

Further Education During Unemployment^{*†}

Pauline Leung
Cornell University

Zhuan Pei
Cornell University

May 2020

Abstract

While policymakers often promote further education for displaced workers, evidence on its effectiveness in the U.S. context primarily comes from evaluations of specific government sponsored training programs, which only represent one narrow avenue for skill acquisition. This paper studies the returns to retraining among unemployed workers, where retraining is broadly defined as enrollment in community colleges, four-year institutions, and technical centers. We link together high quality administrative records from the state of Ohio and estimate the returns using a matching design in which we compare the labor market outcomes of similar workers who do and do not enroll. Our matching specification is informed by a separate validation exercise in the spirit of LaLonde (1986), which evaluates a wide array of estimators using a combination of experimental and non-experimental data in a setting similar to ours. We graphically present the average quarterly earnings trajectories of the enrollees and matched non-enrollees over a nine-year period and show that there is little difference in earnings pre-enrollment, followed by temporarily depressed earnings among enrollees during the first two years after enrolling, and sustained positive returns thereafter. We find that enrollees experience an average earnings gain of seven percent three to four years after enrolling, and that the returns are driven by those who switch industries, particularly to healthcare.

Keywords: Training, Unemployment, Community College
JEL codes: J24, J68, I26

*We thank Amanda Eng, Rebecca Jackson, Suejin Lee, and Katherine Wen for excellent research assistance and Rajeev Darolia, Sue Dynarski, Nathan Grawe, Mike Lovenheim, Jordan Matsudaira, Doug Miller, Ronni Pavan, Haiyuan Wan, Abbie Wozniak, and participants at AASLE, APPAM, CEU, CIRANO-CIREQ, Duke, IZA ed workshop, Northwestern, NTA, Princeton, UC Davis, and Zurich for helpful comments. We are grateful to Lisa Neilson and the staff members at the Center for Human Resource Research at Ohio State University, the Ohio Department of Jobs and Family Services, and the Ohio Department of Higher Education for providing the data and answering our many questions. We thank Jeff Smith for generously sharing the National JTPA Study data. Financial support from the Cornell Institute of Social Sciences is gratefully acknowledged. All errors and opinions are our own.

†The Ohio Longitudinal Data Archive is a project of the Ohio Education Research Center (oerc.osu.edu) and provides researchers with centralized access to administrative data. The OLDA is managed by The Ohio State University's Center for Human Resource Research (chrr.osu.edu) in collaboration with Ohio's state workforce and education agencies (ohioanalytics.gov), with those agencies providing oversight and funding. For information on OLDA sponsors, see <http://chrr.osu.edu/projects/ohio-longitudinal-data-archive>.

1 Introduction

The U.S. labor market has become increasingly polarized in recent decades, as high- and low-skilled jobs grow at the expense of middle-skilled jobs that traditionally employ workers with moderate levels of education (Autor, Katz and Kearney, 2006, 2008). These trends continued through the Great Recession and the recent COVID-19 contraction, resulting in disproportionately high unemployment among workers without a college degree (Katz, 2010; Hoynes, Miller and Schaller, 2012; Bureau of Labor Statistics, 2020). A long line of research shows that job displacement, especially during economic downturns, is associated with large and persistent earnings losses, adverse health outcomes, and negative impacts on the children of the unemployed (Jacobson, LaLonde and Sullivan, 1993; Couch and Placzek, 2010; Sullivan and von Wachter, 2009; Davis and von Wachter, 2011; Oreopoulos, Page and Stevens, 2008; Stevens and Schaller, 2011). To mitigate these social and economic costs, economists and policymakers across the political spectrum have advocated for new skill acquisition through further education.

However, much of our knowledge about the returns to further education for unemployed workers in the U.S. comes from evaluations of specific government-sponsored training programs (e.g., the Workforce Investment Act program, or WIA, now replaced by the Workforce Innovation and Opportunity Act, or WIOA), even though these programs only constitute one narrow avenue through which unemployed workers upgrade their skills. Many workers enroll directly in a local postsecondary institution such as a community college, who potentially account for a much larger share of the retraining population. In the fall of 2017, 4.5 million nontraditional (i.e., 25 years old or above) undergraduate students were enrolled nationwide, compared to 1.6 million participants in the largest U.S. training program in 2016-17 (WIOA Adult and Dislocated Worker programs), of which only a minority received training services (Snyder, de Brey and Dillow, 2019*b*; Social Policy Research Associates, 2018). Recent policy efforts also do not solely focus on government sponsored training programs. At the peak of the Great Recession, the U.S. Departments of Labor and Education created the website opportunity.gov and encouraged state governments to contact unemployment insurance claimants and inform them of resources (e.g., federal financial aid) and institutions (e.g., community colleges) for reskilling. While there is an extensive literature on the returns to community college education (see Kane and Rouse, 1999 and Belfield and Bailey, 2011 for reviews), much less is known about the characteristics of unemployed workers who enroll and their subsequent labor market outcomes.

This study estimates the labor market returns to retraining among unemployed workers, where retraining

is broadly defined as enrollment in a postsecondary institution. We link together high quality administrative data from the state of Ohio, which include unemployment insurance (UI) claims, quarterly wage records, course enrollment and credential data from all in-state public higher education institutions (community colleges, four-year colleges, and technical centers), and WIA records. By following the population of UI claimants from 2004 to 2011 who enroll, we see that the majority of retraining does not occur within the context of a narrowly defined training program: in our data, we observe nearly 88,000 workers enroll in public postsecondary institutions following a layoff, compared with 27,000 workers who train through WIA.

To estimate the returns to retraining, we use a matching design that compares the labor market outcomes of unemployed workers who pursue further education (enrollees) versus observably similar workers who do not (matched non-enrollees). Our matching specification is informed by the large literature that uses selection-on-observables designs to evaluate training programs in both the U.S. and international contexts (see McCall, Smith and Wunsch, 2016 for a comprehensive review). Moreover, we conduct our own validation analysis in the spirit of LaLonde (1986), using data from the National Job Training Partnership Act Study (NJS) which evaluates a predecessor of WIA training programs. Building on the influential work by Heckman, Ichimura and Todd (1997), Heckman et al. (1998), and Heckman and Smith (1999), our validation exercise assesses 1) whether selection-on-observables strategies are likely to recover the true causal effect for recently unemployed workers as in our context, 2) whether flexible machine-learning techniques to model the propensity score improves on the conventional logit-based method, and 3) how the fixed effects approach of Jacobson, Lalonde and Sullivan (2005*a,b*) performs relative to matching methods. We arrive at several important findings that extend the previous analyses of NJS. First, we pinpoint a context in which matching performs well: when the sample has detailed past earnings information and consists of workers previously attached to the labor market (for whom past earnings are likely to be predictive of future prospects). Second, we find that machine learning techniques are no panacea and cannot substitute for the lack of pre-treatment earnings data. Third, the fixed effects approach fares worse than matching methods in this training evaluation. Since we find that conventional logit-based propensity score matching performs competitively in a sample similar to ours, we adapt this approach to our main analysis.

We graphically present the average earnings trajectories of enrollees and matched non-enrollees in the five years before and four years after enrollment. The trajectories reveal little difference in earnings pre-enrollment, followed by temporarily depressed earnings of enrollees while they are in school (the “lock-in” effect), and sustained positive returns thereafter. Overall, the average return among enrollees is \$387 per

quarter, or about seven percent, in the third and fourth years after enrolling. A decomposition of this return reveals that retraining affects earnings mostly at the extensive margin. While the magnitudes of enrollment returns are heterogeneous across various subgroups, we consistently observe positive earnings gains four years after enrolling. Following an early subset of workers for a longer period, we find that the retraining effect persists and widens to 16 percent at the end of a ten-year horizon.

Another advantage of our study relative to existing training program evaluations is our ability to look into the “black box” of retraining. That is, we observe the courses taken and credentials received by enrollees in our sample, which allows us to explore the types of training underlying our returns estimates. Using simple accounting methods, we find that the returns to enrollment are driven by workers who train and subsequently find employment in new industries post-layoff, particularly the healthcare sector.

This paper makes the following contributions. First, it bridges two largely separate strands of empirical literature on training programs and on community colleges by studying the policy-relevant unemployed worker population that intersects with both. As mentioned above, the U.S. training literature focuses mainly on evaluating government sponsored programs, which finds mixed results.¹ In the community college literature, recent studies that use administrative earnings and transcript data show that associate degrees yield returns of about 18 to 26 percent relative to no degree, mixed effects of other credentials, and positive returns to healthcare-related programs (see review by Belfield and Bailey, 2017a). Although unemployed workers are represented in both training program and in community college populations, the degree of overlap and the extent to which prior estimates apply to them is unknown. The only studies that directly examine a broad population of unemployed workers are Jacobson, Lalonde and Sullivan (2005a,b), which measure the returns to attending community college for long-tenured Washington state workers laid off in the early 1990s. Their preferred specification suggests that the returns are between six and eight percent, although the estimates can be sensitive to the model used. Relative to Jacobson, Lalonde and Sullivan (2005a,b), we further connect the literatures by incorporating data from non-community college settings where some WIA training takes place (e.g., technical centers) and estimate the returns for WIA trainees in our sample. We find that the return to WIA training three to four years after enrollment begins is lower (four percent) compare to our overall return (seven percent), largely driven by the longer lock-in period.

¹For WIA, recent experimental and non-experimental evaluations find zero to long-lasting negative effects of training for dislocated workers (Heinrich, Mueser and Troske, 2008; Andersson et al., 2013; McConnell et al., 2016; Fortson et al., 2017). For the Trade Adjustment Assistance (TAA) program, which provides training to workers affected by trade, one non-experimental evaluation finds initially large negative effects that fade to zero over a four-year period, while another study utilizing quasi-random variation on TAA petition approvals finds positive effects (Schochet et al., 2012; Hyman, 2018).

Second, this paper brings together and critically assesses the methodological approaches from the training and community college literatures. Many recent community college studies that use longitudinal earnings data rely on fixed effects models similar to those in Jacobson, Lalonde and Sullivan (2005*a,b*) (see review by Belfield and Bailey, 2017*a* and discussion in Dynarski, Jacob and Kreisman, 2018), who stress the importance of accounting for complex earnings dynamics of displaced workers and present a major improvement over past approaches plagued by identification issues. We find that adopting matching strategies common in the training literature leads to further improvement in our NJS and Ohio samples. Specifically, when using fixed effects models, there appear to be differential pre-trends in both samples, similar to the finding by Jacobson, Lalonde and Sullivan (2005*a*) in their application. Importantly, specifications that mitigate the differential pre-trends exacerbate bias in treatment effect estimates in our validation study.

Finally, this paper sheds light on the returns to retraining for a recent period, which includes the Great Recession and covers a wide range of economic conditions. Recency of data is important: labor market trends such as the rise in automation and trade in past decades (Autor, Dorn and Hanson, 2013; Autor et al., 2014; Acemoglu and Restrepo, 2017) may have impacted training returns, which makes our estimates more informative for current policy-making relative to those by Jacobson, Lalonde and Sullivan (2005*a,b*) from the early 1990s. The economic boom and bust in our sample period allow us to speak to the literature examining the dependence of educational returns on labor market conditions. Consistent with previous studies (Lechner and Wunsch, 2009; Kahn, 2010; Oreopoulos, von Wachter and Heisz, 2012), we find retraining to be more effective for those enrolled during the Great Recession, who sought jobs afterwards in a thawing labor market. Provided that local postsecondary institutions can overcome the COVID-19 interruptions, our findings will be valuable as the U.S. navigates out of the ongoing economic contraction.

The rest of the paper is organized as follows. We provide institutional background on unemployment insurance and the postsecondary education system in Section 2. Section 3 describes the data and analysis sample. We discuss the empirical strategy in Section 4 and present results in Section 5. Section 6 concludes.

2 Institutional Background

2.1 Unemployment Insurance

In this paper, we identify unemployed workers as those who claim unemployment insurance. To be eligible for UI, workers must have lost a job through no fault of their own and have sufficient earnings and work

weeks prior to job loss. In Ohio, workers must have worked at least 20 weeks and have an average weekly wage of about \$200 per week in the one-year period that begins five calendar quarters prior to job loss. As a result, our study population consists of displaced workers previously attached to the labor force.

While workers generally need to actively search for jobs and be available to work in order to continue receiving benefit payments, they can pursue “approved training” opportunities without losing UI eligibility. States vary in their definitions of approved training, though it generally includes vocationally-oriented or basic education training. According to NASWA (2010), the Ohio unemployment agency automatically approves all training through workforce programs and has 7,000 courses listed as approved. It also approves academic courses that do not lead to a specific occupation on a case-by-case basis.

2.2 Postsecondary Institutions

Our study focuses on the impact of classroom training, which can take place in a number of settings. First, workers may choose to attend four-year institutions or community colleges and enroll in courses that may eventually lead to an associate, bachelor’s, or graduate degree. These institutions, especially community colleges, may also have programs that lead to certificates or high school diplomas. Alternatively, workers may enroll in courses offered at technical centers. Technical centers typically offer occupation-specific programs that may lead to a state license or other credentials. Examples include state license for practical nursing or professional certification in welding.

The cost of attendance varies by institution and program. According to the Integrated Postsecondary Education Data System (IPEDS), the average tuition and fees across Ohio institutions were approximately \$6,400 per year in 2010. However, since unemployed workers are often financially constrained, they are likely to be eligible for and rely on several forms of financial assistance. First, workers may be eligible for federal financial aid such as Pell grants, subsidized loans, and tuition tax credits.² Second, they may obtain training funding from workforce programs like WIA or TAA. In WIA, eligible participants receive an Individual Training Account (ITA) voucher that can be used towards approved training (fewer than 20 percent of WIA participants receive training services, while the others only receive “core” or “intensive” services such as assistance in job search, placement, employment and career planning). The TAA program provides tuition assistance to workers affected by import competition. To understand the relative sizes of the

²Although the amount of federal aid typically depends on income from about two years prior, displaced workers may qualify for simplified needs tests or automatic zero expected family contribution starting in 2009.

various sources of financial assistance, Barnow and Smith (2016) report that in 2014, Pell Grants for those who pursued vocational education totaled \$8.2 billion. In contrast, the expenditures for the WIA Dislocated Worker program and TAA were only \$1.2 and \$0.3 billion, respectively.

3 Data, Analysis Sample, and Descriptive Statistics

Our analysis primarily draws on several administrative data sources from Ohio: 1) UI claim records, 2) student records from public postsecondary institutions, including four-year colleges/universities, community colleges, and technical centers, 3) quarterly wage records, and 4) WIA participant records.

For our analysis of labor market returns, we study workers who file an eligible UI claim between 2004 and the third quarter of 2011. To focus on workers who sought further education after unemployment, we exclude those who enrolled at any point within two years prior to layoff. Our analysis sample contains 2 million claims, coming from 1.3 million unique individuals (see Appendix A for details on sample construction and data elements). The UI records contain the claim date, demographics (gender, race, number of dependents, age, and zip code), and prior job information (industry and occupation).³

Our schooling data cover all public postsecondary institutions in Ohio. The Higher Education Information system (HEI) records contain enrollment information for 37 public two- and four-year institutions. We also have student records from 53 publicly funded technical centers through the Ohio Technical Centers (OTC) database. Despite the expansive coverage of the HEI and OTC data, we do not know whether a worker enrolls in a private institution. While we acknowledge this to be a limitation of our paper as our “non-enrollees” may in fact enroll in a program we do not observe, many studies in the training literature also suffer from the same issue. For example, a survey of WIA participants by Fortson et al. (2017) reveals that among those who do not receiving training services (i.e., who receive only “core” or “intensive” services), at least 34 percent pursue non-WIA-funded training opportunities within 30 months of initial participation. In comparison, the share of observed non-enrollees in our data enrolling in a private institution is likely to be smaller—as Turner (2017) reports, undergraduate enrollment in public institutions accounts for 77 percent of enrollment in 2007-08 among students aged 22-39.⁴ In the HEI data, we observe terms enrolled, courses taken, and degrees or other credentials obtained (i.e., graduate or professional, bachelors,

³The number of dependents is recorded because the maximum UI benefit amount is higher when a claimant has more dependents. However, only workers with high enough prior earnings receive the maximum UI amount. Since many workers’ UI benefits do not change with dependents, this measure likely understates the true number of dependents.

⁴Of the remaining 23 percent enrolled in private institutions, 45 percent are in for-profit institutions.

associate, or less than two-year awards). In the OTC data, the courses offered range from one-day courses to certificate programs that last several years. We observe the dates of enrollment in courses as well as any credentials obtained.

We construct our main outcome variables using quarterly earnings data from the state's UI system (we discuss how out-of-state earnings may impact our estimates in Appendix A). In addition to earnings, we also observe, through the third quarter of 2017, number of weeks worked from 2003 and industry for each private sector employer from 1995. The long earnings history allows us to observe at least three years of pre-layoff earnings. We also use this data to construct measures of pre-layoff job tenure and outcomes like industry switching.

Finally, we also observe whether a worker in our analysis sample is in the WIA Standardized Record Data, which cover participants of WIA Adult, Youth, and Dislocated Worker programs. We use information on the dates of WIA training to identify the subset of enrollees observed in the HEI and OTC data who received WIA training services. Next we provide descriptive statistics on enrollee demographics, the timing of enrollment, and enrollment characteristics.

Who enrolls? Panel A of Table 1 presents descriptive statistics for our UI claimant sample by enrollment status. In this table and throughout the paper, an enrollee is a worker who enrolls within two years of layoff.⁵ Of the two million claims in our data, 87,760, or 4.4 percent, are followed by enrollment in a public postsecondary institution. While women make up only 33 percent of UI claimants, they are better represented among enrollees, at 45 percent. Compared with 12 percent among non-enrollees, African Americans make up 18 percent of the enrollee population. In terms of prior job characteristics, enrollees are less likely to have worked in manufacturing, construction, and transportation, have lower job tenure, are younger, and have lower prior earnings (all earnings are expressed in 2012 dollars).

When do workers enroll? We conduct a sequence analysis to graphically represent the temporal patterns of employment and enrollment in Figure 1. Because of the computational burden, we randomly select a 0.5 percent sample from our full set of UI claimants to produce this figure. For workers in the sample, we track their labor market and enrollment history from 12 quarters before UI claim until 12 quarters after. For each quarter, we code workers into one of three states: employed (having positive earnings), not employed

⁵Since we observe enrollment term rather than date of enrollment in the HEI data, we approximate enrollment terms winter, spring, summer, and fall to the first, second, third, and fourth calendar quarters, respectively. Since enrollment typically occurs in the fall and spring terms, and those terms are likely to begin earlier than the corresponding fourth and second calendar quarters, we are likely to report an enrollment start date that is on average later than when workers actually begin schooling. In 2013, all Ohio public colleges switched to the semester system, which eliminated the corresponding winter quarter.

(having no earnings and not enrolled), or in training (enrolled), and represent their history with a vector of length 25 (UI claim quarter plus 24 surrounding quarters). We then define the “optimal matching” distance between any two workers in terms of the similarity of their labor market and enrollment history, which is a standard sequence analysis method and implemented in the R package TraMineR.⁶ Finally, we conduct a standard hierarchical clustering analysis to group workers into four “types.” We choose four types or clusters based on the average distance of fused clusters, which increases much more when we have fewer than four groups (see, for example, Chapter 10 of James et al., 2013). The panels of Figure 1 show that the types are roughly as follows: 1) those who have short employment spells prior to layoff, 2) those who are steadily employed prior to layoff but who move in and out of employment afterward, 3) those who are employed for most quarters both before and after layoff, and 4) those who are steadily employed prior to layoff but stay out of employment after layoff. Enrollees tend to be concentrated in the second group of workers. We also note the heterogeneous timing of enrollment relative to the UI claim filing quarter (“Q0” in the figure)—many of those who enroll do not do so immediately after filing for UI.⁷

How long do workers enroll, where do they enroll, and which courses/credentials do they take/obtain?

Consistent with Figure 1, Table 2 shows that workers take time to enroll—on average 3.7 quarters after layoff—and the mean enrollment length is 4.7 terms. Enrollees in our UI claimant sample mostly attend community colleges (75 percent). 67 percent of enrolled workers take at least one occupational course, and the average proportion of occupational courses is 60 percent, where courses are classified as occupational based on their Classification of Instructional Program (CIP) code following the taxonomy by the National Center of Education Statistics. The overall credential receipt rate within four years of enrollment (including sub-baccalaureate awards, licenses, and industry credentials) is about 26 percent. The majority of degrees or credentials obtained are associate degrees or lower: sub-associate credentials account for 48 percent of total credential receipts, associate degrees for 38 percent, and bachelor’s and graduate degrees make up the remaining 13 percent.

⁶The optimal matching distance between two sequences reflects the minimal cost in terms of the number of operations required to convert one into the other, where the operations are insertion, deletion, and substitution. See Gabadinho et al. (2011) for details.

⁷Some workers appear to regularly enroll every other quarter, which in most cases is a by-product of our approximate conversion of enrollment term to calendar quarter as described in Footnote 5.

4 Empirical Strategy

The empirical challenge in estimating labor market returns to enrollment stems from the differences in the characteristics of workers who do and do not enroll that relate to their future earnings potential. If workers with low labor market prospects are more likely to enroll, we may falsely attribute these lower prospects to enrollment and underestimate the returns. Conversely, if workers with higher ability and motivation select into schooling, we may overstate the returns. Following a long line of research in training program evaluation as reviewed by McCall, Smith and Wunsch (2016), we adopt a matching approach to estimate the causal effect of retraining. Because enrollment timing relative to layoff varies across workers as shown in Section 3, we rely on a dynamic matching framework and discuss the identifying assumptions and treatment effect parameters of interest in Section 4.1 below.

Before we proceed, we highlight several rationales for employing a matching strategy and why it may “work” in our context, given the (often justified) skepticism towards it as a research design. First, our sample construction helps circumvent the “Ashenfelter’s dip” problem, a major challenge in training program evaluation. The Ashenfelter’s dip refers to the phenomenon that only trainees experience an earnings dip prior to training, as the decision to retrain is often a reaction to transitory shocks like job loss (Ashenfelter, 1978; Ashenfelter and Card, 1985). Since we start from the set of recently unemployed workers in Ohio, both enrollees and non-enrollees in our sample have experienced a negative earnings shock, mechanically eliminating the issue of Ashenfelter’s dip.

Second, in our matching specification we take advantage of the information available on past labor market histories from our rich administrative data, which helps to recover the causal effects of training as recent studies have argued. As Andersson et al. (2013) succinctly state, “[m]otivated workers, and high ability workers, should do persistently well in the labor market; if so, conditioning on earlier labor market outcomes will remove any selection bias that results from motivation and ability also helping to determine training receipt.” Andersson et al. (2013) also find that adding firm fixed effects does not change causal estimates relative to specifications which incorporate detailed labor market histories. Similarly, in extensive empirical Monte Carlo simulations based on German administrative data, Lechner and Wunsch (2013) find that including variables on firm characteristics, industry- and occupation- specific experience, health, program compliance, desired job characteristics, and detailed regional information does not further reduce bias relative to only using basic demographics and labor market histories. Finally, Caliendo, Mahlstedt and

Mitnik (2014) find that controlling for typically unobserved non-cognitive traits adds little beyond past labor market histories.

Third, we conduct our own validation exercise using data from the NJS to assess the performance of a wide array of estimators in the spirit of LaLonde (1986). As discussed in Section 4.2 and Appendix C, the validation exercise extends influential studies of Heckman, Ichimura and Todd (1997), Heckman et al. (1998), and Heckman and Smith (1999) by further pinpointing a context where matching is likely to work and by expanding the set of evaluated methods to include machine learning based matching procedures and fixed effects models of Jacobson, Lalonde and Sullivan (2005*a,b*). We find that matching performs especially well in subsamples that resemble our Ohio study population—namely workers with previous labor market attachment—for whom past earnings are likely to be predictive of future prospects. We also find that conventional logit-based propensity score matching methods perform competitively, and that they dominate the commonly used fixed effects models of Jacobson, Lalonde and Sullivan (2005*a,b*).

Fourth, the specification informed by our validation study, which we adapt to our Ohio analysis as described in Section 4.3, leads to high quality of matching. In Section 4.4, we show overlapping support in the propensity score distributions (except for outliers that amount to less than 0.1 percent of the enrollee sample) and covariate balance across enrollees and matched non-enrollees.

Finally, we believe matching to be better suited than other empirical strategies for our setting. In Appendix D, we explore the possibility of adopting alternative approaches, including fixed effects models, a distance-based instrumental variable strategy, and a regression discontinuity design from layoff timing. We discuss why they are not appropriate for our setting.

4.1 Identifying Assumptions and Parameters of Interest

The main assumption underlying our approach is the conditional independence assumption, or “unconfoundedness”. Let T denote whether a worker enrolls in school within two years of job loss, with $T = 1$ if the worker enrolls and $T = 0$ if she does not. Let $Y(1)$ and $Y(0)$ denote the potential future earnings of the worker if she does or does not enroll, respectively. The realized outcome we observe is $Y = Y(1)T + Y(0)(1 - T)$.

The *static* unconfoundedness assumption standard in the matching literature is

$$Y(0) \perp\!\!\!\perp T | X \tag{1}$$

where X are observed covariates realized at or prior to job loss. Together with a common support condition— $\Pr(T = 1|X) < 1$ for all x in the support of X —unconfoundedness implies the identification of the treatment effect on the treated (TOT) parameter:

$$E[Y|T = 1] - E[E[Y|T = 0, X]|T = 1] = \underbrace{E[Y(1) - Y(0)|T = 1]}_{TOT} \quad (2)$$

In our study, the “treatment” (enrollment) is allowed to occur over a period of two years, a complication that necessitates a *dynamic* variant of the unconfoundedness assumption. Similar to the settings studied by Sianesi (2004), Fredriksson and Johansson (2008), and Biewen et al. (2014), workers typically consider enrolling when they are still unemployed, so late enrollees are more likely to have experienced long spells of unemployment or sustained low earnings. To the extent that these spells either directly cause or proxy for labor market conditions that determine future earnings, it is important to incorporate them into the conditioning set beyond information available at layoff.

