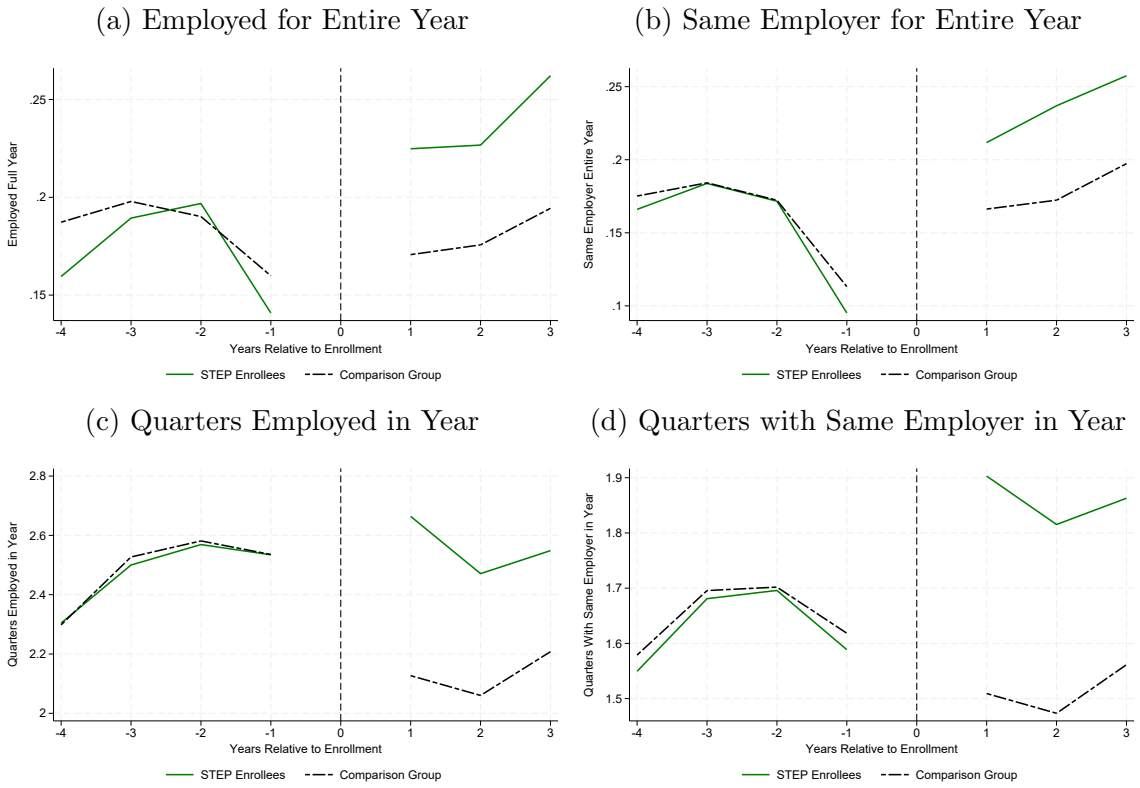
Notes: Data sources are administrative UI earnings data and STEP program tracking data from CDLE. All employment and earnings measures for STEP Enrollees include imputations for STEP placements that do not appear in the UI earnings data (see section 3.1). Sanction/re-engagement designations and barriers/work requirements exemptions are measured as of the current TANF spell up to and including the enrollment month. 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. All estimates in this table use the main propensity score estimated on the full set of covariates reported in Table A-2.

Table A-8: Employment and Earnings Heterogeneity by Age of Youngest Child and TANF History

	(1) Employment	(2) Earnings
Panel A: Age of Youngest Child		
In-Program (Q0-Q3)	0.245***	718.8***
	(0.0189)	(102.6)
Post-Program (Q4-Q16)	0.0949***	764.0***
	(0.0203)	(144.8)
(Age of Youngest)X(In-Program)	0.00220	26.74
	(0.00277)	(16.35)
(Age of Youngest)X(Post-Program)	0.000895	22.85
	(0.00292)	(21.50)
N (Individuals)	22589	22589
Panel B: Cumulative TANF Months		
In-Program (Q0-Q3)	0.247***	859.0***
	(0.0203)	(118.1)
Post-Program (Q4-Q16)	0.109***	925.7***
	(0.0208)	(156.7)
(Cumul. TANF Mo.)X(In-Program)	0.000755	-0.432
	(0.000980)	(5.173)
(Cumul. TANF Mo.)X(Post-Program)	-0.000747	-4.556
	(0.00104)	(7.565)
N (Individuals)	22613	22613

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). Sanction/re-engagement designations and barriers/work requirements exemptions are measured as of the current TANF spell up to and including the enrollment month. 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. All estimates in this table use the main propensity score estimated on the full set of covariates reported in Table A-2.