However, there is tension between allowing for different conditioning sets for workers who enroll at different times and the interpretability of the estimated treatment effect. To illustrate, consider a simple two-period case and let the binary variable T^1 (T^2) be equal to one if a worker begins her enrollment in the first (second) period post-layoff and zero otherwise. Since enrollment can only begin in one period and that to enroll at all means to enroll beginning in one of the two periods, $T^1 = 1$ implies that $T^2 = 0$ and $T = T^1 + T^2$. Using X^1 and X^2 to denote the covariates available at the beginning of the first and second period respectively, we state the dynamic unconfoundedness assumption:

Assumption 1. $Y(0) \perp\!\!\!\perp T^1|X^1$ and $Y(0) \perp\!\!\!\perp T^2|X^2, T^1 = 0$.

Assumption 1 says that within the “at-risk” set of workers who have not yet enrolled at the beginning of period s , $s = 1, 2$ (and are therefore able to begin enrollment in period s), the potential outcome is independent of a worker’s enrollment decision conditional on the information available. Following an argument analogous to the static case, dynamic unconfoundedness allows for the identification of a causal effect. But since the treatment is now “enrollment in period s ”, we can only identify the effect of enrolling in the first period relative to not enrolling in the first period (but possibly enrolling in the second period), which is not the TOT parameter in equation (2).

Since we are interested in the TOT of enrolling versus not enrolling, we invoke an additional assumption:

Assumption 2. $Y(0) \perp T^2 | T^1 = 0, X^1$

Along with an overlap condition, Assumptions 1 and 2 allow us to recover the respective TOT parameter for workers enrolling in period 1 and 2, by using only workers who enroll in *neither* period as comparisons:

$$E [Y | T^1 = 1] - E [E[Y | T^1 = 0, T^2 = 0, X^1] | T^1 = 1] = \underbrace{E[Y(1) - Y(0) | T^1 = 1]}_{TOT \text{ for } T^1=1} \quad (3)$$

$$E [Y | T^2 = 1] - E [E[Y | T^1 = 0, T^2 = 0, X^2] | T^2 = 1] = \underbrace{E[Y(1) - Y(0) | T^2 = 1]}_{TOT \text{ for } T^2=1}. \quad (4)$$

We can further identify the TOT parameter $E[Y(1) - Y(0) | T = 1]$ by aggregating (3) and (4).

While the identification result (4) only relies on Assumption 1, we need Assumption 2 for equation (3)—proofs are in Appendix B. To illustrate Assumption 2's role, note that under Assumption 1, the natural comparison group for those enrolling in period 1 consists of workers with similar characteristics who do not enroll in period 1. However, some of the workers in this comparison group end up enrolling in period 2, whose $Y(0)$ we do not observe. Under Assumption 2, we can replace these period-2 enrollees with similar workers who never enroll—for whom we do observe $Y(0)$ —since the assumption implies

$$E[Y(0) | T^1 = 0, T^2 = 1, X^1] = E[Y(0) | T^1 = 0, T^2 = 0, X^1]. \quad (5)$$

Assumption 2 is violated if period-2 enrollees face a worse labor market between period 1 and 2 than their never-enrolled counterparts. In that case, we expect their (unobserved) $Y(0)$ to be lower than that of never-enrollees (i.e., the left hand side of (5) is less than the right hand side), which biases the enrollment effect downward and against our finding a positive return.

As shown in Appendix B, Assumptions 1 and 2 also imply a dynamic propensity score theorem and, with an overlap assumption, allows us to proceed with propensity score matching. In addition, we state the formal assumptions and identification results for the general S -period case in Appendix B ($S = 8$ quarters in our empirical study). Finally, we note that in the training literature, Sianesi (2004), Fredriksson and Johansson (2008), and Biewen et al. (2014) rely only on the dynamic unconfoundedness assumption to identify the effect of "training now versus later", while van den Berg and Vikström (2019) propose similar identifying assumptions to ours as they are also interested in the TOT effect of training versus not training.⁸

⁸The identification framework in van den Berg and Vikström (2019) is more complicated. They explicitly model the time spent unemployed because the eligibility to participate in a training program hinges on staying unemployed in their context.

4.2 Validation Exercise with the National Job Training Partnership Act Study

To assess whether matching and related research designs are likely to recover causal effects in our empirical setting and to guide our matching specification, we conduct a validation study using data from a separate randomized training program evaluation. Here, the word “matching” encapsulates the class of selection-on-observables strategies, which includes regression and inverse probability weighting estimators. We ask three main questions in this validation analysis. First, can matching deliver the true causal effect of enrollment in a setting similar to ours, where the sample consists of recently unemployed workers whose detailed earnings histories are observed?⁹ Second, does the incorporation of machine learning tools improve the performance of matching estimators? Third, how do fixed effects specifications of Jacobson, Lalonde and Sullivan (2005*a,b*), the main alternative empirical strategy, perform relative to matching?

Our validation exercise uses data from the National Job Training Partnership Act Study, an experimental evaluation of a government training program summarized in Orr et al. (1996). In a series of influential studies, Heckman, Ichimura and Todd (1997), Heckman et al. (1998), and Heckman and Smith (1999) analyze this data to examine the selection process into training programs and its implications for non-experimental causal inference methods. Specifically, researchers collected information not only on program participants randomized into the treatment and control groups, but also on individuals who were eligible for but did not apply to participate in the program (“eligible non-participants”, or ENPs). By comparing the outcomes of the experimental control group and the ENPs, the researchers are able to study the extent to which various matching estimators capture the “treatment effect” of zero, in the spirit of the seminal LaLonde (1986) study.¹⁰ The main challenge, as shown in the upper row of Figure 2, is that the earnings of the experimental controls exhibit a pronounced dip, in contrast to the ENPs. However, these studies find that matching-based methods, coupled with rich data on lagged outcomes, reduce bias and lead to treatment effect estimates insignificantly different from zero in many cases. In the other cases where matching does not perform adequately, a difference-in-differences extension of matching is an improvement.

We extend the analyses of Heckman, Ichimura and Todd (1997) and Heckman et al. (1998) to address the three questions posed above. First, since some of the NJS experimental participants (especially among women) had little or no attachment to the labor market prior to training, we document the performance of

⁹As demonstrated by Smith and Todd (2005), the success of matching estimators may be sensitive to sample restrictions.

¹⁰While one could in principle conduct the equivalent exercise of assessing which estimators capture the true experimental treatment effect by comparing the experimental treatment group and the ENPs, it is not feasible in practice because information on pre-period earnings is not available for the experimental treatment group (Heckman et al., 1998).

various estimators both on the full sample and on a subset of recently employed workers (those with at least four quarters of positive earnings within three years prior to program participation, or the “earnings history sample”) that more closely resembles our study population of unemployed workers. Second, we assess whether using machine learning techniques such as forward stepwise selection, penalized logit, genetic matching, variants of classification trees, or support vector machines to generate a comparison group improves the performance of matching estimators.¹¹ As in the earlier analyses of the NJS, we also study the importance of various types of baseline information by using these methods on covariate sets that include only demographics, earnings, or the combination of both. Third, we implement on the NJS samples variants of the fixed effects earnings regressions proposed by Jacobson, Lalonde and Sullivan (2005a), which are now frequently used to estimate the returns to community colleges (even for populations broader than unemployed workers in the original study). We assess specifications that include: 1) individual fixed effects, 2) individual time trends, and 3) individual time trends with layoff effects that vary by observable characteristics. We describe the validation exercise and the different estimators in detail in Appendix C.

The validation exercise motivated by our three questions above reveals a set of findings useful for guiding our Ohio analysis. First, we find that, among those who have recently worked, including prior earnings in matching specifications generally yields unbiased treatment effect estimates. We present this point graphically: the bottom row of Figure 2 shows the earnings of the experimental controls and their matched ENPs for the male and female earnings history samples respectively, where we match using a logit propensity score model containing demographic characteristics and 12 quarters of past earnings. Unlike that of the entire ENP sample graphed in the top row of Figure 2, the earnings of the *matched* ENPs in the bottom row mimic the pre-program dip and post-program rebound of the experimental controls.¹² The examination of the earnings history sample extends the findings of Heckman, Ichimura and Todd (1997) and Heckman et al. (1998) in that it further pinpoints when matching performs well. Specifically, matching does not appear to meaningfully reduce bias in the full sample of women (in contrast to men). However, once we eliminate the substantial fraction (42 percent) of women with mostly zero pre-program earnings, the bias reduction from matching on past earnings is large and consistent with results for men (only 18 percent of men have

¹¹The pioneering work by Heckman et al. (1998) implements a matching estimator rooted in pruned classification trees, and we branch out to study the performance of other tree-based methods such as bagging, random forest, and boosting.

¹²It is worth noting that although 12 quarters of pre-program earnings enter into the propensity score estimation equation, there is still imbalance in the pre-program earnings between the experimental controls and matched ENPs. More generally, a variable serving as an input into the propensity score model does not mechanically guarantee post-matching balance in that variable. In our analysis of Ohio data below, we provide evidence of higher matching quality.

mostly zero pre-program earnings). Within the earnings history sample, we find no further improvement when using matched difference-in-differences, the best performing method in Heckman, Ichimura and Todd (1997) and Heckman et al. (1998). While broadly consistent with the emphasis on the importance of prior labor market dynamics in existing studies, our findings are more specific: a matching design incorporating past earnings is likely to work well in populations previously attached to the labor market, for whom past earnings are predictive of future prospects.

Second, we find that although machine learning tools offer more flexible functional forms to model the propensity score or use “data-driven” ways to match workers, there does not appear to be a universally dominant method, as the ranking of the matching methods in terms of bias magnitude varies across samples. Much more consistent is the pattern that incorporating prior earnings in the conditioning set substantially reduces the bias across methods. Perhaps surprisingly, once we include prior earnings in the conditioning set, conventional logit propensity score methods where prior earnings enter linearly appear quite competitive.

Third, we find that while individual fixed effect models, which can be thought of as a generalized difference-in-differences estimator, can yield unbiased treatment effect estimates in some cases, specification tests show significant departure from parallel pre-period trends between experimental controls and ENPs. This suggests that these models may not adequately eliminate earnings differences that can result in biased estimates. More importantly, while including linear individual time trends in earnings mitigate the issue of differential pre-trends, doing so may actually exacerbate bias. We conclude that relative to matching, fixed effects specifications are less likely to produce unbiased estimates in our training context.

There are of course limitations to this validation exercise. Other empirical strategies not considered may outperform those that we have tried. Additionally, the sample size for the validation exercise is not very large, and more sophisticated matching methods may perform better in settings with more observations or covariates. Despite these limitations, the competitiveness of logit propensity score matching with detailed earnings history in populations previously attached to the labor force bolsters our confidence, so we proceed with this strategy and provide further evidence on the quality of the matching design below.

4.3 Matching Specification

When translating the lessons from the validation exercise in the previous subsection to our setting, we are mindful of the differences between the NJS and our Ohio data. The biggest advantage the latter offers is the much larger sample size, which allows us to combine exact and propensity score matching with detailed

earnings history to ensure that our comparison sample is as similar as possible to the enrollees. In this subsection, we describe in detail our main matching specification and its rationale.

As described in Section 3, we define enrollees as those who start school within eight quarters after filing a UI claim. Given the heterogeneous enrollment timing shown in Figure 1 and the identification results in Section 4.1, we estimate the TOT separately for those who enroll in the first through eighth quarter relative to layoff. For each of these eight cohorts, the potential comparison pool of workers includes those who do not enroll within the two-year period after layoff.

We perform exact matching along three dimensions. First, we require that enrollees be exactly matched to non-enrollees laid off in the same quarter. This is motivated by the fact that economic conditions and policies varied widely over our study period. Workers laid off in 2004 may differ from those starting unemployment at the peak of the Great Recession and face dramatically different labor market landscapes.¹³ Furthermore, policies enacted to help workers overcome challenges during the Great Recession, such as UI extensions and information campaigns about resources for retraining, may have influenced workers' decisions to enroll (Barr and Turner, 2015, 2016). By comparing workers displaced around the same time, we attempt to control for the influence of time-varying labor market conditions and policies.

Second, following the training evaluation literature, we require that workers be exactly matched on gender, as decisions to enroll may differ between men and women. As argued by Heckman and Smith (1999), men's decisions to enroll may be more heavily influenced by economic prospects, while women's decisions may depend more on family responsibilities. The different motivations for enrolling may translate into different training effects across gender.

Finally, we exactly match enrollees and non-enrollees based on whether they were working in manufacturing at layoff. The manufacturing sector is of particular interest to policymakers, as it has been in rapid decline over the past decades (Autor, Dorn and Hanson, 2013), particularly in "rust belt" states like Ohio.

Within the (exactly matched) layoff quarter, gender, and manufacturing cells for each of the eight enrollment timing cohorts, we estimate a separate propensity score model. As demonstrated in our validation study, it is crucial to include pre-enrollment earnings, and we use three years of pre-layoff quarterly earnings as (linear) inputs into the model. In addition to pre-layoff earnings, we also include quarterly earnings

¹³Although Heckman et al. (1998) stress the importance of matching workers from the same local labor market, evidence from Michalopoulos, Bloom and Hill (2004) and Mueser, Troske and Gorislavsky (2007) suggests that this is less important when comparison groups are drawn from a single state. Since our data come from (moderately sized) Ohio, temporal rather than geographic variation capture most of the variation in labor market conditions within our sample. That said, we also include the unemployment rate in the month and county of layoff in the propensity score model, which accounts for geographic differences.

between layoff and enrollment per our identification framework in Section 4.1, which proxy for the worker’s success in finding reemployment (we further illustrate the importance of incorporating these earnings with empirical evidence in Section 5.1). Finally, we include in our propensity score model demographic and prior job characteristics to the extent that our data allow it.¹⁴ Specifically, we include race indicators (white, African American, other, or unknown), previous sector indicators (construction, wholesale trade, administrative support and waste management, healthcare and social assistance, accommodation and food services, retail trade, transportation, and other), job tenure categories (less than one year, one to six years, and more than six years), age indicators (age below 19, each year from age 19 and 59, and older than 59), whether a worker has a dependent, and county unemployment rate during the month of layoff.

We use the estimated propensity score to pair each enrollee with her nearest neighbor from the comparison sample, where each comparison worker may be matched to more than one enrollee (i.e., matching with replacement). Choosing a larger number of neighbors may further reduce variance in the estimated treatment effect, but the quality of the match may deteriorate as we allow for larger propensity score differences between the enrollee and comparison workers. Since we have a large enough sample size to attain precise estimates and are more concerned with bias, we use only one neighbor in our main specification. We do, however, explore the sensitivity of our results with respect to the number of neighbors in Section 5.1.

After each logit propensity score estimation, we estimate the (cell-specific) TOT with the average difference in outcomes between each enrollee and its match, and the standard error is computed following Abadie and Imbens (2016). We then aggregate the TOT’s across the enrollment-timing \times layoff-quarter \times gender \times sector cells to obtain the overall treatment effect, where the weights are proportional to the number of enrollees in each cell.

4.4 Matching Design Quality Check: Overlapping Support and Covariate Balance

As emphasized by Smith and Todd (2005), we need to assess the validity of the overlapping support assumption. We first point out that some observations indeed appear to violate this assumption, but their number is very small. Specifically, only 73 out of the nearly 88,000 enrollees are dropped because of this violation—enrollment is perfectly predicted for 71 observations, which are assigned an estimated propensity score of one; the other two observations have an estimated propensity score close to one and above the trimming

¹⁴Out of the 980 propensity score models, 24 do not converge to a solution with the full set of covariates. In these cases, we estimate the model by eliminating one covariate (i.e., one industry dummy, demographic variable, or quarter of earnings).

threshold from the algorithm by Imbens and Rubin (2015) (p. 367-368).¹⁵

We show the overlap of the enrollee and non-enrollee distributions in Appendix Figure A.1. Because many of the propensity scores are close to zero, we overlay the histograms of the log odds ratio of the two groups, which is a monotone transformation of the propensity score (i.e. $lor \equiv \log \frac{p(x)}{1-p(x)}$), for ease of visual inspection. In the top row, we plot the estimated log odds ratio distributions for workers who enroll one, four and eight quarters post layoff, and overlay the corresponding distributions for the non-enrollees. We see that the *lor* distributions have little support on the positive range, indicating that the propensity score is significantly below one, and not surprisingly enrollees tend to have a higher propensity to pursue further education. We show the *frequency* plots of *lor* for the two groups in the bottom row, which are more relevant for assessing overlapping support. Because the number of observations in the non-enrollee group is far larger than that in the enrollee group, there appears to be sufficient overlap even in the higher range of the *lor*.

To show covariate balance, we follow Imbens (2015) and report in Panel A of Table 1 the normalized differences between each covariate in the enrollee and non-enrollee samples. That is, for each covariate X , we report $(\bar{X}_E - \bar{X}_N) / \sqrt{(S_{X,E}^2 + S_{X,N}^2) / 2}$, where \bar{X}_E and \bar{X}_N are the respective sample means of X among enrollees and non-enrollees, and $S_{X,E}$ and $S_{X,N}$ are the corresponding sample standard deviations. Even in the full sample, the normalized differences are moderate: except for age, they are all within the 0.3 rule-of-thumb measure used by Imbens (2015) in the exploratory analysis of the Imbens, Rubin and Sacerdote (2001) data. In Panel B of Table 1, we show the average characteristics of the *matched* enrollee and non-enrollee samples. Indeed, covariates are balanced across the two groups: normalized differences are very small, and all t -tests fail to reject equality despite the large sample size. We proceed to present our main empirical results on training impacts in the next section and provide additional graphical evidence to support the balance in the pre-enrollment earnings trajectories.

5 Empirical Results: Returns to Further Education During Unemployment

5.1 Overall Returns

We begin by graphically presenting the earnings effects of pursuing further education using the full matched sample. In Figure 3, the solid line shows the average earnings of enrollees over time, from 20 quarters

¹⁵The remaining difference in the enrollee sample size between Panels A and B of Table 1 is due to the elimination of the winter quarter of 2013 in the HEI data mentioned in Section 3. We do not estimate the propensity score and enrollment effect for (OTC only) enrollees who start in that quarter, which eliminate an additional 24 enrollees.

before until 16 quarters after enrollment begins. The dashed line shows the earnings averaged across each enrollee's nearest neighbor in the non-enrollee sample (see Appendix Figure A.2 for a comparison of earnings trajectories for enrollees and unmatched non-enrollees). Prior to enrollment, both enrollees and their closest comparison workers have similar earnings trajectories, increasing in the period from five years prior to enrollment to approximately two years prior, before dropping to 50 percent of the peak by the time of enrollment. The seemingly slow decline in earnings is due to the fact that we are averaging the earnings of workers who enroll at different times post-layoff, and not because of a drawn-out earnings reduction process for all workers. After enrollment, the two lines begin to diverge. Enrollees have lower earnings relative to their comparison group for approximately two years before returns materialize. This "lock-in" effect, which may come about because enrollees are more constrained in their ability to search for jobs and work while in school, is consistent with the finding in Heinrich et al. (2013) for the WIA Dislocated Worker program. The gains appear to grow after the lock-in period while the earnings of non-enrollees flatten. In the third and fourth years post-enrollment, enrollees earn \$387 more per quarter than non-enrollees as reported in Table 3, a gain of about seven percent. It is notable that even at more than four years after layoff, both enrollees and non-enrollees do not catch up to their pre-layoff earnings.

We conduct several sensitivity checks on our average returns estimate. While the set of non-enrollees in Figure 3 are selected by using 12 quarters of pre-layoff earnings in the propensity score formulation, Panels A and B of Appendix Figure A.3 show that earnings patterns are robust to alternative matching specifications that include fewer quarters of earnings (one pre-enrollment quarter and four pre-enrollment quarters, respectively). However, when we match without earnings between layoff and enrollment in Panel C, the patterns are quite different: the enrollees in this panel have much lower earnings than their matched comparison group heading into enrollment. Appendix Figure A.3 echoes the discussion in Section 4 and highlights the importance of controlling for the most recent information just before treatment, without which the enrollment decision may be negatively correlated with potential labor market prospects, invalidating the matching design. We also explore the sensitivity of our findings to using more neighbors. We show in Panel A of Appendix Figure A.4 the estimated treatment effect against the number of neighbors used for male non-manufacturing workers laid off in the first quarter of 2009, the cell with the largest number of claims (we focus on this subsample to ease the significant computational burden of running the matching analysis 25 times using the full sample). The estimated treatment effect tends to shift upward as we increase the number of neighbors, and the variance is reduced by up to six percent as we show in Panel B of the same

figure. An estimate using the full sample with five neighbors follows this general pattern, but since our sample size allows for precise estimates, we use the lowest bias one-neighbor specification.

The two panels of Appendix Figure A.5 plot the probability of having positive earnings and weeks worked per quarter, respectively. Prior to enrollment, we see a similar pattern of a rise and drop in employment for both groups. After enrollment begins, enrollees are less likely to be employed initially but eventually overtake the comparison group at approximately the two-year mark. A natural question that arises is whether or not the gain in employment can explain all of the gains in earnings, or whether enrollment increases both employment and wage rates. One way to answer this question is to note that in Panel B of Appendix Figure A.5, non-enrollees work for 7.2 weeks per quarter three to four years after enrollment, which implies a weekly wage of approximately $\$5217/7.2 = \729 . Enrollees, on the other hand, are employed 7.8 weeks and have a weekly wage of approximately $\$723$. This indicates that increased wage rates are not driving the returns to enrollment.¹⁶ Another way to see this is to examine the earnings distributions of enrollees and matched non-enrollees, shown in Appendix Figure A.6. While the top panels of the figure document similar earnings distributions before enrollment, the bottom panels show that the distributions start to diverge eight quarters after enrollment and further widens at the end of quarter 16, where the gains are concentrated at the extensive margin. Therefore, we conclude training mainly affects employment and likely has minimal effect on wage rates four years after enrolling.

5.2 Returns By Subgroup

We now present the returns to enrollment by subgroup. For subgroups that have exactly matched participants (i.e., enrollment timing, layoff quarter, gender, and sector), we estimate the returns by restricting to enrollees within the subgroup and examine the difference in outcomes compared to their matched non-enrollees. For subgroups that are not exactly matched (age, tenure, and race groups), we first restrict the analysis sample to the subgroup and then re-match enrollees to non-enrollees using the estimated propensity score described in Section 4.3.¹⁷ That is, we do not implement another matching procedure with re-estimated propensity scores

¹⁶Since Appendix Figure A.5 shows a small gap in weeks worked between enrollees and non-enrollees in the pre-period, the weeks effect may be overstated. To see how this affects our conclusion, consider the following decomposition of the earnings effect: $\Delta \text{Earnings} = \Delta \text{Weeks} \times \frac{\text{Earnings}_N}{\text{Weeks}_N} + \Delta \frac{\text{Earnings}}{\text{Weeks}} \times \text{Weeks}_E$ where Earnings and Weeks are average quarterly earnings and average weeks worked in a quarter, respectively, for enrollees (E) and non-enrollees (N), and Δ denotes the difference between enrollees and non-enrollees. If we adjust the weeks effect downward by 0.15 weeks (the gap in the pre-period), the first term of this decomposition still accounts for 83 percent of the total earnings effect.

¹⁷We cannot simply compare enrollees within a certain subgroup to their original matched comparison groups because this compares the outcomes of enrollees within a subgroup to non-enrollees that are potentially not in the subgroup. A simple way to

conditional on the subgroup, which is computationally expensive and could bring in workers not included set of matched non-enrollees in Section 5.1, resulting in inconsistencies across samples.

A potential downside of this approach is that the nearest neighbor for an enrollee in the subgroup analysis may differ from that in the full sample analysis and may be a worse match. But as we show in Appendix Table A.1, our method still achieves balance in pre-enrollment earnings in the majority of subgroups. Even within the not-exactly-matched categories—age, tenure, and race—where the match quality is more of a concern, after adjusting for multiple hypotheses with the Holm method only one subgroup remains statistically imbalanced at the 5 percent level (Holm is less conservative than the better known Bonferroni method). Even for this subgroup of workers under 40, the imbalance appears moderate as shown in Panel A of Appendix Figure A.11.¹⁸ To account for this imbalance, which manifests in largely parallel trajectories between enrollee and matched non-enrollees before earnings start to decline, we also report returns estimates using difference-in-differences matching in Table 4. These estimates are obtained by first differencing outcome earnings with (symmetrically timed) pre-enrollment earnings within person, and then comparing the resulting differences between enrollees and matched non-enrollees. The reported standard errors in Table 4 are from testing whether the average pairwise difference between enrollees and matched non-enrollees is different from zero, and we abstract away from the sampling variation in forming these matched pairs. This abstraction is inconsequential for the full sample where we also have available the standard errors that account for the sampling variation in forming matched pairs per Abadie and Imbens (2016): compared with their counterparts in Table 3, the two standard errors in columns (1) and (6) of the first row in Table 4 are only slightly larger.

By Enrollment Timing As discussed in Section 3, workers typically do not enroll immediately after lay-off, and some workers take longer than others to go back to school. Consistent with the discussion in Heckman, LaLonde and Smith (1999) that training is a form of job search, enrollees tend to be workers who are not yet reemployed. Correspondingly, those who enroll later have relatively long spells of unemployment. Comparing estimates for those who enroll earlier or later therefore speaks to the question of whether the

think about this is to consider a randomized experiment: if a population is randomly assigned to a treatment or control group with a coin flip, the propensity to be treated for the entire population is 50 percent. If we then wanted to estimate the treatment effect for women only, we cannot simply compare the outcomes of treated women with the entire control population, even though their “propensity scores” are the same at 0.5.

¹⁸Note that, because the nearest neighbor may change from the full sample to a particular subgroup, balance in the full sample and one subgroup does not imply balance in the complement of that subgroup.

returns are larger or smaller for longer-term unemployed workers who may have lower economic prospects.

In Appendix Figure A.7, we show the impacts of enrollment by the quarter in which workers go back to school within two years after layoff. Examining the pre-enrollment period, we see a sharp drop in earnings that is similar across all eight panels, each corresponding to enrolling in a certain quarter (one through eight) since job loss. However, by comparing the matched non-enrollees' earnings trajectories in each graph, it is clear that the groups differ in the extent to which their earnings recover post-layoff. Relative to workers who enroll sooner, the counterfactual earnings profiles of later enrollees are flatter throughout the post-period. As reported in Table 4, we find that workers who enroll later have a smaller “lock-in” effect in the first two years after enrolling, and larger returns to enrollment (up to 14 percent) in the third and fourth years. The lack of lock-in for later enrollees likely reflects their low opportunity cost of schooling. While we find larger returns for those who enroll later, we note that this finding does not imply that it would be more advantageous for workers to delay enrollment. Rather, the exercise compares across different types of workers and quantifies the gains for each group.

By Layoff Year Since our study covers the Great Recession period, there is considerable variation in the labor market conditions at the time of enrollment and reemployment. Prior studies have shown that training programs are more effective when the unemployment rate is high at program entry and low when training ends (Lechner and Wunsch, 2009; Kluve, 2010; Card, Kluve and Weber, 2015). In particular, poor labor market conditions at program entry is associated with smaller lock-in effects due to lower opportunity costs of training, as well as larger gains at the end of training. Outside of the training literature, there is also evidence that graduating college during an economic downturn is associated with persistent earnings losses (Kahn, 2010; Oreopoulos, von Wachter and Heisz, 2012).

In Appendix Figure A.8, we present the earnings of enrollees relative to matched non-enrollees separately by year of layoff. Consistent with the literature, we find that the returns to training do appear to be larger among workers laid off during the Great Recession relative to before: Table 4 shows that workers displaced in 2009-2011 have returns of 8 to 12 percent in the third and fourth years post-enrollment, while the returns for cohorts laid off prior to the recession in 2006 to 2007 (who would be entering the labor market during the recession) have returns that are indistinguishable from zero. Of course, there are differences in the composition of workers who enter unemployment in different years—for example, workers who lost their jobs at the peak of the recession had higher earnings than workers laid off earlier. However, layoff

timing also plays an important role in determining the returns. The flatness of the earnings trajectories of those laid off in 2006 and 2007 likely results from the lack of job opportunities when the workers completed their training during the Great Recession. In comparison, earnings grew much more quickly for later cohorts who faced a thawing labor market.

By Gender Appendix Figure A.9 shows the returns to enrollment by gender. In contrast to Jacobson, Lalonde and Sullivan (2005*a*), we find that the gains to enrollment are larger for men in the four years post enrollment, though we note that women spend more time in school, at an average of 5.2 quarters over four years versus 4.2 quarters for men. This means that women have a longer “lock-in” period relative to men, and a longer follow-up period may reveal larger returns for women. In the third and fourth years after enrollment, we find that the returns to enrollment are nine percent for men and four percent for women. In Section 5.5, we discuss the different types of courses and credentials obtained by women and men, and the subsequent differences in industries of employment.

By Manufacturing vs. Non-manufacturing Due to the decline of the manufacturing sector and the large earnings losses suffered by displaced workers therein, the effects of retraining among this group have attracted particular policy interests (Couch and Placzek, 2010). Appendix Figure A.10 shows separately how further schooling impacts manufacturing and non-manufacturing workers. Consistent with previous studies on the earnings impact of displacement, manufacturing workers, who make up 28 percent of our sample, appear to have higher average pre-layoff earnings than non-manufacturing workers and experience a correspondingly larger drop in earnings post-layoff. We find that in the third and fourth years post-enrollment, manufacturing workers see a four percent return to enrolling on average, smaller than the nine percent for non-manufacturing workers (Table 4). However, there is reason to believe that the returns will increase further for manufacturing workers in the future. The long lock-in period, which largely contributes to the muting of the third to fourth year return, partly reflects the substantially higher (15 percentage points) likelihood of enrollees from manufacturing to switch to a different industry than matched non-enrollees (we discuss industry switching in more in detail in Section 5.5). Since switchers forgo industry specific skills, it may take longer for the human capital investment to pay off. Furthermore, we see in Panel A of Appendix Figure A.10 that the earnings trajectory of the non-enrollees is quite flat for those displaced from manufacturing, allowing for the possibility of higher longer-run enrollment returns.

Other Subgroups The final rows of Table 4, and the accompanying earnings trajectory plots in Appendix Figures A.11-A.13, show the enrollment returns for several other subgroups of interest. We find significant positive returns across workers of different age and racial groups. Older workers (above age 40) see slightly larger returns than younger workers, which are of a similar magnitude as the results by Jacobson, Lalonde and Sullivan (2005*b*). Although Jacobson, Lalonde and Sullivan (2005*a,b*) focus on long-tenured workers who have worked three or more years before layoff, we find that the shortest-tenured group with less than one year of tenure enjoys the highest returns at eight percent. Finally, while there is no consensus on how returns to community college vary by race (Belfield and Bailey, 2011) and estimates are not available in Jacobson, Lalonde and Sullivan (2005*a,b*), we find higher returns for whites than for African Americans.

5.3 Returns For Workforce Investment Act Participants

As discussed in the introduction, much of the evidence on the returns to retraining in the U.S. are from evaluations of specific training programs, such as those funded by WIA. While our study population consists of a broader set of workers who seek retraining, this section focuses on the WIA trainees in our sample to better connect to the literature on government sponsored training programs. Specifically, we provide descriptive statistics for the subset of workers who have trained under WIA and estimate their returns.

For this analysis, we restrict our sample of enrollees to workers who are also observed to have been enrolled in WIA training within two years of layoff. We identify 9,769 WIA trainees within our enrollee sample of almost 88,000 workers.¹⁹ WIA participants differ in characteristics from our overall enrollee sample. As shown in Appendix Table A.2, half of WIA enrollees are formerly manufacturing workers, compared with 28 percent for the main analysis sample. They are also older and have higher earnings prior to layoff. In terms of schooling characteristics, WIA participants are more likely to enroll in technical centers (41 percent versus 13 percent for the main sample), and are enrolled longer (5.3 versus 4.7 quarters). Finally, WIA enrollees are more likely to obtain a credential (58 percent versus 26 percent), mostly at the associate and sub-associate levels.

Using this sample of WIA participants and the non-enrollees from our main analysis sample, we implement the matching strategy discussed in Section 4.3 with some adjustments for the small number of WIA

¹⁹We observe an additional 17,592 WIA trainees in our analysis sample who do not enroll in a public postsecondary institution, which occurs because WIA training may take place at employers, private/online schools, or other institutions not captured in our data. In Fortson et al. (2017), about 24 percent of a nationally representative sample of WIA trainees enrolled in two-year community colleges; the analogous number in our setting is 19 percent.

participants. Instead of exactly matching on claim quarter, gender, and sector, we simply require enrollees be matched with non-enrollees who were laid off within the same year and are of the same gender (we drop four enrollees in a cell with a perfectly predicted logit model).

Figure 4 shows the earnings of WIA participants and their matched non-enrollees. Compared with Figure 3 for all enrollees, we see that WIA participants have higher average pre-layoff earnings and experience a larger drop at layoff. After enrollment, WIA participants endure a significantly larger lock-in effect (-34 percent relative to non-enrollees in the first two years post-enrollment), catch up to matched non-enrollees more slowly, and see smaller returns (4 percent; statistically significant at the 5 percent level) in years three and four. Not surprisingly, these earnings patterns for WIA participants are similar to those for manufacturing workers as depicted in Panel A of Appendix Figure A.10, given the large proportion of manufacturing workers among WIA trainees. And analogous to our discussion of manufacturing workers in Section 5.2, the trajectories indicate that WIA returns may increase further into the future.

Our results are broadly consistent with a recent randomized evaluation of WIA by Fortson et al. (2017). Their experiment compares the outcomes of randomly selected workers who were offered WIA training services versus those offered only other WIA services. Fortson et al. (2017) find no significant difference in earnings between the two groups 12 quarters after randomization, the last time period observed. However, since the difference in actual training receipt is only about 9 percentage points between the treatment and control groups, the 95 percent confidence interval of their implied training effect contains our estimate. More importantly, the point estimates of Fortson et al. (2017) document similar dynamics as Figure 4, in that the earnings of the treatment group catch up to the control group at about two years after randomization.

5.4 Long-Run Returns

While our data do not span a long enough period to allow for the estimation of longer run enrollment effects for our entire analysis sample, we can examine whether returns are likely to persist by following a subsample of early enrollees further out. For this analysis, we restrict our attention to workers who enroll no later than 2007Q3 (19,686 enrollees) and compare their earnings to their matched non-enrollees for ten years post-enrollment. Since all workers in this subsample were laid off before the Great Recession, they differ from our main analysis sample in that their earnings (and “short-run” returns to enrollment) are somewhat depressed in the four-year follow-up period as shown in Table 4. We see this pattern again in Figure 5, where the twin vertical dashed lines denote the two follow-up periods for our main returns estimates. Despite the

dip in earnings that begins around the second to third year after enrollment, the returns appear to increase over time as the labor market rebounds, ultimately resulting in a 16 percent return in the tenth year after enrollment (point estimate \$859, standard error \$66). We find that enrollees are five percentage points more likely to be employed at the end of the ten-year follow-up period, and conditional on employment, enrollees earn seven percent more (the extensive margin explains 60 percent of the overall earnings gain in a formal decomposition exercise). Our estimates are higher than Jacobson, Lalonde and Sullivan (2005a)'s preferred extrapolated long-run effects of retraining from the 1990s (six to eight percent), which the authors suggest may be downwardly biased due to differential pre-trends.²⁰ Our estimates are qualitatively consistent with the long-run positive effects of TAA for high quality (long duration) training by Hyman (2018), although it is hard to draw quantitative comparisons because Hyman (2018) estimates intent-to-treat effects and the TAA treatment includes both benefit payments and training.

5.5 What Does Schooling Do?

In this section, we explore the mechanisms underlying our main returns estimates. We first present results from simple decomposition analyses showing that the returns are primarily driven by increased employment in certain industries, especially healthcare. We then analyze the course and credential data and document that the “excess” employment in these industries can be largely accounted for by the number of workers taking related courses.

We start our analysis by exploring the extent to which enrollees are more likely to leave their pre-layoff industry. The right (left) side of Figure 6 Panel A shows the probability that an enrollee (matched non-enrollee) is employed in their pre-layoff industry, not employed, or employed in a different industry, over the four-year follow-up period. Consistent with Appendix Figure A.5, enrollees are less likely than non-enrollees to be employed immediately after enrollment, and Figure 6 shows that this difference is almost entirely explained by the differential probability of re-employment within the pre-layoff industry, as the two groups are about equally likely to switch industries in quarter zero. Over time, the probability of working in a different industry increases more quickly for enrollees: by the end of the four-year follow-up period, they are 10 percentage points more likely to work in a different industry. Since non-enrollees are 3 percentage points more likely to remain in their pre-layoff industry, enrollees are on net 7 percentage points more likely

²⁰Jacobson, Lalonde and Sullivan (2005a) find returns of 9 and 13 percent for men and women, respectively, per academic year. Since the average number of credits earned in their sample is less than a year's equivalent, we scale down the estimates accordingly.

to be employed in quarter 16 after enrollment.

Panel B of Figure 6 further illustrates the role of industry switching. Analogous to Table 4 of Autor et al. (2014), we decompose the average earnings of enrollees and non-enrollees in each quarter of the follow-up period from Figure 3 into components contributed by industry stayers and industry switchers. Specifically, each group’s quarterly earnings are decomposed as

$$E[\text{Earnings}] = \underbrace{\text{Pr}(\text{Empl. in Same Ind.}) \cdot E[\text{Earnings} | \text{Empl. in Same Ind.}]}_{(i)} + \underbrace{\text{Pr}(\text{Empl. in Diff. Ind.}) \cdot E[\text{Earnings} | \text{Empl. in Diff. Ind.}]}_{(ii)} \quad (6)$$

where “Empl. in Same Ind.” and “Empl. in Diff. Ind” denote employment in the pre-layoff and non-pre-layoff industry, respectively. The gray lines in the figure represent components (i) and (ii) in equation (6) for enrollees (solid) and non-enrollees (dashed), and each pair sums up to their respective black lines. Consistent with Panel A, the earnings difference in quarter 0 is driven by a higher contribution from industry stayers (component i) among non-enrollees. Over time, however, industry switchers among enrollees (component ii) out-contribute their non-enrollee counterparts. In quarter 16, industry switching entirely explains the overall earnings gain: the difference in component (ii) between enrollees and non-enrollees is 117 percent of the overall earnings gain, while the difference in component (i) is -17 percent.

Appendix Figure A.14 shows which industries drive the enrollment effects by plotting the number of non-enrollees and enrollees employed in each sector (or not employed at all) in the final quarter of the follow-up period. Among women (Panel A), enrollees are much more likely to work in healthcare. Among men (Panel B), enrollees are also more likely to work in healthcare (though to a lesser extent), and we see additional gains in construction. Both male and female enrollees are less likely to work in manufacturing. To connect these findings to Figure 6, we show in Appendix Figure A.15 that for both genders, increased employment in healthcare among enrollees is driven by industry switchers, and the decreased employment in manufacturing is the result of a reduction in industry stayers. Industry switchers also contribute to the employment gain in construction for men. This pattern of industry switching driving the effects of enrollment is consistent with Carruthers and Sanford (2018), who study the returns to attending sub-associate level institutions in Tennessee.

Given that increased employment in the healthcare and construction industries appears to explain most of the gains associated with schooling, we now provide evidence that the courses taken and credentials received

by individuals in these sectors are related to their work. We identify courses or credential subjects that are strongly associated with a particular industry, using a mapping of academic subjects to occupations from the National Center of Education Statistics and the joint distribution of occupations and industries from the National Employment Matrix by the Bureau of Labor Statistics. Specifically, we define an academic subject as associated with an industry if the subject prepares a worker for an occupation where more than a quarter of the occupation is employed in a single two-digit industry (see Appendix A.2 for details). For two- and four-year institutions, where individuals typically take multiple courses, we associate an enrollee's coursework with the modal industry across her courses.

Figure 7 shows the fraction of enrollees in each post-layoff industry who have taken associated coursework. We see that a large fraction of female enrollees (Panel A) employed in healthcare have taken related courses. The analogous plot for men (Panel B) shows that a significant fraction of enrollees employed in construction, healthcare, and manufacturing have taken related courses. Strikingly, for both men and women, the number of healthcare or construction course-takers mirrors the enrollee-non-enrollee employment difference in the sector, suggesting that one plausible channel through which training affects earnings is industry-specific skill-building. In addition, we show in Appendix Figure A.16 that enrollees who have taken courses related to health, construction, and manufacturing predominantly work in these sectors.

Figure 8 plots the fraction of enrollees in each industry who obtain a credential related to that industry. For both men and women, healthcare employs a substantial fraction of workers with related credentials. We find that more than two-thirds of those who have taken healthcare courses and are employed in healthcare receive a credential, compared to the overall credential rate of 26 percent. Moreover, we do not find evidence that the gains are driven by a specific occupation within healthcare: of the approximately 4,200 enrollees employed in healthcare with credentials, 24 percent received a credential in Licensed Practical Nursing, 16 percent in Registered Nursing, and 13 percent in Nursing Assistance.

These results on credentials suggest that further education may improve labor market prospects by increasing access to an occupation that requires a license or certification, particularly in healthcare. Using Current Population Survey data from 2016-2019 (Flood et al., 2020), which contain licensing/certification information, we find that among healthcare workers without a bachelor's degree, 38 percent report working in a job that requires a certification or license. In Ohio, that rate is even larger—45 percent report needing a certification or license. However, licensing and certification are likely not the only mechanism for enrollment gains. In the construction industry, for example, we do not see a similar increase in creden-

tials in Figure 8. This is consistent with national statistics indicating that, relative to healthcare, a smaller fraction—18 percent (17 percent in Ohio)—of construction workers without a bachelors degree require a credential. Despite the lack of credential requirements, construction-related course-taking in Figure 7 does appear to drive the enrollment effects for men as shown above. Taken together, the results suggest that the mechanisms through which education increases employment may be heterogeneous across sectors, with both industry-specific skill acquisition and credentialing as potential contributors.

Although one may be tempted to estimate the premium of course-taking in health by examining the differential enrollment effects for workers taking or not taking health courses (where the enrollment effects are estimated by differences in earnings between enrollees and their matched non-enrollees), these returns and premium estimates require additional assumptions to have a causal interpretation. Most importantly, it requires that conditional on observables, the workers who would take health courses if enrolled do not differ systematically from those who would not (formal identification results are in Appendix E, where we augment the potential outcomes framework in Section 4.1). These additional assumptions give rise to strong testable predictions that the propensity scores and baseline covariates are balanced across workers who do and do not take health courses. We find that health course-takers have on average higher propensity scores and lower pre-enrollment earnings relative to those who enroll but do not take health courses, indicating violation of these additional identifying assumptions. In principle, one may calculate the returns to healthcare courses by conducting a separate matching analysis within the enrollee population, where the treatment is healthcare course-taking. Of course, this strategy requires the assumption that conditional on all observed characteristics, enrolling in a health course is as good as random among the set of enrollees. Since it is difficult to find a suitable dataset to conduct a validation exercise for this analysis, estimating the causal effect of taking health courses will require a bigger leap of faith.

To summarize, the results in this subsection suggest that further education improves labor market prospects by allowing workers to acquire industry-specific skills and credentials. For both men and women, training and eventual credential receipt in a healthcare occupation increases the likelihood of switching from non-healthcare employment to healthcare. For men, some gains can also be attributed to reskilling in the construction industry. However, without further assumptions on the selection of enrollees into different types of courses, we cannot conclude that training in these industries offer higher returns than others, as it is possible that workers who ultimately decide to train in health or construction may have similar gains if they pursued other careers.

That said, our finding that the healthcare sector is a key driver of employment growth among those who retrain is consistent with other empirical evidence. Grosz (2020) shows that one specific health program (associate’s degree in nursing) yields significant returns relative to not enrolling at all. Others find evidence that returns to credentials in health appear larger than those for non-health subjects (Bohn, McConville and Gibson, 2016; Stevens, Kurlaender and Grosz, 2019). Finally, anecdotal evidence (e.g., Searcey, Porter and Gebeloff, 2015) and the faster growth of healthcare relative to other sectors both in Ohio and in the U.S. overall (Appendix Figure A.17), are consistent with our finding.

6 Conclusion

In this paper, we estimate the returns to retraining among unemployed workers. Linking together high quality administrative records, we follow the population of Ohio UI claimants between 2004 and 2011 who enroll in public postsecondary institutions. Adopting a matching strategy informed by a validation exercise in the spirit of LaLonde (1986), we find that enrollees experience a “lock-in” effect of depressed earnings immediately after enrolling but see an average return of \$387 per quarter (about seven percent) over the third and fourth years post-enrollment. We find that much of the gain is driven by enrollees who take courses and are subsequently employed in health-related fields. A longer-run follow-up of an early subsample suggests that the returns persist and widen over a ten-year period.

With our earnings impact estimates, we conduct a cost-benefit analysis from the private and social perspectives in Appendix F, similar to Jacobson, Lalonde and Sullivan (2005*b*). We find that the out-of-pocket investment from an average worker breaks even 9 years after enrolling whereas the social investment (including both private investment by the worker and government subsidies) breaks even in 15 years. Assuming an average worker stops working around the normal retirement age as in Jacobson, Lalonde and Sullivan (2005*b*), the private benefit is \$6.60 for every dollar of her out-of-pocket investment in schooling, and the social benefit is \$2.51 for every dollar of total educational investment. The implied marginal private internal rate of return is 15.3 percent, which is in the range of estimates by Heckman, Lochner and Todd (2006) for workers with 12-14 years of education, and the social internal rate of return is 7.9 percent, near the high end of the estimates from Jacobson, Lalonde and Sullivan (2005*b*).

This analysis suggests that policies that encourage and enable unemployed workers to pursue further education, including efforts to expand community college access, can be beneficial in the long run. Fur-

thermore, policies that specifically ease the transition from unemployment to enrollment by targeting UI recipients may be effective. For example, Barr and Turner (2015) show that UI benefit extensions during the Great Recession induced a significant increase in enrollment among the unemployed. Using the same temporal variation in benefit extensions, we replicate their analysis in Appendix G and find a similar effect within Ohio: for every 10-week increase in UI benefits, we estimate a 10 percent increase in enrollment, or an additional 1,200 enrollees per year. Combined with our estimates of positive private and social returns to enrollment, this indicates that UI policies may have social benefits beyond what is typically considered.

Finally, as noted in the previous section, although we document positive returns to enrolling, an open question remains as to whether gains vary significantly across training types. Although we find that the retraining effects are driven by course-taking and subsequent employment in healthcare and construction, evidence that can guide policymakers and unemployed workers on the most effective type of training is a fruitful avenue for future research.

References

- Abadie, Alberto, and Guido W. Imbens.** 2016. "Matching on the Estimated Propensity Score." *Econometrica*, 84(2): 781–807.
- Acemoglu, Daron, and Pascual Restrepo.** 2017. "Robots and Jobs: Evidence from US Labor Markets." National Bureau of Economic Research Working Paper 23285.
- Andersson, Fredrik, Harry J. Holzer, Julia I. Lane, David Rosenblum, and Jeffrey Smith.** 2013. "Does Federally-Funded Job Training Work? Nonexperimental Estimates of WIA Training Impacts Using Longitudinal Data on Workers and Firms." National Bureau of Economics Research Working Paper 19446.
- Ashenfelter, Orley.** 1978. "Estimating the Effect of Training Programs on Earnings." *The Review of Economics and Statistics*, 60(1): 47–57.
- Ashenfelter, Orley, and David Card.** 1985. "Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs." *Review of Economics and Statistics*, 67(4): 648–660.
- Autor, David H., David Dorn, and Gordon H. Hanson.** 2013. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." *American Economic Review*, 103(6): 2121–2168.
- Autor, David H., David Dorn, Gordon H. Hanson, and Jae Song.** 2014. "Trade Adjustment: Worker-Level Evidence." *Quarterly Journal of Economics*, 1799–1860.
- Autor, David H., Lawrence F. Katz, and Melissa S. Kearney.** 2006. "The Polarization of the U.S. Labor Market." *The American Economic Review*, 96(2): 189–194.
- Autor, David H., Lawrence F. Katz, and Melissa S. Kearney.** 2008. "Trends in U.S. Wage Inequality: Revising the Revisionists." *The Review of Economics and Statistics*, 90(2): 300–323.

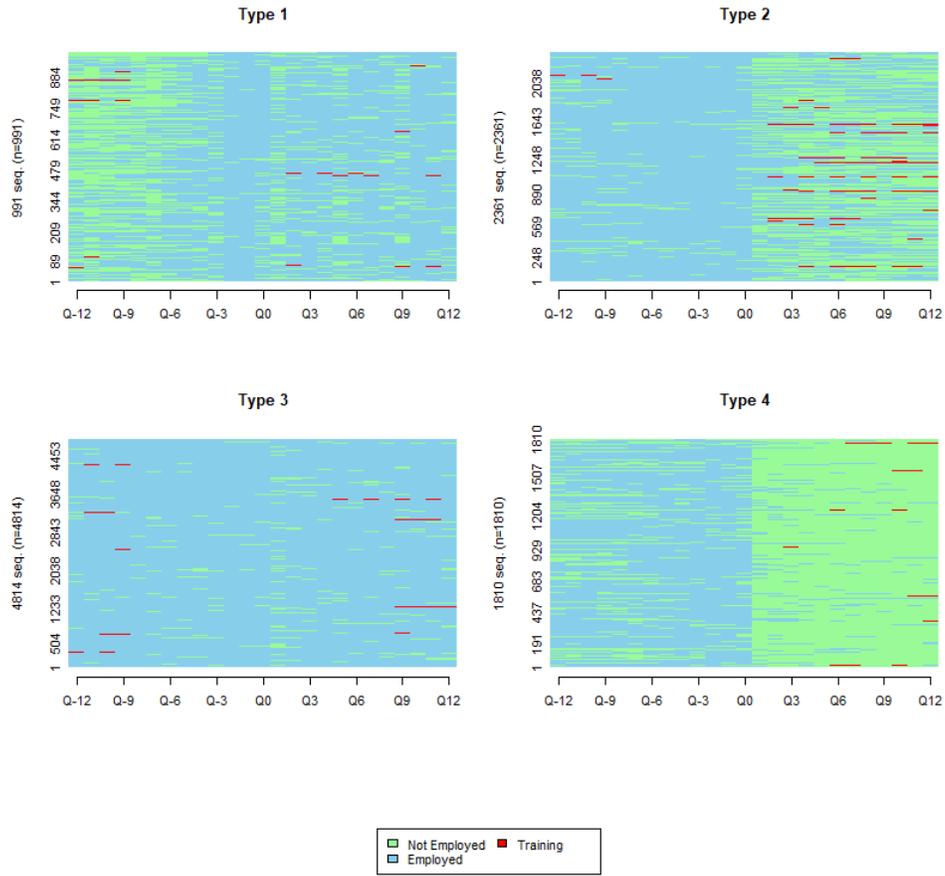
- Barnow, Burt S., and Jeffrey Smith.** 2016. "Employment and Training Programs." In *Economics of Means-Tested Transfer Programs in the United States*. Vol. 2, , ed. Robert A. Moffitt. The University of Chicago Press.
- Barr, Andrew, and Sarah Turner.** 2015. "Out of Work and into School: Labor Market Policies and College Enrollment During the Great Recession." *Journal of Public Economics*, 124: 63–73.
- Barr, Andrew, and Sarah Turner.** 2016. "Aid and Encouragement: Does a Letter Increase Enrollment Among UI Recipients?" Unpublished manuscript.
- Belfield, Clive, and Thomas Bailey.** 2011. "The Benefits of Attending Community College: A Review of the Evidence." *Community College Review*, 39(1): 46–68.
- Belfield, Clive, and Thomas Bailey.** 2017a. "The Labor Market Returns to Sub-Baccalaureate College: A Review." Center for Analysis of Postsecondary Education and Employment.
- Biewen, Martin, Bernd Fitzenberger, Aderonke Osikominu, and Marie Paul.** 2014. "The Effectiveness of Public-Sponsored Training Revisited: The Importance of Data and Methodological Choices." *Journal of Labor Economics*, 32(4): 837–897.
- Bohn, Sarah, Shannon McConville, and Landon Gibson.** 2016. "Career Technical Education in Health." Public Policy Institute of California.
- Bureau of Labor Statistics.** 2020. "The Employment Situation – April 2020." U.S. Department of Labor News Release.
- Caliendo, Marco, Robert Mahlstedt, and Oscar A. Mitnik.** 2014. "Unobservable, but Unimportant? The Influence of Personality Traits (and Other Usually Unobserved Variables) for the Evaluation of Labor Market Policies." IZA Institute for the Study of Labor Discussion Paper 8337. IZA DP No. 8337.
- Card, David, Jochen Kluge, and Andrea Weber.** 2015. "What Works? A Meta Analysis of Recent Active Labor Market Program Evaluations." National Bureau of Economic Research Working Paper 21431.
- Carruthers, Celeste K., and Thomas Sanford.** 2018. "Way station or launching pad? Unpacking the returns to adult technical education." *Journal of Public Economics*, 165.
- Couch, Kenneth A., and Dana W. Placzek.** 2010. "Earnings Losses of Displaced Workers Revisited." *The American Economic Review*, 100(1): 572–589.
- Davis, Steven J., and Till von Wachter.** 2011. "Recession and the Costs of Job Loss." *Brookings Papers on Economic Activity*, 43(2): 1–72.
- Dynarski, Susan, Brian Jacob, and Daniel Kreisman.** 2018. "How important are fixed effects and time trends in estimating returns to schooling? Evidence from a replication of Jacobson, Lalonde, and Sullivan, 2005." *Journal of Applied Econometrics*, 33: 1098–1108.
- Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles, and J. Robert Warren.** 2020. "Integrated Public Use Microdata Series, Current Population Survey: Version 7.0 [dataset]."
- Fortson, Kenneth, Dana Rotz, Paul Burkander, Annalisa Mastro, Peter Schochet, Linda Rosenberg, Sheena McConnell, and Ronald D'Amico.** 2017. "Providing Public Workforce Services to Job Seekers: 30-month Impact Findings on the WIA Adult and Dislocated Worker Programs." Mathematica Policy Research and Social Policy Research Associates.