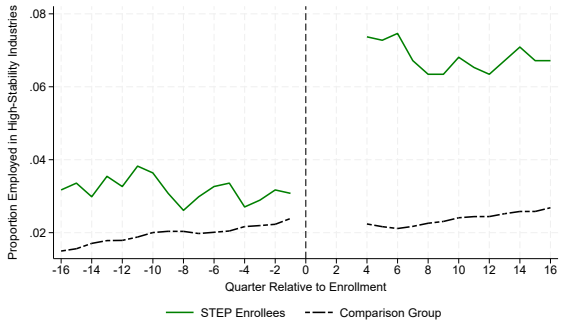
Figure A-10: Trends in Job Stability Measures



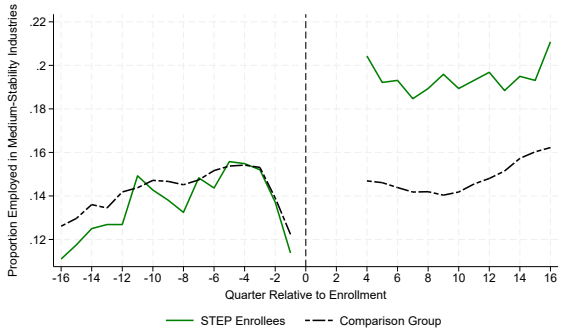
Notes: Data come from administrative UI earnings data from the CUBS data system and administrative TANF data from CDHS. Each quarterly sample includes 1072 STEP participants who applied between January 2018 and March 2020. Quarter or Month 0 represents the quarter or month in which an individual enrolled 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.

Figure A-11: Trends in Employment by Industry Stability

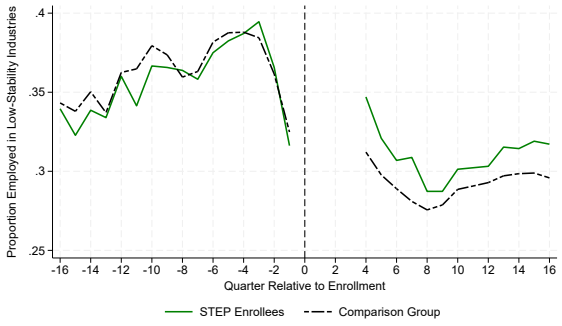
(a) High-Stability Industries



(b) Medium-Stability Industries



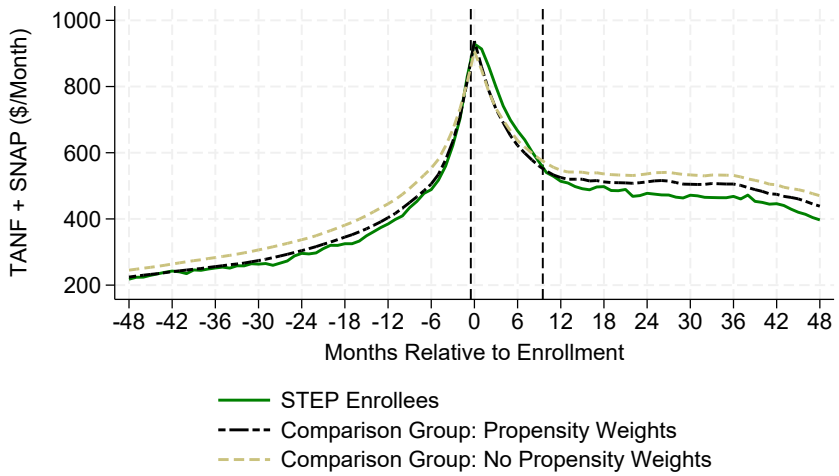
(c) Low-Stability Industries



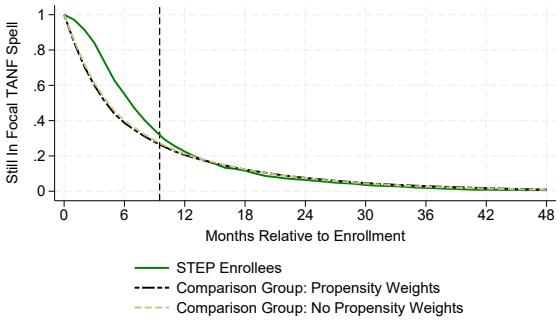
Notes: Data come from administrative UI earnings data from the CUBS data system and administrative TANF data from CDHS. Each quarterly sample includes 1072 STEP participants who applied between January 2018 and March 2020. Quarter or Month 0 represents the quarter or month in which an individual enrolled 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. High-stability industries are those with average job durations of 4-5 quarters among the comparison group; medium-stability have durations of 3-4 quarters and low-stability have durations of 2-3 quarters. See industry details in Table 7.

Figure A-12: Additional Trends in Benefit Receipt

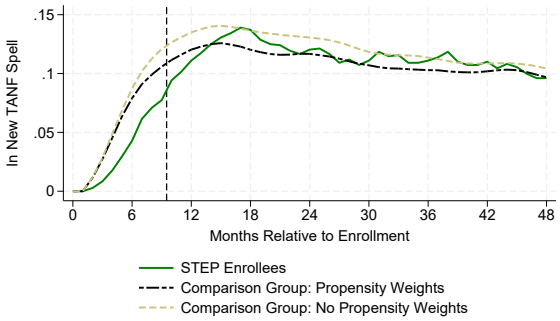
(a) Combined Monthly TANF + SNAP Amount



(b) TANF Receipt - Initial Spell



(c) TANF Receipt - Subsequent Spell



Notes: Data come from administrative TANF data from CDHS. Each quarterly sample includes 1072 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.