- Fredriksson, Peter, and Per Johansson.** 2008. "Dynamic Treatment Assignment." *Journal of Business & Economic Statistics*, 26(4): 435–445.
- Gabadinho, Alexis, Gilbert Ritschard, Nicolas S. Müller, and Matthias Studer.** 2011. "Analyzing and Visualizing State Sequences in R with TraMineR." *Journal of Statistical Software*, 40(4): 1–37.
- Grosz, Michel.** 2020. "The Returns to a Large Community College Program: Evidence from Admissions Lotteries." *American Economic Journal: Economic Policy*, 12(1): 226–253.
- Heckman, James, and Jeffrey Smith.** 1999. "The Pre-Programme Earnings Dip and the Determinants of Participation in a Social Programme, Implications for Simple Programme Evaluation Strategies." *The Economic Journal*, 109: 313–348.
- Heckman, James, Hidehiko Ichimura, and Petra Todd.** 1997. "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme." *The Review of Economic Studies*, 64(4): 605–654.
- Heckman, James, Hidehiko Ichimura, Jeffrey Smith, and Petra Todd.** 1998. "Characterizing Selection Bias Using Experimental Data." *Econometrica*, 66(5): 1017–1098.
- Heckman, James, Lance Lochner, and Petra Todd.** 2006. "Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond." In *Handbook of the Economics of Education*. Vol. 1, , ed. Eric A. Hanushek and Finis Welch, Chapter 7, 307–458. Elsevier.
- Heckman, James, Robert LaLonde, and Jeffrey Smith.** 1999. "The Economics and Econometrics of Active Labor Market Programs." In *Handbook of Labor Economics*. Vol. 3, , ed. Orley Ashenfelter and David Card, Chapter 31, 1865–2097. Elsevier Science.
- Heinrich, Carolyn J., Peter R. Mueser, and Kenneth R. Troske.** 2008. "Workforce Investment Act Non-Experimental Net Impact Evaluation." Impaq International Final Report.
- Heinrich, Carolyn J., Peter R. Mueser, Kenneth R. Troske, Kyung-Seong Jeon, and Daver C. Kahvecioglu.** 2013. "Do Public Employment and Training Programs Work?" *IZA Journal of Labor Economics*, 2(6).
- Hoynes, Hilary, Douglas L. Miller, and Jessamyn Schaller.** 2012. "Who Suffers During Recessions?" *The Journal of Economic Perspectives*, 26(3): 27–47.
- Hyman, Benjamin.** 2018. "Can Displaced Labor Be Retrained? Evidence from Quasi-Random Assignment to Trade Adjustment Assistance."
- Imbens, Guido W.** 2015. "Matching Methods in Practice: Three Examples." *Journal of Human Resources*, 50(2): 373–419.
- Imbens, Guido W., and Donald B. Rubin.** 2015. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press.
- Imbens, Guido W., Donald B. Rubin, and Bruce I. Sacerdote.** 2001. "Estimating the Effect of Unearned Income on Labor Earnings, Savings, and Consumption: Evidence from a Survey of Lottery Players." *American Economic Review*, 91(4): 778–794.
- Jacobson, Louis, Robert J. Lalonde, and Daniel Sullivan.** 2005a. "Estimating the Returns to Community College Schooling for Displaced Workers." *Journal of Econometrics*, 125(1): 271–304.

- Jacobson, Louis, Robert J. Lalonde, and Daniel Sullivan.** 2005b. "The Impact of Community College Retraining on Older Displaced Workers: Should We Teach Old Dogs New Tricks?" *ILR Review*, 58(3): 398–415.
- Jacobson, Louis S., Robert J. LaLonde, and Daniel G. Sullivan.** 1993. "Earnings Losses of Displaced Workers." *The American Economic Review*, 83(4): 685–709.
- James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani.** 2013. *An Introduction to Statistical Learning. Springer Texts in Statistics*, Springer.
- Kahn, Lisa B.** 2010. "The long-term labor market consequences of graduating from college in a bad economy." *Labour Economics*, 17(2): 303–316.
- Kane, Thomas J., and Cecilia Elena Rouse.** 1999. "The Community College: Educating Students at the Margin Between College and Work." *Journal of Economic Perspectives*, 13(1): 63–84.
- Katz, Lawrence F.** 2010. "Long-Term Unemployment in the Great Recession." *Testimony for the Joint Economic Committee U.S. Congress*.
- Klueve, Jochen.** 2010. "The effectiveness of European active labor market programs." *Labour Economics*, 17(6): 904–918.
- LaLonde, Robert.** 1986. "Evaluating the Econometric Evaluations of Training Programs with Experimental Data." *The American Economic Review*, 76(4): 604–620.
- Lechner, Michael, and Conny Wunsch.** 2009. "Are Training Programs More Effective When Unemployment Is High?" *Journal of Labor Economics*, 27(4): 653–692.
- Lechner, Michael, and Conny Wunsch.** 2013. "Sensitivity of matching-based program evaluations to the availability of control variables." *Labour Economics*, 21: 111–121.
- McCall, Brian, Jeffrey Smith, and Conny Wunsch.** 2016. "Government-Sponsored Vocational Education for Adults." In *Handbook of the Economics of Education*. Vol. 5, , ed. Eric A. Hanushek, Stephen Machin and Ludger Woessmann, Chapter 9, 479–652. Elsevier B.V.
- McConnell, Sheena, Kenneth Fortson, Dana Rotz, Peter Schochet, Paul Burkander, Linda Rosenberg, Annalisa Matri, and Ronald D'Amico.** 2016. "Providing Public Workforce Services to Job Seekers: 15-month Impact Findings on the WIA Adult and Dislocated Worker Programs." Mathematica Policy Research and Social Policy Research Associates.
- Michalopoulos, Charles, Howard S. Bloom, and Carolyn J. Hill.** 2004. "Can Propensity-Score Methods Match the Findings from a Random Assignment Evaluation of Mandatory Welfare-to-Work Programs?" *Review of Economics and Statistics*, 86(1): 159–179.
- Mueser, Peter R., Kenneth R. Troske, and Alexey Gorislavsky.** 2007. "Using State Administrative Data to Measure Program Performance." *The Review of Economics and Statistics*, 89(4): 761–783.
- NASWA.** 2010. "NASWA Survey on Pell Grants and Approved Training for UI." National Association of State Workforce Agencies.
- Oreopoulos, Philip, Marianne Page, and Ann Huff Stevens.** 2008. "The Intergenerational Effects of Worker Displacement." *Journal of Labor Economics*, 26(3).

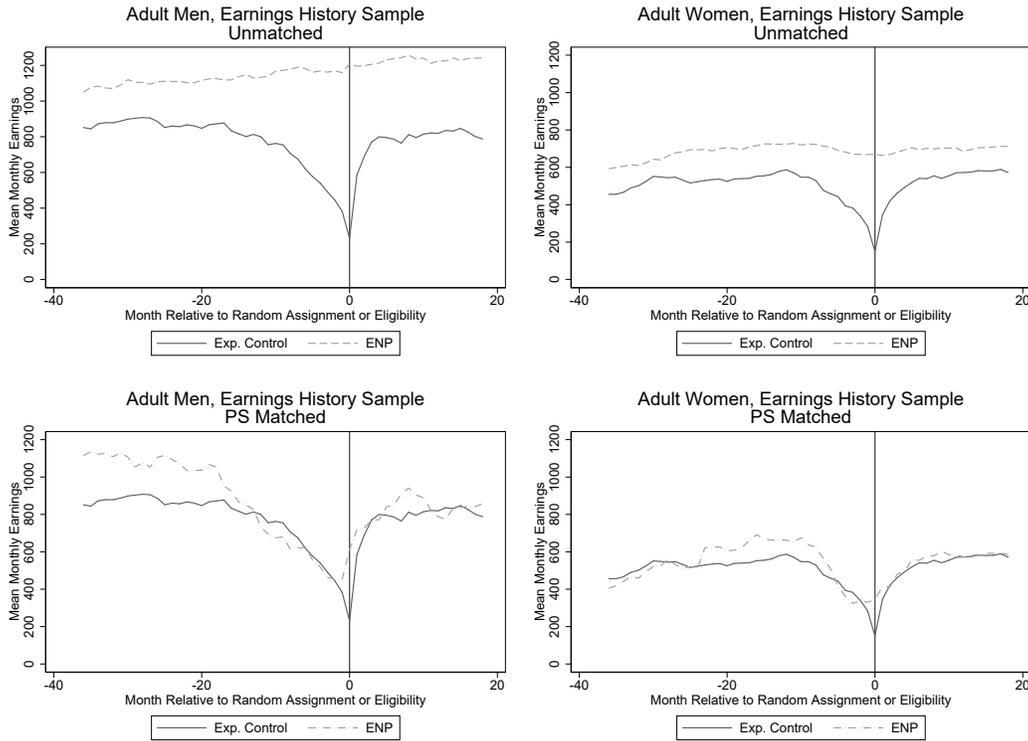
- Oreopoulos, Philip, Till von Wachter, and Andrew Heisz.** 2012. “The Short-and Long-Term Career Effects of Graduating in a Recession.” *American Economic Journal: Applied Economics*, 4(1): 1–29.
- Orr, Larry L., Howard S. Bloom, Stephen H. Bell, Fred Doolittle, Winston Lin, and George Cave.** 1996. *Does Training for the Disadvantaged Work? Evidence from the National JTPA Study*. The Urban Institute Press.
- Schochet, Peter Z., Ronald D’Amico, Jillian Berk, Sarah Dolfin, and Nathan Wozny.** 2012. “Estimated Impacts for Participants in the Trade Adjustment Assistance (TAA) Program Under the 2002 Amendments.” Mathematica Final Report.
- Searcey, Dionne, Eduardo Porter, and Robert Gebeloff.** 2015. “Health Care Opens Stable Career Path, Taken Mainly by Women.”
- Sianesi, Barbara.** 2004. “An Evaluation of the Swedish System of Active Labor Market Programs in the 1990s.” *Review of Economics and Statistics*, 86(1): 133–155.
- Smith, Jeffrey, and Petra Todd.** 2005. “Does matching overcome LaLonde’s critique of nonexperimental estimators?” *Journal of Econometrics*, 125: 305–353.
- Snyder, Thomas D., Cristobal de Brey, and Sally A. Dillow.** 2019b. “Digest of Education Statistics, 2018.” National Center for Education Statistics.
- Social Policy Research Associates.** 2018. “PY 2016 WIOA and Wagner-Peyser Data Book.”
- Stevens, Ann Huff, and Jassamyn Schaller.** 2011. “Short-run Effects of Parental Job Loss on Children’s Academic Achievement.” *Economics of Education Review*, 30: 289–299.
- Stevens, Ann Huff, Michal Kurlaender, and Michel Grosz.** 2019. “Career Technical Education and Labor Market Outcomes: Evidence from California Community Colleges.” *Journal of Human Resources*, 54(4): 986–1034.
- Sullivan, Daniel, and Till von Wachter.** 2009. “Job Displacement and Mortality: An Analysis Using Administrative Data.” *Quarterly Journal of Economics*, 124(3): 1265–1306.
- Turner, Sarah.** 2017. “Labor Force to Lecture Hall: Postsecondary Policies in Response to Job Loss.” The Hamilton Project Policy Proposal 2017-06.
- van den Berg, Gerard J., and Johan Vikström.** 2019. “Long-Run Effects of Dynamically Assigned Treatments: A New Methodology and an Evaluation of Training Effects on Earnings.” IZA Institute of Labor Economics Discussion Paper 12470.

Figure 1: Labor Market and Enrollment Dynamics



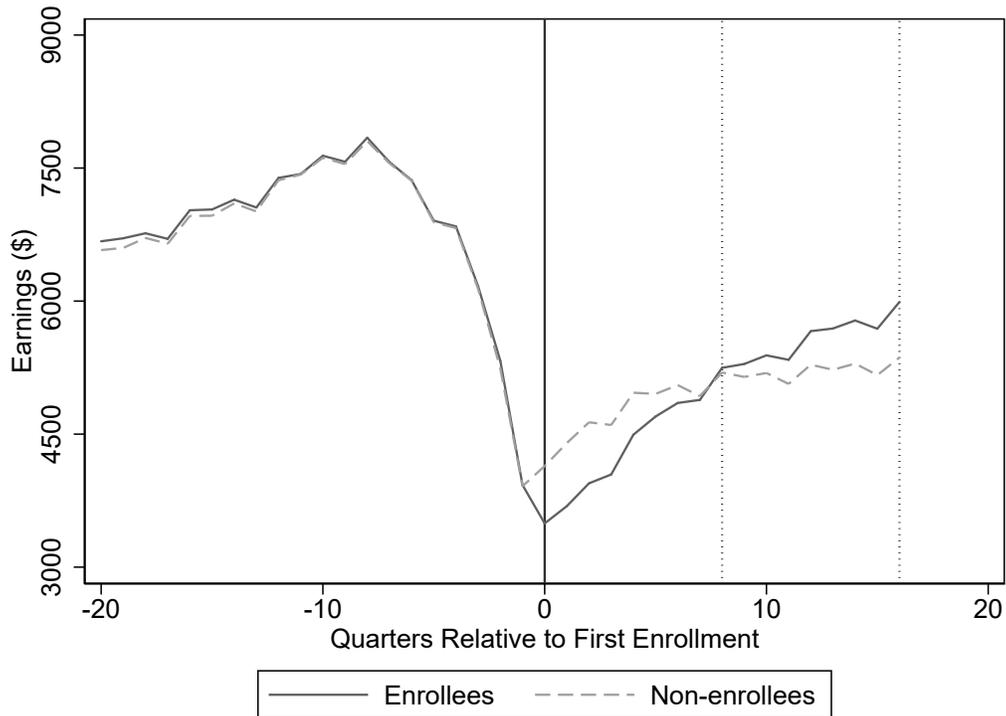
Notes: This figure plots the labor market/enrollment status of UI claimants 12 quarters before and after the UI claim quarter (denoted Q0) for 0.5 percent of our analysis sample. Each “type” groups similar labor market/enrollment sequences. $N = 9,976$ UI claims, corresponding to 9,951 unique individuals.

Figure 2: Earnings of NJS Validation Samples, Unmatched and Propensity-Score-Matched



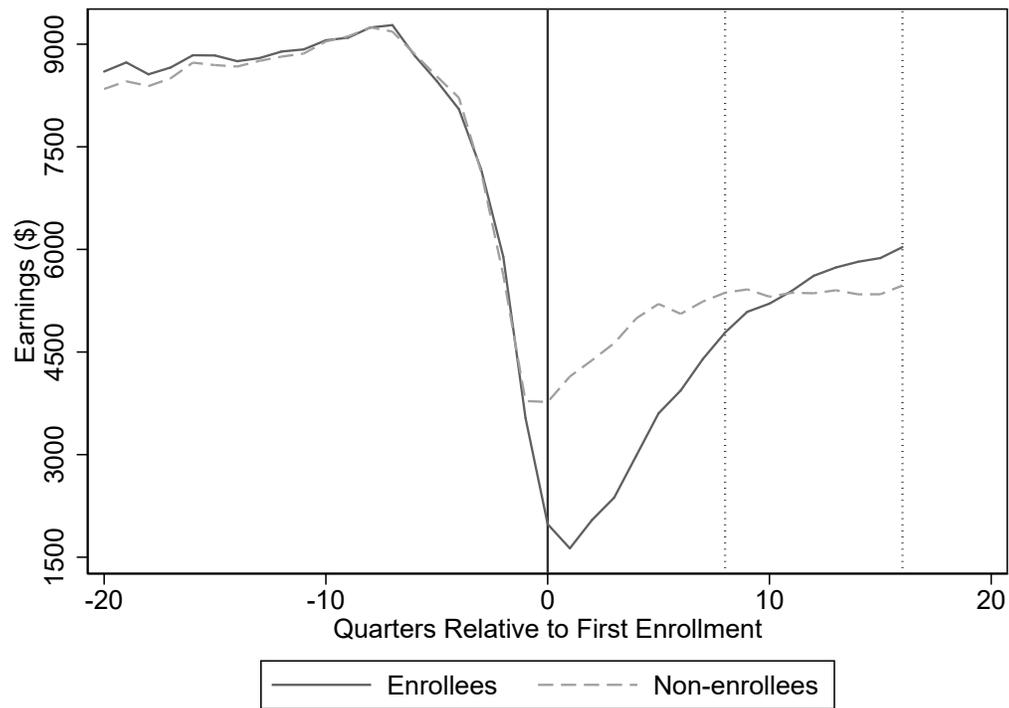
Notes: The upper plots show the mean monthly earnings of the experimental controls and eligible non-participants (ENPs) for the samples of NJS adult men and women with earnings history used in our validation analysis. The lower plots show the mean monthly earnings of experimental controls and their matched ENPs. Observations are matched on propensity scores estimated using a logit model containing: geographic location, 10-year age categories, race/ethnicity indicators, marital status, indicator for having a child under the age of 6, educational attainment, and 12 quarters of pre-period earnings. One nearest neighbor was used for each experimental control observation.

Figure 3: Earnings of Enrollees and Matched Non-enrollees



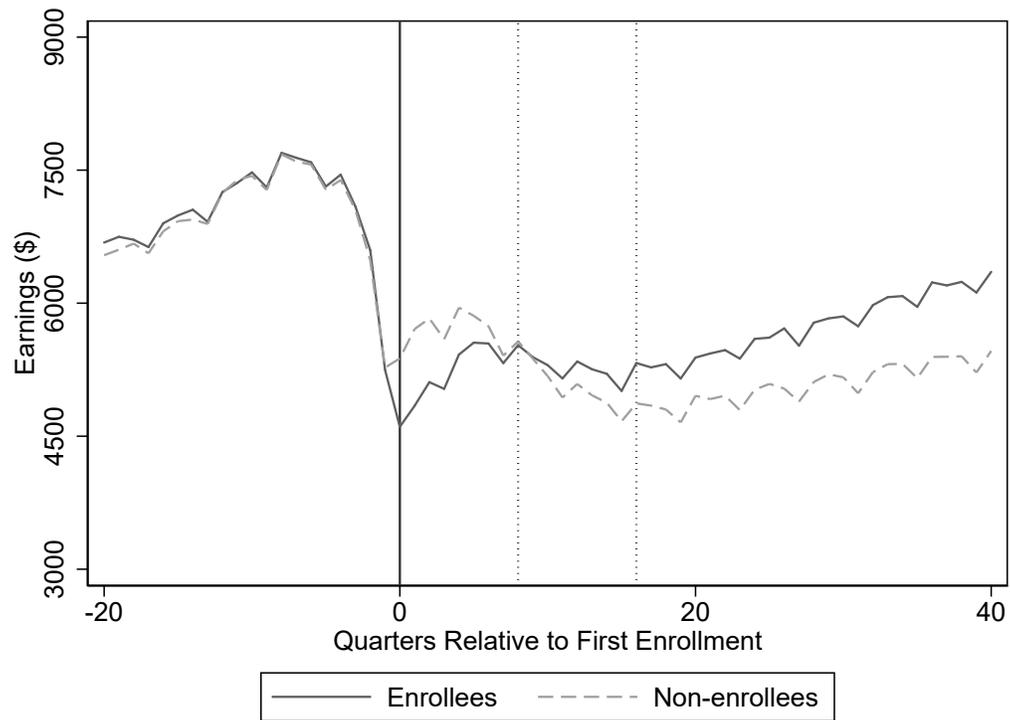
Notes: This figure plots the average quarterly earnings of enrollee and matched non-enrollee UI claimants. The solid vertical line denotes the quarter of first enrollment, and the vertical dashed lines denote eight and 16 quarters after first enrollment. $N = 175,326$, corresponding to 167,422 unique individuals.

Figure 4: Earnings of Enrollees Who Participated in WIA and Matched Non-enrollees



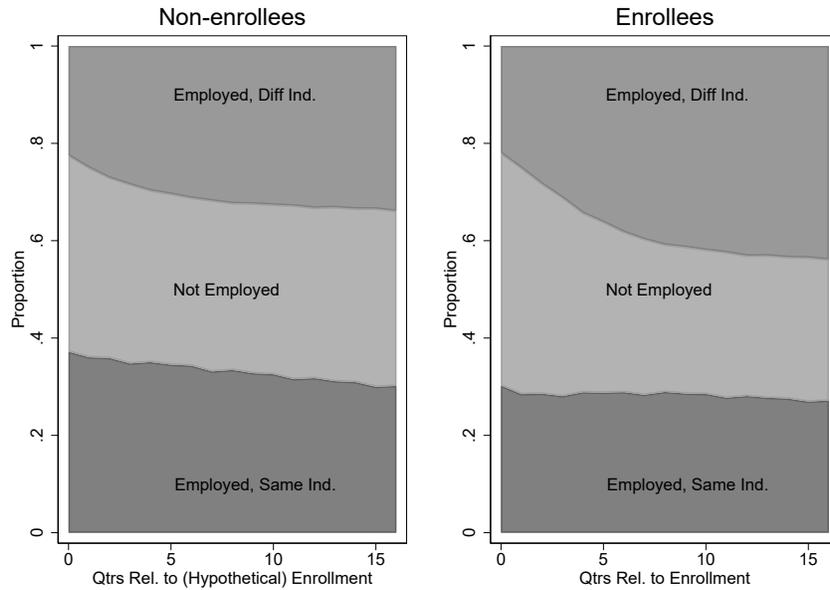
Notes: This graph shows the average quarterly earnings of enrollees who received WIA training services, and their matched non-enrollees. The solid vertical line denotes the quarter of first enrollment, and the vertical dashed lines denote eight and 16 quarters after first enrollment. $N = 19,530$ UI claims, corresponding to 19,387 unique individuals.

Figure 5: Long-run Earnings of Enrollees and Matched Non-enrollees

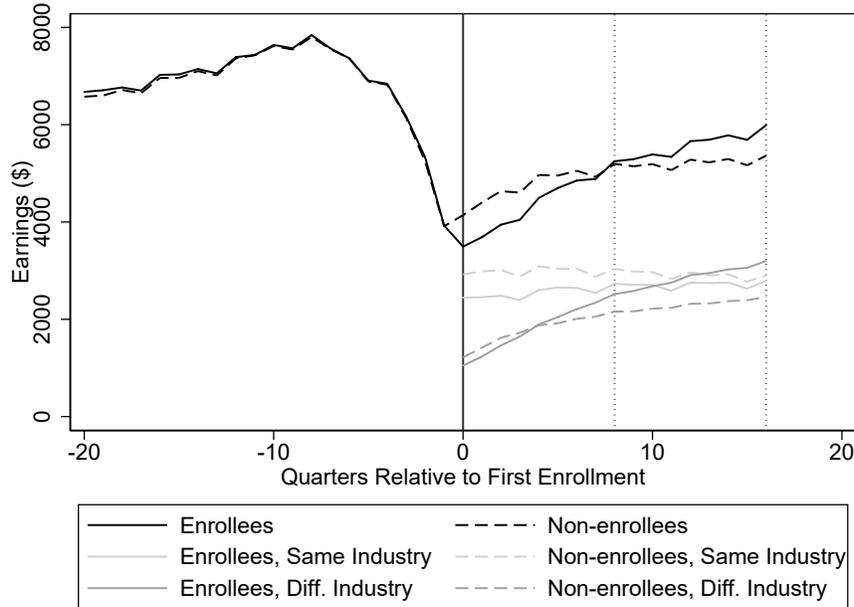


Notes: This graph shows the average quarterly earnings of UI claimants who enrolled from 2004 through the third quarter of 2007 and their matched non-enrollees. The solid vertical line denotes the quarter of first enrollment, and the vertical dashed lines denote eight and 16 quarters after first enrollment. $N = 39,372$ UI claims, corresponding to 38,340 unique individuals.

Figure 6: Industry Switching Among Enrollees and Matched Non-enrollees
 (A) Probability of Switching Industries Over Time



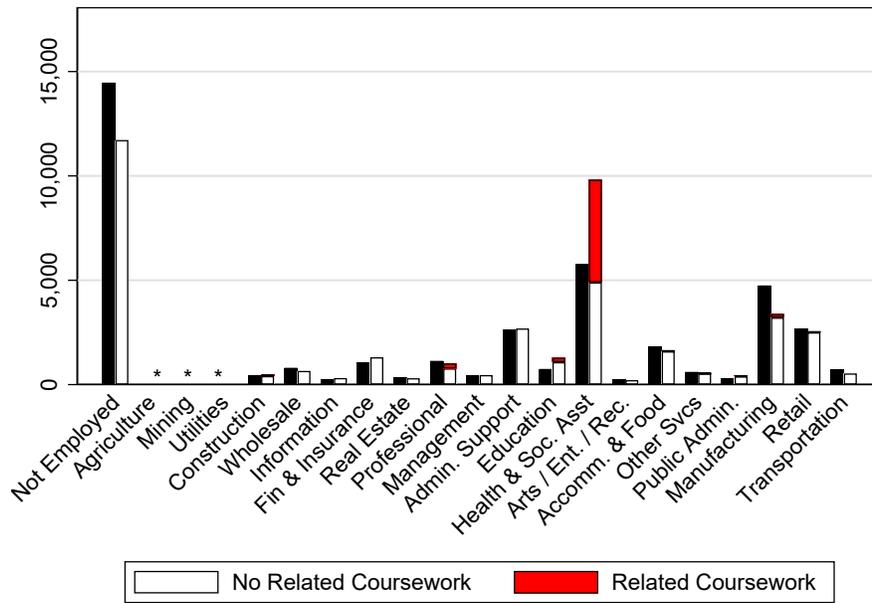
(B) Decomposition of Earnings: the Role of Industry Switching



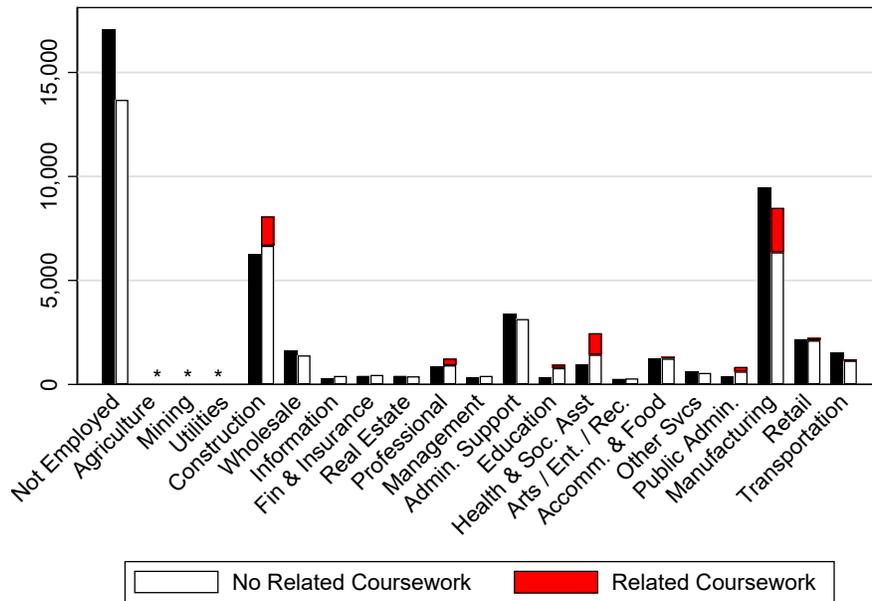
Notes: The figures in Panel A plot the probability of employment in the same (pre-layoff) industry, employment in a different industry, and non-employment over time for enrollees and matched non-enrollee UI claimants. Panel B plots the average quarterly earnings of enrollee and matched non-enrollee UI claimants (black solid and dashed lines). The gray lines disaggregate the post-enrollment earnings into two components: average quarterly earnings from the pre-layoff (“same”) industry and from a different industry, each scaled by the probability of employment in either the same or different industries. The solid (dashed) gray lines sum up to the solid (dashed) black lines. $N = 175,326$, corresponding to 167,422 unique individuals.

Figure 7: Industry-Related Course-Taking, by Industry of Employment at 16th Quarter Post-Enrollment

(A) Women



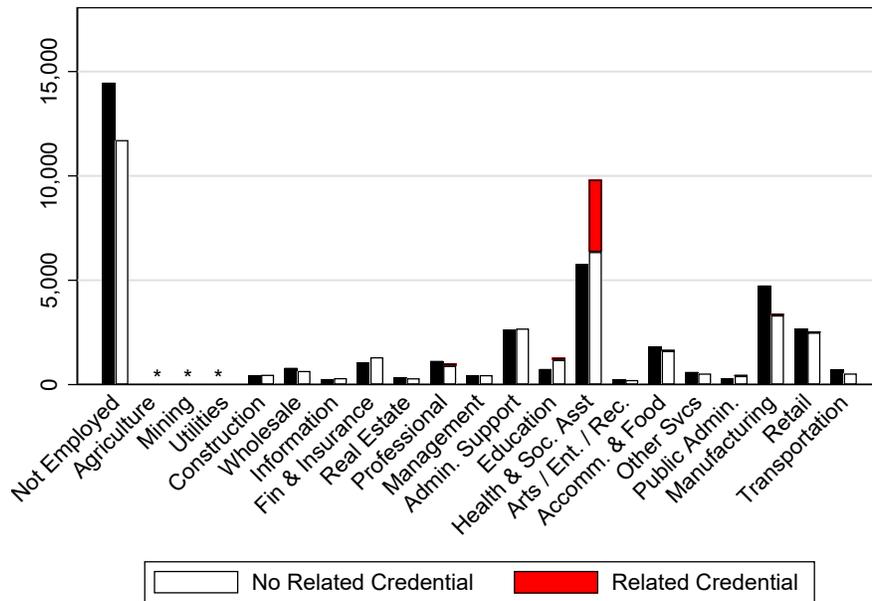
(B) Men



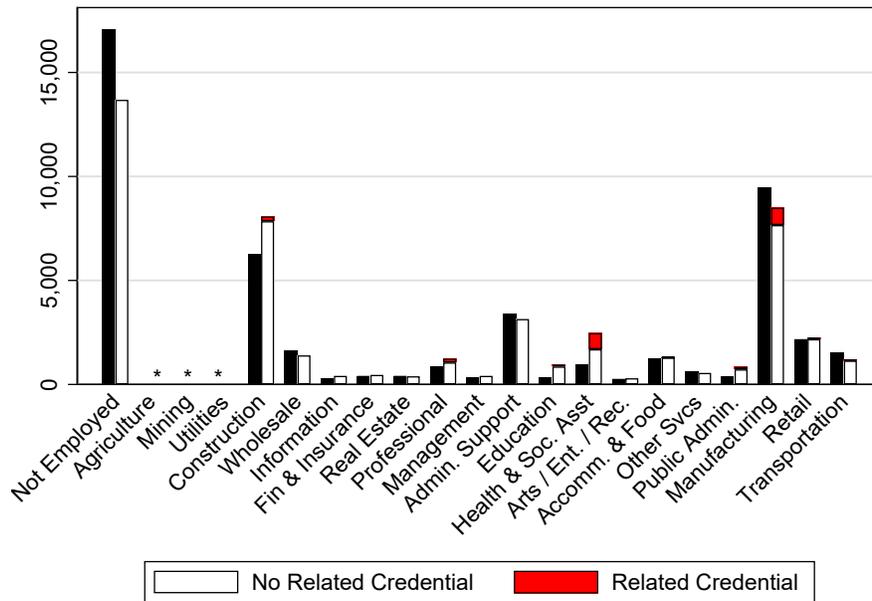
Notes: In each plot, the right side of each pair of bars shows the number of enrollees who take or do not take sector-related coursework. The left bar shows the number of matched non-enrollees in each sector. Agriculture, Mining, and Utilities sectors have fewer than 250 workers in each enrollee/non-enrollee cell and are not plotted. If there are 10 or fewer enrollees in a sector who take sector-related coursework, it is not plotted.

Figure 8: Industry-Related Credential Receipt, by Industry of Employment at 16th Quarter Post-Enrollment

(A) Women



(B) Men



Notes: In each plot, the right side of each pair of bars shows the number of enrollees who obtain or do not obtain sector-related credentials. The left bar shows the number of matched non-enrollees in each sector. Agriculture, Mining, and Utilities sectors have fewer than 250 workers in each enrollee/non-enrollee cell and are not plotted. If there are 10 or fewer enrollees in a sector who have sector-related credentials, it is not plotted.

Table 1: (A) Descriptive Characteristics of Enrollees and Non-enrollees

	Enrollees		Non-enrollees		Norm. Diff.
	Mean	SD	Mean	SD	
Female	0.45	0.50	0.33	0.47	0.24
Race					
White	0.75	0.43	0.81	0.40	-0.13
Black	0.18	0.38	0.12	0.33	0.14
Other	0.02	0.15	0.02	0.15	0.00
Unknown	0.05	0.22	0.05	0.21	0.02
Prior Industry					
Manufacturing	0.28	0.45	0.32	0.47	-0.08
Construction	0.15	0.36	0.17	0.37	-0.05
Admin. Support & Waste	0.12	0.32	0.11	0.31	0.03
Healthcare and Social Assistance	0.09	0.29	0.06	0.23	0.14
Retail Trade	0.08	0.27	0.07	0.26	0.04
Accommodation and Food Services	0.04	0.20	0.04	0.19	0.03
Wholesale Trade	0.04	0.19	0.04	0.20	-0.03
Transportation	0.03	0.18	0.05	0.21	-0.06
Tenure at Recent Employer					
<=1 year	0.35	0.48	0.27	0.45	0.16
>1 years to <=3 years	0.29	0.45	0.26	0.44	0.06
>3 years to <=6 years	0.16	0.37	0.18	0.38	-0.04
>6 years	0.20	0.40	0.29	0.45	-0.21
Age	35.46	10.74	42.33	11.79	-0.61
Cty Unempl. Rate at Layoff (%)	8.26	2.69	7.85	2.62	0.15
Earnings					
1 year before layoff	32690	21740	37202	27442	-0.18
2 years before layoff	29977	23297	35252	28708	-0.20
3 years before layoff	28302	24248	34111	29073	-0.22
Observations	87,760		1,894,145		

Notes: “Enrollees” (“Non-enrollees”) are UI claimants who enroll (do not enroll) in a public post-secondary institution in Ohio within two years of filing a UI claim. “SD” denotes standard deviation. “Norm. Diff.” is the normalized difference defined in Section 4.4.

Table 1: (B) Descriptive Characteristics of Enrollees and Matched Non-enrollees

	Enrollees		Matched Non-enrollees		Norm. Diff.	t-stat
	Mean	SD	Mean	SD		
	Female	0.45	0.50	0.45		
Race						
White	0.75	0.43	0.75	0.43	0.00	0.20
Black	0.18	0.38	0.18	0.38	0.00	-0.47
Other	0.02	0.15	0.02	0.15	0.00	-0.06
Unknown	0.05	0.22	0.05	0.22	0.00	0.47
Prior Industry						
Manufacturing	0.28	0.45	0.28	0.45	-	-
Construction	0.15	0.36	0.15	0.35	0.00	0.63
Admin. Support & Waste	0.12	0.32	0.12	0.32	0.00	0.43
Healthcare and Social Assistance	0.09	0.29	0.09	0.29	0.00	0.12
Retail Trade	0.08	0.27	0.08	0.27	0.00	-0.24
Accommodation and Food Services	0.04	0.20	0.04	0.20	0.00	-0.11
Wholesale Trade	0.04	0.19	0.04	0.19	0.00	-0.72
Transportation	0.03	0.18	0.03	0.18	0.00	-0.01
Tenure at Recent Employer						
<=1 year	0.35	0.48	0.35	0.48	0.00	0.55
>1 years to <=3 years	0.29	0.45	0.29	0.45	0.00	0.28
>3 years to <=6 years	0.16	0.37	0.16	0.37	0.00	0.17
>6 years	0.20	0.40	0.20	0.40	-0.01	-1.12
Age	35.46	10.74	35.43	10.76	0.00	0.51
Cty Unempl. Rate at Layoff (%)	8.26	2.68	8.26	2.71	0.00	-0.29
Earnings						
1 year before layoff	32694	21745	32595	27034	0.00	0.84
2 years before layoff	29978	23298	29862	28341	0.00	0.94
3 years before layoff	28305	24254	28181	28642	0.00	0.98
Observations	87,663		87,663			

Notes: “Enrollees” are UI claimants who enroll in a public post-secondary institution in Ohio within two years of filing a UI claim, and “Matched Non-enrollees” are their matched comparison group. “SD” denotes standard deviation. “Norm. Diff.” is the normalized difference defined in Section 4.4. “t-stat” is the t-statistic corresponding to the difference in means between enrollees and matched non-enrollees.

Table 2: Enrollment Characteristics

Time from job loss to enrollment (quarters)	3.7
Terms/Quarters Enrolled	4.7
Type of Institution Attended (%)	
Technical Center	13.0
Community College	74.5
Four-year Institution	19.5
Types of Courses	
Taken at least one occupational course (%)	66.8
Avg proportion of courses occupational	0.6
Credential (%)	
Graduate / Professional	1.1
Bachelors	2.4
Associate	10.0
Less Than Associate	12.5
Observations	87,663

Notes: Type of Institution Attended, Terms/Quarters Enrolled, Types of Courses, and Credential are calculated within four years of first enrollment. Enrollees may attend more than one type of institution over the four-year period. “Less than Associate” credentials include less than two-year awards from HEI and any credential from OTC.

Table 3: Enrollment Effect Estimates

	TOT Estimates (1)	Matched Non-enrollee Mean (2)
Post-Enrollment Quarterly Earnings		
1-2 Yrs Post-Enrollment	-359.79 (18.70)	4841.75
3-4 Yrs Post-Enrollment	387.01 (20.90)	5216.83
Pre-Layoff Quarterly Earnings		
1 Year Before Layoff	24.74 (21.98)	8148.71
2 Years Before Layoff	29.07 (22.79)	7465.41
Observations	175326	87663

Notes: Column (1) shows the estimated effect of enrollment on earnings in the time period denoted by the row headings. Column (2) shows the mean earnings in the matched non-enrollee sample. Abadie and Imbens (2016) standard errors are in parentheses.

Table 4: Enrollment Effect Estimates By Subgroup

Subgroup	Quarterly Earnings, 1-2 Yrs Post-Enrollment				Quarterly Earnings, 3-4 Yrs Post-Enrollment				No. of Enrollees (11)		
	Simple Diff.	Diff-Diff	Nonenr. Mean	Pct. Ch. Diff.	Simple Diff.	Diff-Diff	Nonenr. Mean	Pct. Ch. Diff.			
	TOT Estimate (1)	TOT Estimate (2)	Mean (3)	Diff. (4)	DD (5)	TOT Estimate (6)	TOT Estimate (7)	Diff. (9)		DD (10)	
All	-360 (23)	-397 (23)	4842	-7%	-8%	387 (26)	368 (29)	5217	7%	7%	87663
Quarters From Layoff to Enrollment											
1	-786 (58)	-908 (64)	5337	-15%	-17%	119 (64)	66 (72)	5697	2%	1%	14519
2	-728 (52)	-767 (56)	5005	-15%	-15%	171 (59)	117 (68)	5457	3%	2%	17021
3	-393 (54)	-361 (58)	4682	-8%	-8%	382 (62)	436 (74)	5149	7%	8%	14440
4	-253 (63)	-298 (64)	4725	-5%	-6%	409 (70)	400 (81)	5115	8%	8%	11610
5	-123 (69)	-136 (65)	4673	-3%	-3%	577 (77)	526 (93)	5009	12%	10%	9395
6	75 (76)	8 (67)	4686	2%	0%	674 (84)	664 (94)	4923	14%	13%	8102
7	60 (81)	42 (68)	4534	1%	1%	603 (90)	632 (94)	4843	12%	13%	6947
8	180 (97)	182 (81)	4604	4%	4%	704 (103)	650 (105)	4866	14%	13%	5629
Layoff Year											
2004	-437 (76)	-516 (67)	5642	-8%	-9%	318 (81)	331 (88)	5186	6%	6%	7603
2005	-97 (94)	-195 (81)	5856	-2%	-3%	452 (93)	280 (103)	4836	9%	6%	5783
2006	-517 (79)	-440 (85)	5441	-10%	-8%	48 (83)	141 (95)	4977	1%	3%	8395
2007	-403 (67)	-471 (73)	4392	-9%	-11%	103 (76)	61 (89)	4790	2%	1%	9137
2008	-464 (50)	-515 (53)	4282	-11%	-12%	425 (58)	365 (65)	5042	8%	7%	17067
2009	-422 (48)	-427 (49)	4794	-9%	-9%	444 (55)	437 (62)	5578	8%	8%	21106
2010	-175 (64)	-234 (63)	4693	-4%	-5%	551 (71)	563 (84)	5408	10%	10%	11872
2011	-108 (81)	-178 (79)	4759	-2%	-4%	654 (94)	636 (109)	5434	12%	12%	6700
Male	-190 (35)	-235 (34)	5657	-3%	-4%	574 (39)	545 (44)	6077	9%	9%	48168
Female	-567 (28)	-594 (31)	3848	-15%	-15%	159 (31)	153 (37)	4168	4%	4%	39495
Manufacturing	-969 (49)	-1008 (54)	5919	-16%	-17%	217 (52)	195 (60)	6098	4%	3%	24914
Non-manuf.	-118 (26)	-154 (24)	4414	-3%	-3%	455 (29)	437 (33)	4867	9%	9%	62749
Age <40	-368 (25)	-404 (24)	4556	-8%	-9%	364 (29)	258 (31)	5032	7%	5%	57809
Age >=40	-363 (46)	-390 (48)	5410	-7%	-7%	428 (51)	424 (59)	5574	8%	8%	29925
Tenure											
<=1 Year	-155 (33)	-201 (33)	4038	-4%	-5%	427 (38)	367 (43)	4405	10%	8%	30487
1-6 Years	-428 (34)	-358 (33)	4769	-9%	-8%	355 (38)	407 (43)	5243	7%	8%	39527
>6 Years	-511 (61)	-681 (66)	6334	-8%	-11%	397 (66)	284 (76)	6542	6%	4%	17720
White	-439 (27)	-466 (28)	5150	-9%	-9%	399 (31)	373 (35)	5543	7%	7%	65925
Black	-176 (44)	-203 (47)	3726	-5%	-5%	272 (48)	137 (56)	4021	7%	3%	15407

Notes: Columns (1) and (2) show the estimated effect of enrollment on quarterly earnings in the two years after first enrollment, computed by taking the difference between enrollees and matched non-enrollees, and by matched difference-in-differences, respectively, for each subgroup denoted by the row headings. Columns (6) and (7) show the analogous results for the third and fourth years after enrolling. Columns (3) and (8) show the mean quarterly earnings of the matched non-enrollees. Columns (4), (5), (9), and (10) express the columns (1), (2), (6), and (7) as percentages of the matched non-enrollee mean. Standard errors for the mean pairwise difference between enrollees and matched non-enrollees are reported in parentheses.

Appendix (For Online Publication Only)

A Data Sources and Sample Construction: Additional Details

A.1 Construction of Analysis Sample and Variables

Our unemployment insurance (UI) claims sample consists of eligible, regular UI claims (i.e., claims that are filed when a worker is first unemployed). To create the sample, we first construct UI spells by grouping together regular claims and all associated extension claims (e.g., benefits under the Extended Benefit, Emergency Unemployment Compensation, and Trade Readjustment Allowance programs), which have the same benefit year beginning date. We then check for cases where an individual has two overlapping UI spells (i.e., one spell begins before the previous one ends), which we further group together as part of the same UI spell. These steps help ensure that the beginning of each spell corresponds to when a worker is first laid off, rather than a continuation of an ongoing unemployment spell. Out of the approximately two million claims, we drop a tiny fraction with missing gender (5,456 observations), or if the age reported is under 16 or over 100 (88 observations). For the small proportion of observations that do not have an in-state zip code, we assign them the state-wide unemployment rate in the matching analysis.

Several of our key variables come from the quarterly wage database that is part of the Ohio UI system. Our main outcome measures are quarterly earnings. Although the earnings data we receive are topcoded, the censoring points are sufficiently high so as to be not very relevant for our sample (as is evident from the summary statistics of Table 1): prior to 2009, earnings above \$99,999 were censored at \$99,999; for all quarters starting in 2009, the top one percent of earnings have been topcoded to the average of the top one percent. The wage data also contain employer pseudo-IDs and industry, which allow us to construct job tenure measures. Since these employer pseudo-IDs are not in the UI claims data, we first search in the wage data for an employer that matches the industry reported in the UI claims data in the five quarters before layoff, starting with the most recent quarter. Once we have identified the employer that matches the industry in the claims data, we define tenure as the number of quarters since the first time a worker is observed to have worked for the employer in the wage data. We then create three tenure categories: less than one year, one to six years, and more than six years (since we only observe wages starting in 1995Q2, the maximum tenure for the earliest claims in the sample is 8.75 years). If we do not find an employer with a matching industry or if the industry is missing in the claims data, we report the tenure at the most recent employer

in the wage data. If the industry is missing in the claims data, we fill in the industry of the most recent employer from the wage data. Only a small fraction of the wage data have missing industry (0.5 percent).

Since we only observe earnings within Ohio, it is possible that we are understating the earnings of workers who work out-of-state. To gauge the extent of this issue, we examine the out-migration rates of workers who claim UI in the Survey of Income and Program Participation (SIPP). In the 2004 (2008) SIPP panels, there were 139 (261) individuals who reported receiving UI benefits while residing in Ohio at some point during the four years of the panel. Of these individuals, 8 (4) individuals moved out-of-state, and 3 (2) found jobs in their new destination in the 2004 (2008) panels. Although the samples of UI recipients are small in the SIPP, these rates of out-migration are similar to the overall yearly migration rates for Ohio found using IRS Statistics of Income data (2005-2016) and the American Community Survey (2010-2016), which is about two percent. To get a sense for how out-migration may affect our main estimates, consider the extreme scenario where two percent of the non-enrollee control group have no in-state earnings due to out-migration and that *none* of the enrollee treatment group has migrated. Then replacing the zero earnings of the bottom two percent of the control group with their mean earnings will increase the average control earnings by about \$104 (two percent times \$5,217 as reported in Table 3) three to four years post-enrollment, resulting in a treatment effect estimate of \$283 per quarter. Even in this very conservative scenario, there is a positive return to enrollment of more than five percent.

A.2 Enrollment, Course, and Credential Data

Our enrollment data come from two sources: 1) The Higher Education Information (HEI) data, which cover two- and four- year public colleges, and 2) the Ohio Technical Center (OTC) data, which contain information on technical centers (the OTC data also include training that takes place in correctional facilities and high schools, though these account for less than one percent of our enrollee sample). The HEI data, available from the summer of 1999 to the spring of 2017, are all reported at the person-institution-term level so we can observe enrollment, courses taken, and credentials/degrees obtained in every term for each individual. Terms are “Winter”, “Spring”, “Summer”, and “Autumn”, which we map to the first, second, third, and fourth quarters of the calendar year in our analysis. Starting in 2013, “Winter” terms were eliminated as all public colleges moved to a semester system. The OTC data, spanning years between 2002 and 2017, contain start and end dates for each course, and credentials are often associated with specific courses.

Each course and credential has a Classification of Instructional Programs (CIP) code that denotes the

subject area. In the HEI data, 0.3 percent of courses taken by enrollees in our sample are missing CIP codes (none of the credentials are missing CIP codes). In the OTC data, although we see that 9.8 percent of courses in our enrollee sample have missing CIP codes, there is another variable containing an internal subject code that we can use to fill in most of the missing CIP codes. That is, for courses that have missing CIP codes but non-missing internal subject codes, we fill in the most common CIP code associated with that internal subject code. This procedure effectively reduces the percentage of OTC courses with missing CIP codes to 0.7 percent. Also, the HEI course information for 2006 appears to be incomplete relative to the enrollment data: 17 percent of the enrollee-by-term observations in 2006 are not associated with courses. This missing data issue only affects results in Section 5.5 and may slightly undercount the number of enrollees that take coursework related to specific industries or areas.

To categorize courses and credentials to specific (two-digit) industries, we create a mapping of CIP codes to industries using the data from the National Center of Education Statistics (NCES) and the Bureau of Labor Statistics (BLS). The NCES provides a list of occupations that each CIP subject prepares for, while the National Employment Matrix (2018) from the BLS has data on the share of workers of a particular occupation in a specific industry (National Center for Education Statistics, 2011; Bureau of Labor Statistics). To create this mapping, we restrict to the set of “occupational” (as opposed to “academic”) CIP codes, as defined by the NCES. We also remove “post-secondary teacher” occupations (SOC 25-1000) because many subjects list educators of that subject as a possible occupation, and these tend to be categorized as post-secondary teachers, most of whom work in the Educational Services sector. Since we believe most workers who train in a particular subject are not aiming to teach in the area, we are eliminating the likelihood of falsely attributing a subject to the education sector. We then order the occupations within CIP subject by the total employment of the occupation according to the National Employment Matrix, and map each CIP code to the most populous occupation. We further map a CIP code to a two-digit NAICS industry if, for the occupation the CIP maps to, more than a quarter of the employment is within that industry. We use the largest industry if there are multiple industries that meet this criterion. We designate a CIP code as not associated with an industry if 1) the largest industry accounts for less than 25 percent of the occupation the CIP maps to, 2) the CIP is not associated with any occupation, or 3) the mapped “industry” is self-employment (i.e., more workers in the mapped occupation are self-employed than working in any particular sector). An example of a course not mapped to an industry is Human Resources Management (CIP 52.1001): although it prepares a worker for the Human Resource Specialist occupation, human resource specialists work in many different

industries and cannot be assigned to a specific one.

Courses and credentials in the OTC data are mapped to industries via their CIP code. When a worker takes more than one course, we use the course with the most course hours (and randomly pick a course with the same largest number of course hours). In the HEI data, where each enrollee typically takes more than one course, we first map each of her courses to an industry and then assign to her coursework the modal industry over all her courses (for this assignment, we require that she take at least three courses in that industry). When we observe a worker earn more than one credential, we map her credentials to the modal (non-missing) industry over all her credentials, and in the case of ties, we keep the industry mapped from the highest credential. When a worker is observed in both OTC and HEI data, and the industries of her coursework do not match, we use the one in the OTC data (because courses in OTC are more easily matched to an industry as there are fewer courses); if the industries of the credential do not match, we use the one in the HEI data (because institutions in the HEI data tend to confer higher credentials).

A.3 Workforce Investment Act Data

To analyze the returns to enrollment for workers who trained under the Workforce Investment Act (WIA) program, our analysis sample has been merged with the WIA administrative program data (from the WIA Standardized Record Data system). The data contain quarterly snapshots of WIA participants and exiters between 2006Q1 and 2015Q4. The earliest snapshot (2006Q1) contains participants who exited the program starting in 2004Q1. The analysis sample in Section 5.3 contains only WIA participants who received job training from WIA. We observe the month of WIA registration, the beginning and end months of WIA training, program funding stream (e.g., adult, dislocated worker, or youth), and type of training (e.g., on-the-job, skill upgrading, entrepreneurial skills). We say that an enrollee from our main analysis sample is a trainee in the WIA program if we observe her starting WIA training within 24 months after the UI claim date. We only observe WIA participation for workers in our main analysis sample of UI claims.

B Generalized Identification Results on Enrollment Effects and Proofs

In this section, we provide additional details to Section 4.1 with generalized results and proofs. First, we prove the statements in equations (3) and (4). We then state and prove an identification result based on propensity score matching for the two-period case ($S = 2$). Finally, we generalize the assumptions and

identifications to the dynamic context with an arbitrary S .

Proof of equations (3) and (4)

Proof. Equation (4) directly follows from Assumption 1 and the overlap condition $\Pr(T^2 = 1|T^1 = 0, X^2 = x^2) < 1$, and the proof is similar to that of the static identification result (2).

For equation (3), we apply the law of iterated expectations and write the LHS of (3) as

$$\begin{aligned}
& E[Y|T^1 = 1] - E[E[Y|T^1 = 0, T^2 = 0, X^1]|T^1 = 1] \\
&= E[E[Y|T^1 = 1, X^1]|T^1 = 1] - E[E[Y|T^1 = 0, T^2 = 0, X^1]|T^1 = 1] \\
&= E[E[Y|T^1 = 1, X^1] - E[Y|T^1 = 0, T^2 = 0, X^1]|T^1 = 1]. \tag{A1}
\end{aligned}$$

Focusing on the quantity inside the outer conditional expectation operator, we have

$$\begin{aligned}
& E[Y|T^1 = 1, X^1] - E[Y|T^1 = 0, T^2 = 0, X^1] \\
&= E[Y(1)|T^1 = 1, X^1] - E[Y(0)|T^1 = 0, X^1] - \{E[Y|T^1 = 0, T^2 = 0, X^1] - E[Y(0)|T^1 = 0, X^1]\} \\
&= E[Y(1) - Y(0)|X^1] - \{E[Y|T^1 = 0, T^2 = 0, X^1] - E[Y(0)|T^1 = 0, X^1]\}, \tag{A2}
\end{aligned}$$

where Assumption 1 implies the last equality. Assumption 2 implies that the expression inside the curly bracket is zero:

$$\begin{aligned}
& E[Y|T^1 = 0, T^2 = 0, X^1] - E[Y(0)|T^1 = 0, X^1] \\
&= E[Y|T^1 = 0, T^2 = 0, X^1] - E[Y(0)|T^1 = 0, T^2 = 0, X^1] \Pr(T^2 = 0|T^1 = 0, X^1) \\
&\quad - E[Y(0)|T^1 = 0, T^2 = 1, X^1] \Pr(T^2 = 1|T^1 = 0, X^1) \\
&= \{E[Y(0)|T^1 = 0, T^2 = 0, X^1] - E[Y(0)|T^1 = 0, T^2 = 1, X^1]\} \cdot \Pr(T^2 = 1|T^1 = 0, X^1) = 0 \tag{A3}
\end{aligned}$$

The identification result (3) follows from the combination of equations (A1), (A2), and (A3). \square

Identification of the TOT parameters via propensity score matching: 2-period case

We begin by defining the dynamic propensity scores:

$$\begin{aligned}
p^1(X^1) &\equiv \Pr(T^1 = 1|T^2 = 0, X^1) \\
p^2(X^2) &\equiv \Pr(T^2 = 1|T^1 = 0, X^2),
\end{aligned}$$

for which the conditioning set is information available at treatment variable realization for workers who do not enroll in any other period.

The following lemma establishes the dynamic propensity score theorem needed for identification:

Lemma 1. *Under Assumptions 1 and 2,*

$$Y(0) \perp\!\!\!\perp T^1 | T^2 = 0, p^1(X^1) \quad (\text{A4})$$

$$Y(0) \perp\!\!\!\perp T^2 | T^1 = 0, p^2(X^2) \quad (\text{A5})$$

Proof. Since statement (A5) follows directly from the second part of Assumption 1 by the same argument in Rosenbaum and Rubin (1983), we prove statement (A4). First, we prove that Assumptions 1 and 2 imply

$$Y(0) \perp\!\!\!\perp T^1 | X^1, T^2 = 0 \quad (\text{A6})$$

by showing

$$\Pr(T^1 = 0 | Y(0), X^1, T^2 = 0) = \Pr(T^1 = 0 | X^1, T^2 = 0). \quad (\text{A7})$$

We write

$$\begin{aligned} & \Pr(T^1 = 0 | Y(0), X^1, T^2 = 0) \\ & \stackrel{\text{i}}{=} \frac{\Pr(T^2 = 0 | Y(0), X^1, T^1 = 0) \Pr(T^1 = 0 | Y(0), X^1)}{\sum_{t=0}^1 \Pr(T^2 = 0 | Y(0), X^1, T^1 = t) \Pr(T^1 = t | Y(0), X^1)} \\ & \stackrel{\text{ii}}{=} \frac{\Pr(T^2 = 0 | X^1, T^1 = 0) \Pr(T^1 = 0 | X^1)}{\sum_{t=0}^1 \Pr(T^2 = 0 | Y(0), X^1, T^1 = t) \Pr(T^1 = t | Y(0), X^1)} \end{aligned} \quad (\text{A8})$$

where equality i follows from Bayes' Rule and equality ii follows from Assumptions 1 and 2. Note that the denominator of (A8) is

$$\begin{aligned} & \Pr(T^2 = 0 | Y(0), X^1, T^1 = 0) \Pr(T^1 = 0 | Y(0), X^1) + \Pr(T^2 = 0 | Y(0), X^1, T^1 = 1) \Pr(T^1 = 1 | Y(0), X^1) \\ & \stackrel{\text{iii}}{=} \Pr(T^2 = 0 | X^1, T^1 = 0) \Pr(T^1 = 0 | X^1) + \Pr(T^1 = 1 | X^1) \\ & \stackrel{\text{iv}}{=} \Pr(T^1 = 0, T^2 = 0 | X^1) + \Pr(T^1 = 1, T^2 = 0 | X^1) \\ & = \Pr(T^2 = 0 | X^1) \end{aligned} \quad (\text{A9})$$

where equalities iii and iv hold because $T^1 = 1$ implies that $T^2 = 0$ and by Assumptions 1 and 2. Equation (A7) follows from the combination of (A8) and (A9), and it implies the dynamic conditional independence statement (A6).

With (A6), we can establish the dynamic propensity score statement (A4). The proof is analogous to that in Rosenbaum and Rubin (1983) of the static propensity score theorem and is omitted here for brevity. \square

We now state the propensity score based identification results for the 2-period case.

Proposition 1. Under Assumptions 1 and 2 and provided that $p^1(X^1), p^2(X^2) < 1$,

(a):

$$E[Y|T^1 = 1] - E[E[Y|T^1 = 0, T^2 = 0, p^1(X^1)]|T^1 = 1] = \underbrace{E[Y(1) - Y(0)|T^1 = 1]}_{TOT \text{ for } T^1=1} \quad (\text{A10})$$

$$E[Y|T^2 = 1] - E[E[Y|T^1 = 0, T^2 = 0, p^2(X^2)]|T^2 = 1] = \underbrace{E[Y(1) - Y(0)|T^2 = 1]}_{TOT \text{ for } T^2=1}; \quad (\text{A11})$$

(b):

$$\sum_{s=1}^2 \{E[Y|T^s = 1] - E[E[Y|T^1 = 0, T^2 = 0, p^s(X^s)]|T^s = 1]\} \Pr(T^s = 1|T = 1) = \underbrace{E[Y(1) - Y(0)|T = 1]}_{\text{overall } TOT}.$$

Proof. Since $T^1 = 1 \Rightarrow T^2 = 0$, and by law of iterated expectations and Lemma 1,

$$\begin{aligned} & E[Y|T^1 = 1] - E[E[Y|T^1 = 0, T^2 = 0, p^1(X^1)]|T^1 = 1] \\ &= E[E[Y(1)|T^1 = 1, T^2 = 0, p^1(X^1)]|T^1 = 1] - E[E[Y(0)|T^1 = 0, T^2 = 0, p^1(X^1)]|T^1 = 1] \\ &= E[E[Y(1) - Y(0)|T^1 = 1, T^2 = 0, p^1(X^1)]|T^1 = 1] \\ &= E[Y(1) - Y(0)|T^1 = 1], \end{aligned}$$

which proves equation (A10) in part (a). The proof of equation (A11) is analogous.

Because $T = T^1 + T^2$, part (b) follows easily from part (a) and the identity

$$E[Y(1) - Y(0)|T = 1] = \sum_{s=1}^2 E[Y(1) - Y(0)|T^s = 1] \Pr(T^s = 1|T = 1).$$

□

Identification of the TOT parameters via propensity score matching: S -period case

Now we generalize Assumptions 1 and 2 and Proposition 1 from two periods to S periods. First, define the period- s propensity score as

$$p^s(X^s) \equiv \Pr(T^s = 1 | T^{s'} = 0 \text{ for all } s' \neq s, X^s)$$

for $s = 1, \dots, S$.

Assumption 3. $Y(0) \perp\!\!\!\perp T^1 | X^1$ and $Y(0) \perp\!\!\!\perp T^s | T^{s-1} = \dots = T^1 = 0, X^s$ for $s = 2, \dots, S$

Assumption 4. $Y(0) \perp\!\!\!\perp T^r | T^{r-1} = \dots = T^1 = 0, X^s$ for all $s < r$ and $r = 2, \dots, S$

Proposition 2. Under Assumptions 3 and 4 and provided that $p^s(X^s) < 1$ for all s , we have

(a): for $s = 1, \dots, S$

$$E[Y|T^s = 1] - E[E[Y|T^s = 0 \forall s, p^s(X^s)]|T^s = 1] = \underbrace{E[Y(1) - Y(0)|T^s = 1]}_{TOT \text{ for } T^s=1};$$

(b):

$$\sum_{s=1}^S \{E[Y|T^s = 1] - E[E[Y|T^s = 0 \forall s, p^s(X^s)]|T^s = 1]\} \Pr(T^s = 1|T = 1) = \underbrace{E[Y(1) - Y(0)|T = 1]}_{\text{overall } TOT}.$$

Similar to the proof of Lemma 1, we can show that Assumptions 3 and 4 imply dynamic independence conditional on the propensity score:

$$Y(0) \perp\!\!\!\perp T^s | T^{s'} = 0 \text{ for all } s' \neq s, p^s(X^s)$$

for $s = 1, \dots, S$, from which Proposition 2 follows easily.

C Validation Exercise Using the National Job Training Partnership Act Study

C.1 The National Job Training Partnership Act Study

The National Job Training Partnership Act Study (NJS) was an experimental evaluation of training programs and services that were funded under Title II of the Job Training Partnership Act (JTPA) of 1982, a predecessor of WIA. The NJS was conducted in 16 locations across the U.S. from 1987 to 1989. Over 20,000 applicants were randomly assigned to a treatment group with access to JTPA services or a control group barred from these services for 18 months. The treatment and control group members were followed for 12 to 37 months post randomization (21 months, on average).

As mentioned in Section 4.2, a unique aspect of the NJS is that the researchers also collected information for an “eligible non-participant” (ENP) group from four of the sixteen experimental sites, which consists of workers who met the income eligibility requirement but did not apply to JTPA at the time of randomization. Importantly, the same background and outcome data were obtained for the ENP and the experimental control groups. By comparing the pre- and post- program characteristics of the ENP and experimental control groups, researchers can study the selection process into training programs and evaluate the performance of alternative causal inference methods, in the spirit of LaLonde (1986). The main advantage of the NJS data over the National Supported Work (NSW) data used in LaLonde (1986) is that the NJS data contain a much richer set of information on workers’ labor market history both before and after randomization. While the NSW data only contain earnings at the annual frequency two years before and one year after the intervention,

the NJS data contain earnings at the monthly frequency more than three years before and approximately 18 months after randomization.¹

Heckman, Ichimura and Todd (1997) (hereafter, HIT), Heckman et al. (1998) (hereafter, HIST), and Heckman and Smith (1999) (hereafter, HS) have used the NJS data to evaluate econometric methods widely applied in the 1980s and 1990s. In this section, we extend their analyses in three ways. First, we examine the performance of non-experimental methods on a subset of recently employed workers that is closer to our population of interest. Second, we expand upon the set of methods evaluated by HIST by incorporating recent machine learning based matching approaches, such as forward stepwise selection, penalized-logit propensity score matching, genetic matching (GenMatch), classification tree based methods such as bagging, random forest and boosting, and support vector machines (SVM). Third, we assess variants of the fixed effects earnings regression specifications proposed by Jacobson, Lalonde and Sullivan (2005a).

C.2 Sample Construction

Our analysis sample contains the experimental controls and the ENPs. The experimental controls were randomized out of the program and lived in the same geographical areas as the ENPs. Since neither group received treatment of JTPA training, the true causal effect is zero. An estimate of the “treatment effect” very different from zero reflects differences between those who applied for training and those who did not. Since the focus of our paper is the impact of education for unemployed workers who have some recent work experience, we only use the adult (i.e., ages 22-54) samples in the validation study, as many in the youth samples are new labor market entrants.

We restrict our sample to workers who have at least one month of non-missing earnings (missing earnings are due to item nonresponse and are different from zero earnings) for each of the 12 quarters before and 6 quarters after the date of random assignment (for experimental controls) or eligibility determination (for ENPs). Workers in our sample must also have non-missing earnings during the month of random assignment or eligibility determination. We coarsen the monthly data to quarters for two reasons. First, the earnings data in our main study are quarterly. Second, this balances the tension between keeping our sample as large as possible while keeping only workers with complete earnings histories. We further restrict the sample to workers who had no missing demographic information (race, ethnicity, age, sex, geographic loca-

¹Specifically, the experimental controls and ENPs were given a survey detailing the start and end dates of each job, wage, and hours worked, which were then converted to monthly earnings.

tion, marital status, household size, the presence of children under six, and educational attainment), which comes from the “Long Baseline Survey” administered to both ENPs and experimental controls. We drop 11 observations with monthly earnings exceeding \$10,000 in any given month, which are most likely due to reporting error. The resulting sample contains 770 adult men and 1,375 adult women. The first four columns of Appendix Table A.3 present summary statistics of each sample.

Since our Ohio analysis focuses on a population of recently unemployed workers, we create two additional NJS subsamples with recent employment. Specifically, we define the “earnings history” subsamples as those in the adult men and adult women groups who have at least four quarters of positive earnings within the previous 12 quarters, which contain 635 men and 791 women respectively. Restricting workers to have positive earnings histories greatly reduces the sample of women, as many women were out of the labor force before joining the study. As expected to be the case mechanically, the earnings history subsamples have higher average earnings in the pre-period than their full sample counterparts (24 and 55 percent higher for the male and female experimental controls, respectively; see Appendix Table A.3). As a result, while the earnings of the NJS experimental controls in the earnings history samples still had lower earnings than our Ohio enrollees (about 195 and 115 percent of the 1986 individual federal poverty line for an average male and female NJS experimental control compared to 260 percent of the 2009 individual federal poverty line for an Ohio enrollee), the earnings history samples better resemble our target population.

Appendix Figure A.18 shows the earnings trajectories of the full men and women samples (upper panels) as well as the earnings history subsamples (lower panels). In all four samples, the experimental controls exhibit a drop in earnings prior to random assignment while the ENPs do not, indicating selection into the program on transitory earnings declines as pointed out by HIST. Furthermore, in both samples for men, the ENPs had higher average earnings prior to the study period than the experimental controls. In the full women sample, the two groups had similar earnings, but consistent with the pattern of their male counterparts, the female ENPs had higher pre-period income in the earnings histories subsample.

C.3 Selection-On-Observable Estimators Evaluated

We first use the NJS data to evaluate a series of matching estimators. The estimators include conventional methods such as ordinary least squares (OLS), logit-based propensity score matching (L-PSM), Mahalanobis matching, and inverse probability weighting, as well as machine-learning type matching methods that have attracted recent attention (see Westreich, Lessler and Funk, 2010 for a review).

While some of the estimators have been extensively studied, others may be unfamiliar to empirical economists. Here, we briefly review the construction of conventional estimators and introduce the more novel methods. For the OLS estimator, we use the Stata `regress` command with heteroskedasticity robust standard errors. For all estimators involving actual matching, we conduct nearest neighbor matching with one and five neighbors. For L-PSM, Mahalanobis matching, and inverse probability weighting, we use the Stata `teffects` command with default standard error estimators. For other propensity score or distance calculations, we state the R software package used for implementation, in which we employ the “default” options whenever possible. All matching estimates using R packages are obtained using the `Match` command of the `Matching` package (Sekhon, 2011), with Abadie and Imbens (2006) standard errors.

- OLS: Although the OLS estimator is in general not consistent for the TOT parameter, we include it for completeness given its familiarity. The OLS estimator comes from a regression of future earnings on the “treatment” indicator (i.e. whether an observation was in the experimental control group) controlling for all variables in each covariate set linearly.
- L-PSM: We estimate propensity score using a logit regression of treatment on each covariate set linearly.
 - Difference-in-Differences Matching (D-in-D Matching): We follow HIT and HIST and estimate the treatment effect by measuring the difference-in-differences in earnings for matched treatment and control pairs (based on the logit propensity score), before and after random assignment/eligibility determination.
- Mahalanobis Matching: Instead of estimating the propensity score, we define the distance between observations i and j with corresponding covariate vectors X_i and X_j as $M(X_i, X_j) = \sqrt{(X_i - X_j)'V^{-1}(X_i - X_j)}$ where $V = \text{var}(X)$. We match observations that are closest in terms of distance M .
 - Inverse Variance Matching: We also implement inverse variance matching suggested by Imbens (2004), which replaces V above with $\text{diag}(V)$, so that the distance between X_i and X_j is $IVM(X_i, X_j) = \sqrt{(X_i - X_j)'\text{diag}(V)^{-1}(X_i - X_j)}$.
- Inverse Probability Weighting: We use the logit-based propensity score to re-weight ENP observations. The estimator is a sample analogue of the Horvitz-Thompson TOT estimand $E[Y|T = 1] - E\left(E\left[\frac{p(X)Y}{1-p(X)}|T = 0, X\right]\right)$, where $p(X)$ is the propensity score and T indicates treatment.

- L-PSM with covariate selection: We incorporate quadratic and interaction terms of covariates as potential inputs into the logit-based propensity score model. Since the number of these additional terms can be large, we use two methods to discipline the selection of relevant first and second-order terms within each covariate set into the propensity score model.
 - Forward stepwise covariate selection algorithm per Imbens and Rubin (2015): Starting with no covariates, we begin the algorithm by considering linear terms to include in the propensity score model. We estimate the model using each of the covariates and test whether the added covariate matters using a likelihood ratio test. If the largest likelihood ratio statistic across all covariates is less than $C_l = 1$ (a tuning parameter recommended by Imbens and Rubin, 2015), no additional covariates are added. Otherwise, the covariate with the largest likelihood ratio statistic is included. The linear terms are added one at a time in this manner. Quadratic and interaction terms are similarly added, but the threshold likelihood ratio value is $C_q = 2.71$.
 - Penalized-logit: A penalized logit is similar to a standard logistic regression, but regularizes the coefficients with an L1 penalty commonly used in LASSO regressions, which effectively selects a subset of first and second order terms to include in the propensity score model. We use the `glmnet` command in the `glmnet` package to implement and select the penalty parameter using 10-fold cross-validation (Friedman, Hastie and Tibshirani, 2010).
- GenMatch: Genetic matching, developed by Diamond and Sekhon (2013), generalizes the Mahalanobis distance metric by defining the distance between observations i and j as $GMD(X_i, X_j, W) = \sqrt{(X_i - X_j)'(V^{-1/2})'W(V^{-1/2})(X_i - X_j)}$, where W is a weight matrix and $V^{-1/2}$ is the Cholesky decomposition of $V = var(X)$. As in Mahalanobis distance, X_i is the covariate vector of observation i but can also include an estimate of the propensity score. A “genetic” search algorithm is used to find weight W that minimizes imbalance across matched treatment and control observations (as measured by the minimum p -value from t -tests on individual covariates, for example). We implement the algorithm using the `GenMatch` command of the `Matching` package (Sekhon, 2011).
- Classification Tree: A classification tree partitions the covariate space into rectangular cells based on threshold split rules. The estimated propensity score is simply the proportion of treated units in each cell. The standard algorithm for fitting a tree only stops after a minimum cell size rule is violated, and

therefore is subject to overfitting. We adopt several standard approaches to reduce the variance of the classification tree, including pruning, bootstrap aggregating (bagging), random forest, and boosting.

- Pruning: We pare back a tree by trading off its complexity as measured by the number of terminal nodes (leaves) with classification error. The penalty parameter on the number of leaves is determined by 10-fold cross validation. We implement the pruning procedure with the `autoPrune` command in the `adabag` package (Alfaro, Gámez and García, 2013).
- Bagging: Bagging is based on a bootstrap procedure, in which a tree is fitted to each of the bootstrapped samples. Using this fitted tree, we can predict, for an observation that is not included in the bootstrap sample (out-of-the-bag, or OOB), whether it belongs to the treatment or the control group. The propensity score is then obtained by averaging all these OOB predictions for a given observation—each observation is out of the bag approximately 36 percent of the time when the sample size exceeds 30. We use the `randomForest` command in the `randomForest` package with the `mtry` parameter set to be the number of covariates over 500 bootstrap samples (default) (Liaw and Wiener, 2002).
- Random Forest: Random Forest is a variant of bagging. In each bootstrap sample, instead of using the entire set of covariates to build the tree, we randomly select a subset at each split. This decorrelates the trees across bootstrap samples and further reduces the variance. We again use the `randomForest` command in the `randomForest` package but set the `mtry` parameter to be its default, i.e. $\text{ceiling}(\sqrt{J})$ where J is the total number of covariates (Liaw and Wiener, 2002).
- Boosting: While bagging and random forest average over parallel trees across bootstrap samples, boosting grows the trees sequentially using the following algorithm. First, a simple tree is used to predict the treatment status of each observation. In subsequent iterations, misclassified observations receive larger weight when fitting a new tree. The binary predictions from each tree are linearly combined to produce the final classification, where prediction from trees with lower classification errors are upweighted. We use the `gbm` command in the `gbm` package to implement boosting, where each tree is a stump (i.e. it only has one split and therefore two leaves) by default, and the number of trees is selected using 5-fold cross validation (Ridgeway, 2017).
- SVM: A support vector machine partitions the covariate space using potentially nonlinear decision boundaries with the help of a kernel function $K(x_i, x_j)$. The kernel function collapses the multi-

dimensional characteristics of observations i and j into a measure of similarity that is computationally inexpensive to obtain, and it captures interactions and nonlinearities of the covariate space. Decision boundaries are defined by a function $f(x) = \beta_0 + \sum_{i=1}^N \alpha_i K(x, x_i) = 0$, and it can be shown that α_i is only non-zero for a subset of observations close to the decision boundary (“support vectors”). Observations are classified as “treated” if and only if $f(x) > 0$. To convert the classifier into a propensity score, a logit is fitted to the distance measure $f(x)$. We use the `svm` command in package `e1071` for implementation (Meyer et al., 2017), with the default radial kernel. The kernel and cost parameters are determined by following recommendations from Hsu, Chang and Lin (2016).

C.4 Performance of Matching Estimators

Our validation study results for matching estimators are summarized in Appendix Table A.4 (full samples by gender) and Appendix Table A.5 (earnings history samples by gender). For each method listed above, we evaluate its performance using the following covariate sets:

1. Demographic (non-earnings) covariates only: geographic location (3 dummies), 10-year age categories (3 dummies), race/ethnicity (indicators for Black, other race, and Hispanic), indicator for married, indicator for having a child under the age of 6, and educational attainment (4 dummies).
2. Demographic covariates *and* full (12-quarter) earnings history: all covariates above and the average monthly earnings in each of the 12 quarters preceding random assignment (for experimental controls) or eligibility determination (for ENPs).
3. Full (12-quarter) earnings history only.
4. One quarter of earnings prior to random assignment/eligibility determination.

For each estimator, we report the estimated “treatment effect” of belonging to the experimental control group (relative to ENP) on average monthly earnings over the 18-month followup period (i.e., the bias, because the true treatment effect is zero), the standard error of the estimated treatment effect, and the t -statistic.

The first lines of Panels A and B of Appendix Table A.4 show that for the full sample, taking the simple difference between experimental control and ENP groups yields different amounts of bias for men and women. For men (Panel A), the ENP group had higher earnings in the follow-up period than the experimental controls, leading to a large negative “treatment effect” of almost 400 dollars. In contrast, for

women (Panel B), simple comparisons between the experimental controls and ENPs yield little bias that is difficult to improve upon. Column (1) of Panel A shows that for men, none of the estimators purge the large and statistically significant negative bias when we use demographic covariates only, with the exception of difference-in-differences matching. However, once we incorporate earnings history into the covariate set (column 2), the bias shrinks substantially and becomes statistically insignificant in almost all cases (the exceptions are Mahalanobis matching and SVM). Notably, with the inclusion of a full earnings history, simple L-PSM performs comparably to D-in-D matching. Appendix Figures A.19 and A.20 show visually the importance of including earnings in the propensity score. Each panel plots the average earnings of the experimental controls and their nearest neighbor ENP matches, using different sets of matching variables. Comparing the upper left and upper right panels of both figures, it appears that the inclusion of earnings in the propensity score ensures that ENP workers who exhibit the same downward dip in earnings are matched to the experimental controls. This reduces much of the bias for the men sample.

Looking down column (2) of Panel A of Appendix Table A.4, we see that logit-based propensity score with variable selection algorithms which allow for quadratic and interactions of covariates (Imbens and Rubin, 2015 covariate selector and penalized logit) do not necessarily perform better than simply using a fixed set of linear terms. Similarly, other machine learning methods that allow for flexible functional forms (i.e., GenMatch, SVM, and tree-based methods) do not consistently perform better either. By comparing within matching methods using one and five neighbors, we see that increasing the number of matches typically reduces the estimated variance. The final two columns, which only consider earnings in the covariate space (12 quarter and one quarter in columns 3 and 4, respectively) further highlight the role of prior earnings in reducing the bias, often performing comparably to when we use full set of covariates in column (2)—see the bottom panels of Appendix Figure A.19 for visual corroboration in the case of L-PSM. In Panel B, we show the analogous results for the full women sample. In general, the matching and reweighting methods do not improve on the simple difference estimates. Using only non-earnings covariates results in unbiased estimates for almost all estimators, and inclusion of earnings can even increase bias.

As mentioned above, a substantial fraction of women had tenuous labor market attachment prior to random assignment/eligibility determination. As argued by HS, these women were trying to transition into the labor market, unlike most of the men who apply for training after job loss. Therefore, their past earnings are not as informative for predicting future prospects. As we show in Appendix Table A.5 and Appendix Figures A.21-A.22, this is not the case in the earnings history samples, which better resemble the enrollees

in our Ohio analysis.

While the results using the men’s earnings history sample (Panel A of Appendix Table A.5) are expectedly similar to those from the men’s full sample (the vast majority of men had a history of positive earnings), the earnings history restriction makes a difference for women.² In contrast to Panel B of Appendix Table A.4, Panel B of Appendix Table A.5 shows that the simple difference in post-period earnings between experimental controls and ENPs is now statistically significant and negative. Similar to the men’s results, using only non-earnings covariates in any estimator does not reduce the bias, except for D-in-D matching. Once we include earnings, however, much of the bias is removed, and simple methods like L-PSM are competitive with D-in-D matching as well as more flexible machine learning methods. Looking across columns, we find that including even one quarter of earnings improves the bias for women with earnings histories in most cases. These findings are corroborated by the visual evidence in Appendix Figures A.21 to A.24, which plot the earnings of the experimental controls and their nearest neighbor ENP matches from using alternative covariate sets and methods.

We draw several broad lessons from this exercise that we apply to our specific empirical setting. First, among those who have recently worked, the inclusion of earnings in the conditioning set dramatically improves the performance of most estimators. Since many of those with recent earnings only choose to participate in training after experiencing a decline in earnings (i.e., a job loss), accounting for earnings greatly reduces the selection problem. Another lesson from this validation exercise is that although machine learning tools offer more flexible and data-driven ways to model the propensity score, there does not appear to be a universally dominant strategy in this context. Perhaps surprisingly, conventional methods that estimate logit-based propensity scores using linear covariate terms appear quite competitive in this setting.

C.5 Performance of Jacobson, Lalonde and Sullivan (2005a) Fixed Effects Specifications

Moving beyond selection-on-observables strategies, we also use the NJS to evaluate the performances of commonly adopted fixed effects models proposed by Jacobson, Lalonde and Sullivan (2005a) (hereafter, JLS). JLS use longitudinal earnings data to estimate the effects of attending community colleges for a population of UI claimants. To account for unobserved individual characteristics (that are either constant or

²One notable result for men in Appendix Table A.5 is that OLS yields a larger bias than L-PSM when earnings are included.

evolving linearly), JLS estimate models of the form

$$Y_{it} = \beta E_{it} + \alpha_i + \omega_i t + \gamma_t + \delta_{it}(s_i, z_i) + \varepsilon_{it} \quad (\text{A12})$$

where Y_{it} is the earnings of individual i at time t , E_{it} is an indicator for whether the individual has started school as of time t , α_i and $\omega_i t$ are individual fixed effects and linear time trends, γ_t denote time fixed effects, and $\delta_{it}(s_i, z_i)$ are layoff effects that depend on layoff date s_i and fixed individual characteristics z_i .³

Adapting the JLS regression model to the NJS study, we define $E_{it} = 1$ if individual i is in the experimental control group and t is a month after random assignment (and 0 otherwise). Although we do not observe layoffs directly in the NJS data, we define the layoff date (if any) as the most recent change in job search status (from “not searching” to “searching”) before random assignment or eligibility determination.⁴ In our most basic specification, $\delta_{it}(s_i, z_i) = \sum_{k=-35}^{29} D_{it}^k \delta_k$, where D_{it}^k are a full set of dummy variables for month relative to layoff (35 and 29 correspond to the maximum pre- and post-layoff period observed): $D_{it}^k = 1$ if individual i had been laid off in month $t - k$. We also follow JLS and include a set of heterogeneous layoff effects that allow the earnings patterns relative to layoff to parametrically depend on demographic variables z_i , which include region, ten-year age bins, race, and education categories—see Jacobson, Lalonde and Sullivan (2005a) for the exact form of these heterogeneous layoff patterns.

In Appendix Table A.6, we present the estimated β from equation (A12), or bias, obtained by using the Stata package `reghdfe` by Correia (2016). In all specifications, we account for calendar month effects γ_t and a basic set of layoff effects $\delta_{it}(s_i, z_i) = \sum_{k=-35}^{29} D_{it}^k \delta_k$. We also include in various specifications 1) individual fixed effects α_i , 2) individual time trends $\alpha_i + \omega_i t$, and 3) individual time trends and heterogeneous layoff effects. We find that the bias from individual fixed effects specifications are statistically insignificant in three of the four samples, while specifications that included individual time trends leads to positive bias in all samples. Allowing for more flexible layoff effects do not appear to reduce bias.

Finally, we examine the pre-trends in earnings from the fixed effects models. We estimate a version of equation (A12) that replaces E_{it} with a set of dummy variables denoting time relative to random assignment (which all equal to 0 for ENPs). In Appendix Figure A.25, we plot the “event study” coefficients on these

³This is a simplification of Jacobson, Lalonde and Sullivan (2005a)’s model in that enrollment here is a binary state and we are estimating the average earnings effects after first enrolling. JLS also consider the incremental earnings effects of credits earned as well as earnings dynamics during and after school by replacing βE_{it} with $\tau_{it}(c_i, f_i, l_i, z_i)$ where c_i is credits earned, and f_i and l_i denote the enrollment entry and exit periods.

⁴Job search status is available for the twelve months before random assignment/eligibility determination. We observe 261 (228) and 351 (252) layoffs in the men’s and women’s (earnings history) samples, respectively.

dummy variables for each specification, for the earnings history samples.⁵ We find that even though some of the individual fixed effects specifications yield little bias, there are significant differences in earnings patterns between experimental controls and ENPs prior to random assignment that are not captured by the model. Including individual time trends appears to reduce the pre-trend problem, but, as noted above, it actually exacerbates bias.

D Alternative Identification Strategies

D.1 Jacobson, Lalonde and Sullivan (2005a) Fixed Effects Specifications

Although our validation exercise with NJS data in Appendix C.5 shows that individual fixed effects models perform worse than matching, it is possible that this is due to the lack of direct layoff information in the NJS. In our empirical setting, we do observe layoff timing accurately for all workers, which may improve the performance of fixed effects models. In Appendix Figure A.26, we plot the estimated β_k from the regressions

$$Y_{it} = \sum_{k=-20}^{23} R_{it}^k \beta_k + \alpha_i + \omega_{it} + \gamma_t + \delta_{it}(s_i, z_i) + \varepsilon_{it}$$

where Y_{it} is earnings for person i in quarter t , R_{it}^k is an indicator for beginning enrollment in quarter $t - k$ (and is equal to 0 for all non-enrollee observations). Layoff effects $\delta_{it}(s_i, z_i)$ include a set of dummies for quarter relative to layoff (as described in Appendix C.5 except the time unit is quarter instead of month). The regressions are estimated on a 5 percent random sample to speed up computation.

The specification shown in Panel A of the figure includes individual fixed effects α_i , calendar quarter fixed effects γ_t , and quarter relative to layoff dummies; Panel B adds individual time trends $\alpha_i + \omega_{it}$; Panel C adds heterogeneous layoff effects per the second term of equation (2) in Jacobson, Lalonde and Sullivan (2005a), which we allow to vary by gender, minority status, major county, tenure, and age. Similar to what we see in the NJS data, this “event study” specification check on our Ohio UI data yields problematic pre-enrollment earnings differences between enrollees and non-enrollees, regardless of whether we account for individual trends. Therefore, we do not rely on these fixed effects models for our analysis.

⁵Note that if equation (A12) does not include time varying controls ω_{it} , γ_t and $\delta_{it}(s_i, z_i)$ and the data are balanced in event time, the coefficient on E_{it} is equal to the difference in the average event study coefficients between the “post” and “pre” period. But this is not true in general, which makes it difficult to “visually estimate” the coefficient on E_{it} in Appendix Table A.6 based on the event study plot of Appendix Figure A.25.

D.2 Distance-Based Instrumental Variables

Another identification strategy that has been used to estimate educational returns relies on the idea that, all else equal, students are more likely to enroll if they live close to a school (Card, 1993). Therefore, the distance to the nearest school can be used as an instrumental variable (IV) for enrollment. Since we observe the zip codes of UI claimants in our sample, we can adapt this strategy to our setting by computing the linear distance from a worker's zip code to the nearest community college.⁶

Since the validity of the distance IV approach hinges on the exogeneity of the distance measure, it is important to control for all other potential determinants of earnings that may also be correlated with distance to schools. In fact, we find workers who live close to community college to be different along observable dimensions from those who live farther away. Therefore, we control for a similar set of covariates as in our matching specification: 12 quarters of pre-layoff earnings, calendar quarter of layoff, prior industry (eight indicators), prior job tenure category (less than one year, one to six years, and more than six years), age indicators (age below 19, each year from age 19 and 59, and older than 59), whether a worker has a dependent, county unemployment rate during the month of layoff, and county fixed effects. As in our main analysis, we estimate models separately for men and women, and by whether or not they previously worked in manufacturing.

However, we do not find consistently strong first-stage relationships between distance and enrollment. Moreover, we find that the results are sensitive to transformations of the instrument (e.g., including a quadratic distance term), casting doubt on an IV framework in which the instrument is assumed to be independent to the reduced-form equation error term. We conclude that an instrumental variables strategy to estimate the returns to enrollment is not desirable for our context.

D.3 Timing of Layoff

One may conjecture that the timing of layoff may induce a discontinuous change in enrollment probability. That is, workers laid off just before the beginning of a semester may be much more likely to enroll than those laid off just after. This sudden decrease in probability may be leveraged to identify enrollment effects.

However, there do not appear to be salient discontinuities in Appendix Figure A.27, which plots fraction of enrollees by UI claim date. This should not be surprising given that schooling decisions may take time to

⁶We use the nearest community college because three quarters of enrollment are in community colleges.

materialize. In fact, the modal number of quarters between UI claim and enrollment is two as documented in the notes of Appendix Figure A.7. Therefore, implementing a regression discontinuity design in layoff date does not seem to be a viable path to estimating the returns to enrollment for unemployed workers.

E Identification of Enrollment Effects by Courses and Course Effects

In this section, we extend the potential outcomes framework in Section 4.1 and consider nonparametric identification of both the effects of enrollment by course type and the effects of taking courses related to healthcare among enrollees. Our results analogously apply to the identification of effects of other schooling characteristics such as institutions attended and degree/credential receipts—like course-taking, these variables are realized simultaneously or after enrollment and are only observed if a worker enrolls. For ease of exposition, we only consider the static framework, but our analysis can easily extend to the dynamic case investigated in Section 4.1.

First, we introduce additional notation. Let H be the binary variable indicating whether a worker pursues courses related to healthcare if she enrolls. Note that H is only observed when $T = 1$. That is, we cannot observe in the non-enrollee sample who *would have taken* a health course, which is a key distinction from the subgroup analysis in Section 5.2. The introduction of H calls for an additional argument to the potential outcome function: $Y(t, h)$ for $t, h = 0, 1$.

As we mention in Section 5.5, one may be tempted to estimate the effects of taking health courses by examining the differential enrollment effects for workers taking or not taking health courses, where the enrollment effects are estimated by differences in earnings between enrollees and their matched non-enrollees. We list assumptions below under which these quantities have a causal interpretation.

Assumption 5. $\{Y(0, 0), Y(0, 1), Y(1, 0), Y(1, 1), H\} \perp\!\!\!\perp T | p(X)$

Assumption 6. $E[Y(0, 1) | T = 0, p(X)] = E[Y(0, 0) | T = 0, p(X)]$

Assumption 7. For $t = 0, 1$, $\{Y(0, 0), Y(0, 1), Y(1, 0), Y(1, 1)\} \perp\!\!\!\perp H | T = t, p(X)$

Assumption 5 requires conditional independence between the treatment variable and the joint distribution of potential outcomes and H . Like the potential outcomes, we think of H as realized before the treatment, but it is only observed for $T = 1$. Assumption 6 is akin to an exclusion restriction and states that when workers with the same observable characteristics do not enroll ($T = 0$) their average outcome is

the same regardless of what courses they would have taken had they enrolled in the counterfactual parallel universe. Assumption 7 mandates that within each treatment group and conditional on observed characteristics, the health course-takers are a random sample of the population in that their potential outcomes do not systematically differ from non-health course-takers. We state and prove the following two lemmas before proceeding to the identification results.

Lemma 2. *Under Assumption 5, for $t = 0, 1$:*

$$p(X) \perp\!\!\!\perp H|T = t$$

Proof. We prove the lemma by showing that for any constant π ,

$$\begin{aligned} & \Pr(p(X) < \pi|T = t, H = 1) \\ &= \frac{\Pr(H = 1|T = t, p(X) < \pi) \Pr(T = t, p(X) < \pi)}{\int_0^1 \Pr(H = 1|T = t, p(X) = p) dF_{p(X)|T=t}(p) \cdot \Pr(T = t)} \\ &= \frac{\Pr(H = 0|T = t, p(X) < \pi) \Pr(T = t, p(X) < \pi)}{\int_0^1 \Pr(H = 0|T = t, p(X) = p) dF_{p(X)|T=t}(p) \cdot \Pr(T = t)} \\ &= \Pr(p(X) < \pi|T = t, H = 0), \end{aligned}$$

where the second equality follows from Assumption 5. □

Lemma 3. *Under Assumptions 5 and 7,*

$$\{Y(0, 0), Y(0, 1), Y(1, 0), Y(1, 1)\} \perp\!\!\!\perp H|T.$$

Proof. We show that $Y(0, 0) \perp\!\!\!\perp H|T$, and the rest is analogous. For any constant y , and $h, t = 0, 1$

$$\begin{aligned} & \Pr(Y(0, 0) < y|H = h, T = t) \\ &= \int \Pr(Y(0, 0) < y|H = h, T = t, p(X) = p) dF_{p(X)|H=h, T=t}(p) \\ &= \int \Pr(Y(0, 0) < y|T = t, p(X) = p) dF_{p(X)|T=t}(p) \\ &= \Pr(Y(0, 0) < y|T = t), \end{aligned}$$

where the second equality follows from Assumption 7 and Lemma 2. □

Proposition 3. *Under Assumptions 5-7 and provided that $p(X) < 1$,*

$$E[Y|T = 1, H = 1] - E[E[Y|T = 0, p(X)]|T = 1, H = 1] = E[Y(1, 1) - Y(0, 1)|T = 1] \quad (\text{A13})$$

$$E[Y|T = 1, H = 0] - E[E[Y|T = 0, p(X)]|T = 1, H = 0] = E[Y(1, 0) - Y(0, 0)|T = 1]. \quad (\text{A14})$$

Proof. By Assumptions 6 and 7,

$$E[Y(0,1)|T = 0, H = 1, p(X)] = E[Y(0,0)|T = 0, H = 0, p(X)],$$

and it follows that

$$\begin{aligned} & E[Y|T = 0, p(X)] \\ &= \sum_{h=0,1} E[Y(0,h)|T = 0, H = h, p(X)] \Pr(H = h|T = 0, p(X)) \\ &= E[Y(0,h)|T = 0, H = h, p(X)] \end{aligned}$$

for $h = 0, 1$. Therefore,

$$\begin{aligned} & E[Y|T = 1, H = h, p(X)] - E[Y|T = 0, p(X)] \\ &= E[Y(1,h)|T = 1, H = h, p(X)] - E[Y(0,h)|T = 0, H = h, p(X)]. \end{aligned} \tag{A15}$$

By Assumption 5, for any constant y and for $t, t', h = 0, 1$

$$\begin{aligned} & \Pr(Y(t,h) \leq y | T = t', H = h, p(X)) \\ &= \frac{\Pr(Y(t,h) \leq y, H = h | T = t', p(X))}{\Pr(H = h | T = t', p(X))} \\ &= \frac{\Pr(Y(t,h) \leq y, H = h | p(X))}{\Pr(H = h | p(X))} \\ &= \Pr(Y(t,h) \leq y | H = h, p(X)), \end{aligned}$$

and thus

$$Y(t,h) \perp\!\!\!\perp T | H = h, p(X) \tag{A16}$$

Combining (A15) and (A16) we have

$$\begin{aligned} & E[Y|T = 1, H = h, p(X)] - E[Y|T = 0, p(X)] \\ &= E[Y(1,h) - Y(0,h) | T = 1, H = h, p(X)], \end{aligned}$$

and as a result,

$$\begin{aligned}
& E[Y|T = 1, H = h] - E[E[Y|T = 0, p(X)]|T = 1, H = h] \\
&= E[E[Y|T = 1, H = h, p(X)] - E[Y|T = 0, p(X)]|T = 1, H = h] \\
&= E[E[Y(1, h) - Y(0, h)|T = 1, H = h, p(X)]|T = 1, H = h] \\
&= E[Y(1, h) - Y(0, h)|T = 1, H = h] \\
&= E[Y(1, h) - Y(0, h)|T = 1]
\end{aligned}$$

where the last equality follows from Lemma 3. □

Proposition 3 states the identification of enrollment effect by course-taking: the difference in earnings between enrollees taking (not taking) health courses and their matched non-enrollees identifies the causal effect of enrollment for that population, under Assumptions 5-7. Next, we show that the difference between these enrollment effects identifies the effect of health courses among enrollees.

Proposition 4. *Under Assumptions 5 and 6,*

$$E[Y(1, 1) - Y(0, 1)|T = 1] - E[Y(1, 0) - Y(0, 0)|T = 1] = E[(1, 1) - Y(1, 0)|T = 1]$$

Proof. Assumptions 5 and 6 imply

$$\begin{aligned}
& E[Y(1, 1) - Y(0, 1)|T = 1] - E[Y(1, 0) - Y(0, 0)|T = 1] \\
&= E[E[\{Y(1, 1) - Y(1, 0)\} - \{Y(0, 1) - Y(0, 0)\}|T = 1, p(X)]|T = 1] \\
&= E[E[Y(1, 1) - Y(1, 0)|T = 1, p(X)] - E[Y(0, 1) - Y(0, 0)|T = 0, p(X)]|T = 1] \\
&= E[Y(1, 1) - Y(1, 0)|T = 1].
\end{aligned}$$

□

Following Proposition 4, we can use the difference in the left hand sides of (A13) and (A14) to identify the causal effect of health courses $E[Y(1, 1) - Y(1, 0)|T = 1]$. But if we are simply interested in the course effect and not enrollment effects, we can obtain identification more easily and under fewer assumptions.

Proposition 5. *Under Assumptions 5 and 7,*

$$E[Y|T = 1, H = 1] - E[Y|T = 1, H = 0] = E[Y(1, 1) - Y(1, 0)|T = 1]$$

Proof.

$$\begin{aligned} & E[Y|T = 1, H = 1] - E[Y|T = 1, H = 0] \\ &= E[Y(1, 1)|T = 1, H = 1] - E[Y(1, 0)|T = 1, H = 0] \\ &= E[Y(1, 1) - Y(1, 0)|T = 1], \end{aligned}$$

where the second equality follows from Lemma 3. □

Proposition 5 states that the simple difference between the outcomes of those who do and do not take health courses identifies the effect of taking health courses under Assumptions 5 and 7. Given this simple result, it is not surprising that Assumptions 5 and 7 give rise to strong predictions. First, within the enrollee population, the distribution of the propensity score across workers that do and do not take health courses should be the same (Lemma 2). Second, baseline covariates should be balanced across these two groups. As mentioned in Section 5.5, we find that health course takers have on average higher propensity scores and lower pre-enrollment earnings, indicating violations of Assumptions 5 and 7.

F Cost-Benefit Analysis of Further Education

In light of the estimated average earnings impacts, can we justify the investment in further education? In this section, we provide back-of-the-envelope calculations on the private and social returns to retraining an additional worker. The private return compares the net present value of the stream of *after-tax* earnings impacts of retraining against the out-of-pocket education expenses an enrollee has to pay upfront. The social return compares the net present value of *pre-tax* earnings impacts of retraining against the overall cost of enrolling an additional unemployed worker.

These returns calculations require inputs on long-run earnings impacts, years of work life remaining, real interest rate for discounting, tax rates, and various cost measures of education investment. We adopt the following approximations in our calculations of the net present value of long-term earnings impacts. First, we use the 10-year post-enrollment real earnings impacts graphed in Figure 5 for workers who began schooling before the fall quarter of 2007 and assume that the earnings gains in the long run will stay at the level in the tenth year. Second, we assume that an average enrollee has 30 years remaining in her work life just prior to retraining. Given that an average enrollee in our sample is about 35 years old at layoff (Table 1) and that it takes slightly less than a year from layoff to enrollment (Table 2), our assumption implies

that the enrollee will stop working at around 66, which is consistent with Jacobson, Lalonde and Sullivan (2005*b*) and just under the normal retirement age of 67 faced by most of our enrolled workers. Third, we use a real interest rate of two percent to discount future earnings. This is based on the fact that the average of the daily Treasury real long-term rates, as calculated from the yields of outstanding long-term Treasury Inflation-Protected Securities (TIPS), is 1.98 percent between 2009 and 2010, the two years during which enrollment in our sample peaked. Our real interest rate is lower than the four percent used by Jacobson, Lalonde and Sullivan (2005*b*) due to the different time period examined. Jacobson, Lalonde and Sullivan (2005*b*) study workers who enrolled in the 1990s when the long-term Treasury bill yields were around 7 percent and inflation rates around 3 percent.⁷ Finally, we follow Jacobson, Lalonde and Sullivan (2005*b*) and assume that workers pay 25 percent of their earnings in various taxes. This average tax rate may be on the higher end given the lower federal income tax rates relative to those in the 1990s, which will imply a conservative private return estimate. The tax rate does not enter the social returns calculations.

To compute the cost of attending post-secondary institutions, we rely on the Integrated Postsecondary Education Data System (IPEDS) and the Digest of Education Statistics by NCES. To estimate the yearly out-of-pocket expenses for a typical enrollee, we first subtract the expected amount of grant—the product of $E[\text{grant amount}|\text{receiving grant}]$ and $\text{Pr}(\text{receiving grant})$ —from the sum of annual tuition and book costs during the 2010-2011 academic year (we use the reported average grant amount among students receiving a grant and the fraction of students receiving a grant to estimate $E[\text{grant amount}|\text{receiving grant}]$ and $\text{Pr}(\text{receiving grant})$ respectively; to the extent that displaced workers are more likely to qualify for grants and receive larger grants, we overstate the out-of-pocket costs and understate the returns to enrollment). We do this separately for the three types of Ohio institutions (two-year college, four-year college/university and technical center) that appear in the IPEDS database, and we average the yearly out-of-pocket expenses across the three types weighted by the proportion of enrollees attending each (unlike two- and four-year institutions that report cost information for each academic year, most technical centers report cost statistics by program, and we use the average costs corresponding to the largest program offered).

To estimate the social investment on an additional enrollee, which encapsulates both the enrollee's expenses and government subsidies, we follow Rouse (1998) and proxy it with the variable cost per full-time-equivalent (FTE) student. Like Rouse (1998), who estimates the cost of educating a student in an

⁷TIPS yields would be the most germane measures as they encapsulate long-run inflation expectations, but they were not introduced until the late 1990s; the published Treasury real long-term rates date back to early 2000, at which point they were just above four percent—consistent with Jacobson, Lalonde and Sullivan (2005*b*).

associate degree program, we obtain our variable cost by excluding from the total expenditure any fixed costs (spending on administration, public service, operation and maintenance) and research outlays (which is not important for training the enrollees in our sample and accounts for less than 0.1 percent of the total expenditure). Rouse (1998) also proposes an adjustment for the variable cost of four-year colleges because it is cheaper to add a two-year student than an upper class undergraduate or graduate student; we adopt the same adjustment. We compute the cost measures using information for the 2010-2011 academic year from Snyder and Dillow (2013), and the annual social investment per FTE student is \$8,084 in a community college and \$14,122 in a four-year college.⁸ Since expenditures are not available for technical centers in Snyder and Dillow (2013), we proxy them with the larger four-year college figure. Lastly, since the average length of enrollment is 4.7 terms, we multiply the yearly cost measures by $4.7/2 = 2.35$, assuming two academic terms per year. There could be more than two terms per year, and enrollees in our sample might not have enrolled full time, both of which imply an overestimate of cost and conservative estimate of the returns.

The net present value of the average post-tax earnings impacts is \$43,827 against an out-of-pocket cost of \$6,636, and the resulting private net benefit an enrollee accrues is \$37,190. Another way to interpret these quantities is that the worker gets \$6.60 for every dollar invested, and the private internal annual rate of return (IRR) is 15.3 percent. The net present value of the average pre-tax earnings impacts is \$58,436 against a social investment of \$23,307, and the resulting social net benefit an enrollee accrues is \$35,129. The social return is \$2.51 for every dollar invested, and the annual social IRR is 7.9 percent. Our private IRR is in the range of those reported by Heckman, Lochner and Todd (2006), and our social IRR is near the high end of the estimates by Jacobson, Lalonde and Sullivan (2005*b*) that pass a specification check (their Table 5B column 3). We have mentioned above several reasons to believe that these estimates may be conservative (e.g., the approximations of tax rate, cost of technical centers, and the use of investment per FTE student we adopt). Additionally, the earnings estimates used for this calculation correspond to an early cohort of enrollees who faced the labor market at the height of the Great Recession and likely saw lower long-term returns compared to later cohorts, as seen in Appendix Figure A.8.

The only assumption we invoke that may substantially overstate the returns is the persistence of training effects over the 30-year horizon remaining in a typical enrollee's work life, even though it is consistent with prior work by Jacobson, Lalonde and Sullivan (2005*b*). For example, many of our enrollees may have lost

⁸Since Snyder and Dillow (2013) only report total expenditure that are not broken down into detailed categories for each state, these figures represent the national averages. That said, the reported total expenditure per two-year FTE student is \$12,398 nationally and \$12,346 in Ohio. Given the small difference, the national averages are likely to serve as good substitutes.

jobs again in the COVID-19 pandemic, including those working in healthcare due to the postponement of routine and elective procedures (Scott, 2020). Another way to interpret these results without this strong assumption is to ask how long it takes to break even on the educational investment. The impacts graphed in Figure 5 suggest private (social) investment breaks even at 9 (15) years after enrollment begins.

G Impact of UI Benefit Policies on Enrollment

In this section, we examine the role of UI benefit policies on enrollment decisions. Barr and Turner (2015) show that generous benefit durations induce more unemployed workers to pursue schooling. We replicate this finding for Ohio and discuss the implications for UI policy in light of our returns estimates. We note that while UI benefit durations affect enrollment, this policy variation cannot be readily used to estimate the returns to enrollment in our main analysis because benefit durations can directly impact unemployment and labor force participation (Rothstein, 2011).

G.1 Background and Summary of Barr and Turner (2015)

Under normal economic conditions, unemployed workers in Ohio (and in most other states) are eligible to receive 26 weeks of UI benefits, which replace 50 percent of past earnings up to a cap (the cap ranges from \$323 to \$524 per week during our sample period, depending on the year and number of dependents a worker has). The duration of benefits may be increased during economic downturns. As shown in Appendix Figure A.28, benefit durations varied substantially over our sample period in Ohio, reaching 99 weeks in 2009 and persisting for a few years afterward.⁹ The degree to which benefits were extended depended on the state unemployment rate and policies in place at a given time (see Rothstein, 2011 for details). Using the October Education Supplement of the Current Population Survey, Barr and Turner (2015) show that these changes in UI policy across states and over time increased unemployed workers' propensity to pursue post-secondary education. We use the same temporal variation to estimate the magnitude of the effect within Ohio.

One major difference between our analysis and Barr and Turner (2015)'s is our ability to observe the timing of enrollment (and unemployment), which allows us to relate enrollment to the UI policy in place at the time of the enrollment decision.¹⁰ This is important because, as discussed in Rothstein (2011), sudden

⁹The narrow valleys are due to the failure to extend EUC08 legislation before a scheduled policy expiration. However, claimants were retroactively compensated after these lapses.

¹⁰Barr and Turner (2015)'s main regressor is the UI duration in the August of the year enrollment is observed, under the assump-

changes in federal legislation during the Great Recession resulted in changing expectations on benefit duration over the unemployment spell. One way to see this is in Appendix Figure A.29. For this figure, we simulate the number of UI benefit weeks remaining for workers at various quarters of the unemployment spell using only variation in UI extension policies over time, and we plot the averages for different cohorts of claimants.¹¹ Before the recession, workers expected 26 weeks of benefits at the beginning of unemployment and would run out of benefits after two quarters. For cohorts laid off in 2008, workers began unemployment under the assumption that they were eligible for 26 weeks of benefits, but as new extensions began in June 2008, workers with relatively long spells of unemployment were eligible for the new extensions. Throughout 2009 and 2010, benefits were continually extended such that although workers used their benefits, the average number of UI benefit weeks remaining declined less quickly than would be expected mechanically with the passage of time—in fact, for the 2009 cohort, the remaining benefit weeks actually increased at times. To the extent that workers base enrollment decisions on expectations of UI remaining available, it is important to consider this policy variation over the course of unemployment.

G.2 Effect of UI Benefit Duration on Enrollment

Our analysis uses a 5 percent subset of our UI claims sample, which covers claimants from 2004-2011Q3.¹² For each claimant, we have a balanced panel of eight quarters starting in the first quarter after her claim. We are interested in the effect of the expected UI benefit duration on enrollment over the first two years of layoff. We estimate equations of the following form:

$$E_{it} = \beta P_{it} + p(UR_t; \rho) + \sum_{k=1}^8 \delta^k D_{it}^k + \lambda_t + \mathbf{X}_i \gamma + \varepsilon_{it}$$

where E_{it} is the enrollment indicator for worker i in quarter t , P_{it} is the UI potential benefit duration in quarter t , D_{it}^k are a set of indicators denoting the k th quarter since layoff, UR_t is the state unemployment rate at time t , and \mathbf{X}_i contain demographic and pre-layoff job characteristics. The main coefficient of interest is β , which is the effect of the benefit duration on enrollment. Since benefit durations are partially determined by state economic conditions, we include a quadratic function $p(UR_t; \rho)$ of the unemployment rate, following Barr and Turner (2015). We also flexibly control for time since layoff (δ^k), year and quarter-in-year effects (λ_t),

tion that workers make decisions to enroll at the start of the academic year.

¹¹This simulation assumes that all workers receive 26 weeks of regular benefits and is continuously unemployed starting on their claim date. The simulation code is lightly adapted from Rothstein (2011).

¹²One difference between our sample and Barr and Turner (2015)'s is that ours is a sample of *layoffs* while Barr and Turner (2015)'s is a sample of *unemployment spells*. The latter likely contains disproportionately more long-term unemployed workers.

and worker characteristics including gender, age category, race, whether a worker reports having dependents, past wage quintile, tenure at last employer (four categories), past industry (two-digit NAICS), and past occupation (two-digit SOC).

Appendix Table A.7 presents our regression results. The first column shows that a ten-week increase in potential benefit duration raises the probability of enrollment by 0.15 percentage points, or a 10 percent increase. Although this effect is smaller than the point estimate found in Barr and Turner (2015), it does fall inside their 95 percent confidence interval. As noted above, since potential durations were changing throughout the unemployment spell during the recession, the benefit duration presumed at the beginning of a worker's unemployment spell differs from the benefit duration actually experienced. In the second column of Appendix Table A.7, we show that the potential duration at the beginning of the spell has no impact on enrollment beyond the impact of the potential duration at the time of enrollment. The third column estimates the effect using an individual fixed effects model, utilizing only the variation in potential benefit expectations over time for each spell, and finds that the effect of potential duration is slightly smaller, though still statistically significant. Finally the fourth column only includes workers who are on their first unemployment spell, defined as those who have not yet experienced a full 13-week quarter of employment. Although the probability of enrollment is higher in this sample, the estimate is similar in percentage terms: a 10-week increase in potential benefit duration increases enrollment by 11 percent.

G.3 Policy Implications

These estimates indicate that a 10-week increase in UI potential durations induces approximately 1,200 more workers to enroll annually. Assuming that our estimated return of \$387 per quarter in the third and fourth year represent an estimate of the long-run returns, the increased enrollment would imply about \$1.8 million per year in earnings gain starting in the fifth year after enrollment (since the lock-in effects in the first two years roughly equal the positive returns in the third and fourth years).

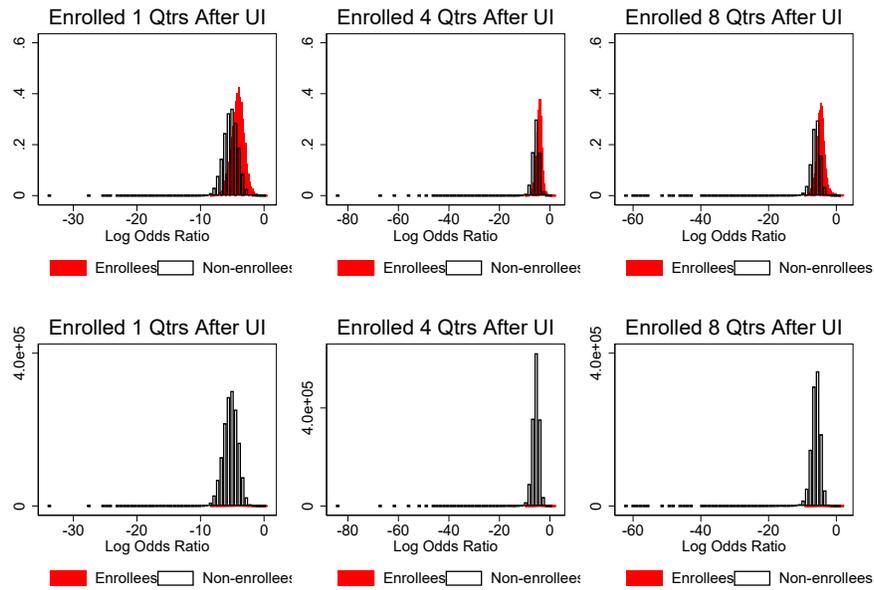
This finding also suggests that there is an externality associated with extending UI. As discussed in Schmieder and von Wachter (2016) and Lee et al. (2019), the overall impact of a policy on the government budget is a critical parameter in optimal policy-making. Therefore, estimates of the implied increases in tax revenues and government expenditures related to financial aid or tuition subsidies should be accounted for in UI policy analysis.

Appendix References

- Abadie, Alberto, and Guido W. Imbens.** 2006. "Large Sample Properties of Matching Estimators for Average Treatment Effects." *Econometrica*, 74(1): 235–267.
- Alfaro, Esteban, Matías Gámez, and Noelia García.** 2013. "adabag: An R Package for Classification with Boosting and Bagging." *Journal of Statistical Software*, 54(2): 1–35.
- Barr, Andrew, and Sarah Turner.** 2015. "Out of Work and into School: Labor Market Policies and College Enrollment During the Great Recession." *Journal of Public Economics*, 124: 63–73.
- Bureau of Labor Statistics.** n.d.. "National Employment Matrix 2018-28."
- Card, David.** 1993. "Using Geographic Variation in College Proximity to Estimate the Return to Schooling." National Bureau of Economic Research Working Paper 4483.
- Correia, Sergio.** 2016. "Linear Models with High-Dimensional Fixed Effects: An Efficient and Feasible Estimator." Working Paper.
- Diamond, Alexis, and Jasjeet S. Sekhon.** 2013. "Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies." *Review of Economics and Statistics*, 95(3): 932–945.
- Friedman, Jerome, Trevor Hastie, and Robert Tibshirani.** 2010. "Regularization Paths for Generalized Linear Models via Coordinate Descent." *Journal of Statistical Software*, 33(1).
- Heckman, James, and Jeffrey Smith.** 1999. "The Pre-Programme Earnings Dip and the Determinants of Participation in a Social Programme, Implications for Simple Programme Evaluation Strategies." *The Economic Journal*, 109: 313–348.
- Heckman, James, Hidehiko Ichimura, and Petra Todd.** 1997. "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme." *The Review of Economic Studies*, 64(4): 605–654.
- Heckman, James, Hidehiko Ichimura, Jeffrey Smith, and Petra Todd.** 1998. "Characterizing Selection Bias Using Experimental Data." *Econometrica*, 66(5): 1017–1098.
- Heckman, James, Lance Lochner, and Petra Todd.** 2006. "Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond." In *Handbook of the Economics of Education*. Vol. 1, , ed. Eric A. Hanushek and Finis Welch, Chapter 7, 307–458. Elsevier.
- Hsu, Chih-Wei, Chih-Chung Chang, and Chih-Jen Lin.** 2016. "A Practical Guide to Support Vector Classification." National Taiwan University.
- Imbens, Guido W.** 2004. "Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review." *The Review of Economics and Statistics*, 86(1): 4–29.
- Imbens, Guido W., and Donald B. Rubin.** 2015. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press.
- Jacobson, Louis, Robert J. Lalonde, and Daniel Sullivan.** 2005a. "Estimating the Returns to Community College Schooling for Displaced Workers." *Journal of Econometrics*, 125(1): 271–304.

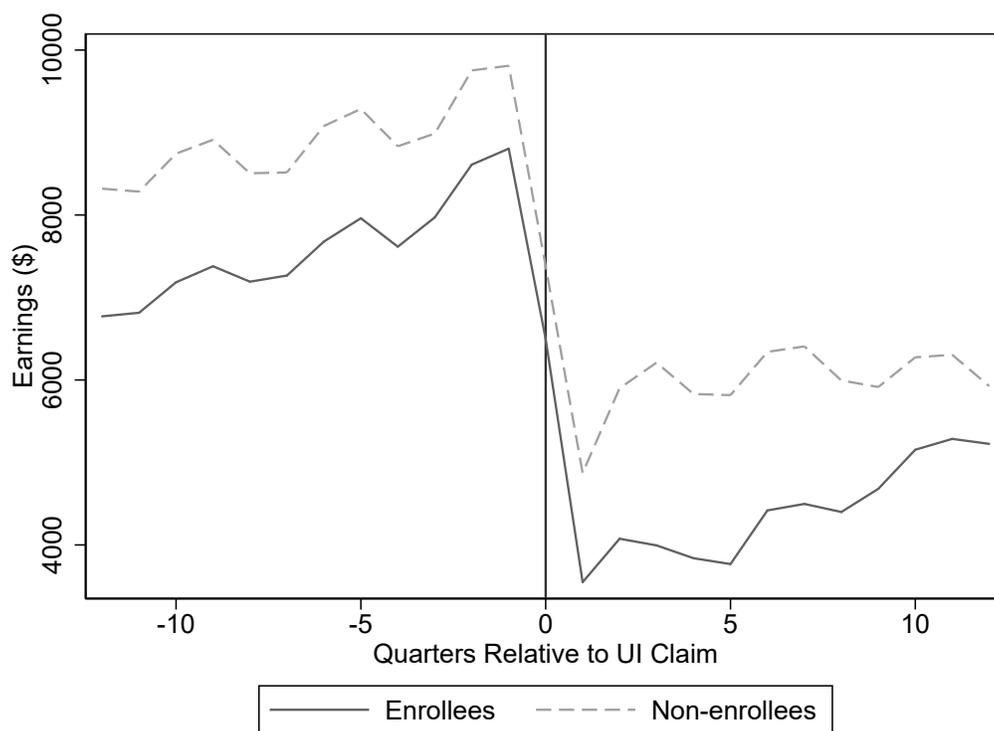
- Jacobson, Louis, Robert J. Lalonde, and Daniel Sullivan.** 2005*b*. “The Impact of Community College Retraining on Older Displaced Workers: Should We Teach Old Dogs New Tricks?” *ILR Review*, 58(3): 398–415.
- LaLonde, Robert.** 1986. “Evaluating the Econometric Evaluations of Training Programs with Experimental Data.” *The American Economic Review*, 76(4): 604–620.
- Lee, David S., Pauline Leung, Christopher J. O’Leary, Zhuan Pei, and Simon Quach.** 2019. “Are Sufficient Statistics Necessary? Nonparametric Measurement of Deadweight Loss from Unemployment Insurance.” National Bureau of Economic Research Working Paper 25574.
- Liaw, Andy, and Matthew Wiener.** 2002. “Classification and Regression by randomForest.” *R News*, 2(3): 18–22.
- Meyer, David, Evgenia Dimitriadou, Kurt Hornik, Andreas Weingessel, and Friedrich Leisch.** 2017. “e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien.” R package version 1.6-8.
- National Center for Education Statistics.** 2011. “CIP 2010 to SOC 2010 Crosswalk.”
- Ridgeway, Greg.** 2017. “gbm: Generalized Boosted Regression Models.” R package version 2.1.3.
- Rosenbaum, Paul R., and Donald B. Rubin.** 1983. “The Central Role of the Propensity Score in Observational Studies for Causal Effects.” *Biometrika*, 70(1): 41–55.
- Rothstein, Jesse.** 2011. “Unemployment Insurance and Job Search in the Great Recession.” *Brookings Papers on Economic Activity*, 43(2): 143–213.
- Rouse, Cecilia Elena.** 1998. “Do Two-Year Colleges Increase Overall Educational Attainment? Evidence from the States.” *Journal of Policy Analysis and Management*, 17(4): 595–620.
- Schmieder, Johannes F., and Till von Wachter.** 2016. “The Effects of Unemployment Insurance Benefits: New Evidence and Interpretation.” *Annual Review of Economics*, 8.
- Scott, Dylan.** 2020. “Hospital Are Laying Off Workers in the Middle of the Coronavirus Pandemic.” *Vox*.
- Sekhon, Jasjeet S.** 2011. “Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching Package for R.” *Journal of Statistical Software*, 42(7): 1–52.
- Snyder, Thomas D., and Sally A. Dillow.** 2013. “Digest of Education Statistics, 2012.” National Center for Education Statistics.
- Westreich, Daniel, Justin Lessler, and Michele Jonsson Funk.** 2010. “Propensity score estimation: neural networks, support vector machines, decision trees (CART), and meta-classifiers as alternatives to logistic regression.” *Journal of Clinical Epidemiology*, 63: 826–833.

Figure A.1: Propensity Score Distributions



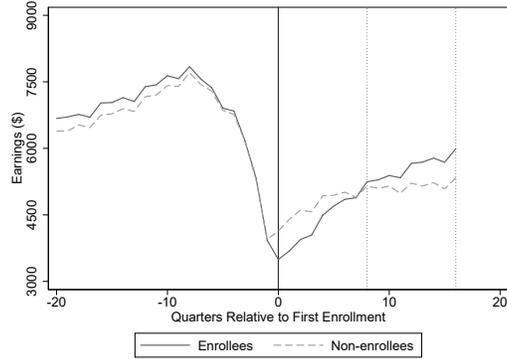
Notes: These figures show the distributions of the estimated propensity to enroll (expressed in log-odds ratios), for UI claimants who enroll in the first, fourth, and eighth quarters after layoff, and non-enrollees. The first row of figures show the distributions conditional on being an enrollee/non-enrollee; the second row shows the distributions unconditionally. There are $N = 1,579,251$; $1,496,274$; $1,124,105$ UI claims in the graphs in each column, corresponding to $1,141,278$; $1,093,273$; and $877,157$ unique individuals.

Figure A.2: Earnings of Enrollees and Non-enrollees

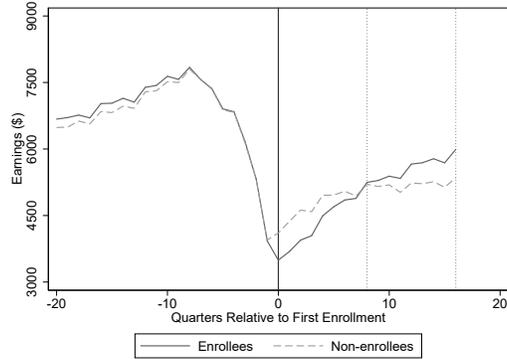


Notes: This figure plots the average quarterly earnings of enrollee and non-enrollee UI claimants for five percent of our analysis sample. The vertical line denotes the UI claim quarter. $N = 99,478$ UI claims (corresponding to 96,874 unique individuals).

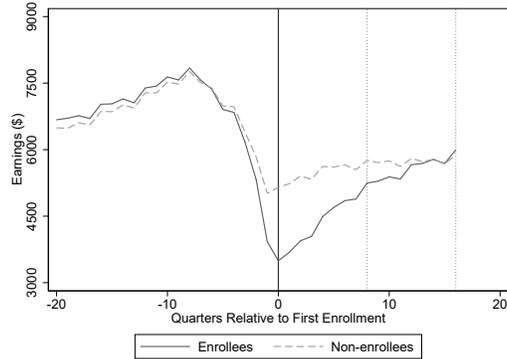
Figure A.3: Earnings of Enrollees and Matched Non-enrollees Using Alternative Matching Specifications
 (A) Matched on One Quarter of Earnings Before Layoff and Earnings Between Layoff and Enrollment



(B) Matched on Four Quarters of Earnings Before Layoff and Earnings Between Layoff and Enrollment

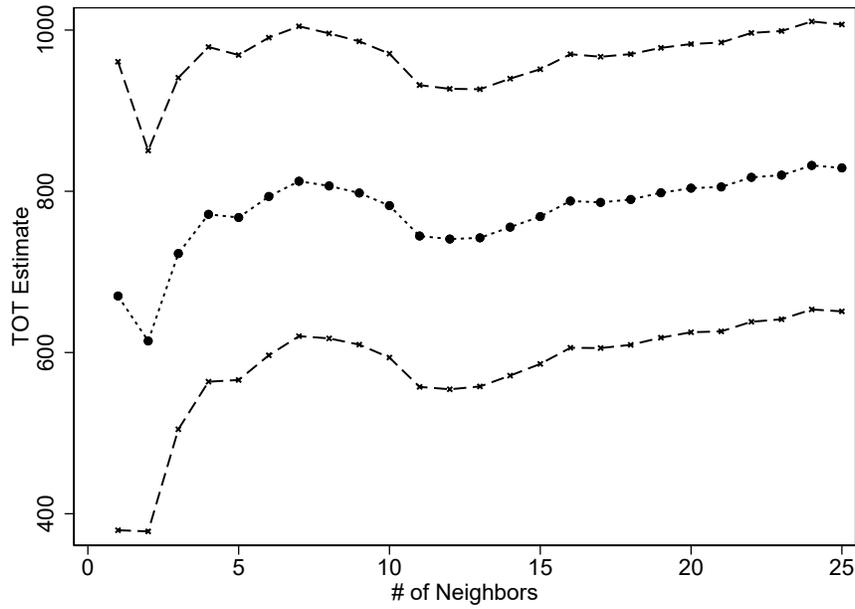


(C) Matched on Four Quarters of Earnings Before Layoff

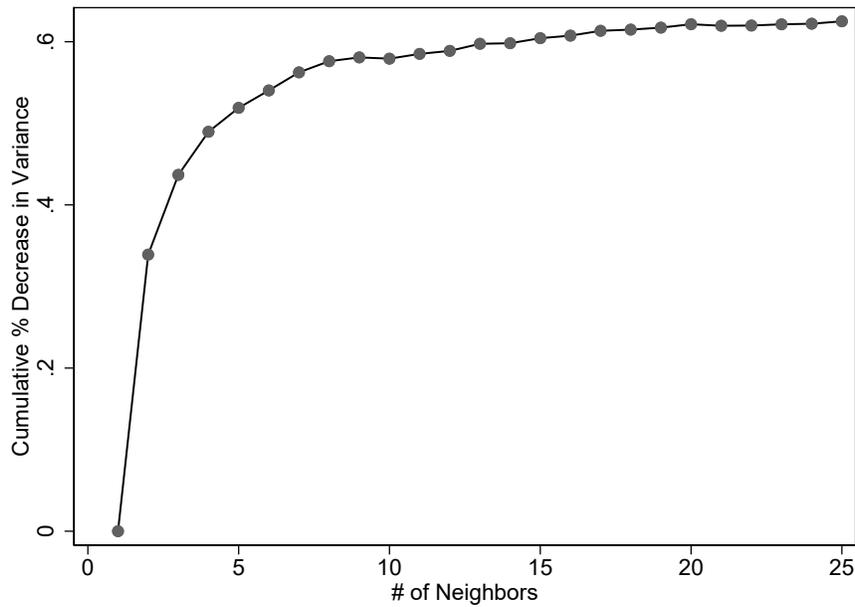


Notes: These figures plot the average quarterly earnings of enrollee and matched non-enrollee UI claimants, where the matching is done using all demographic variables described in Section 4.3 and the following alternative sets of earnings variables: one quarter of earnings before layoff and earnings between layoff and enrollment (Panel A), four quarters of earnings before layoff and earnings between layoff and enrollment (Panel B), and four quarters of earnings before layoff (Panel C). The solid vertical line denotes the quarter of first enrollment, and the vertical dashed lines denote eight and 16 quarters after first enrollment. These figures contain $N = 175,436$; $175,414$; and $175,468$ UI claims, respectively, corresponding to 167,705; 167,669; and 167,606 unique individuals.

Figure A.4: Estimated Enrollment Effects, By Number of Neighbors
 (A) Estimated Enrollment Effect



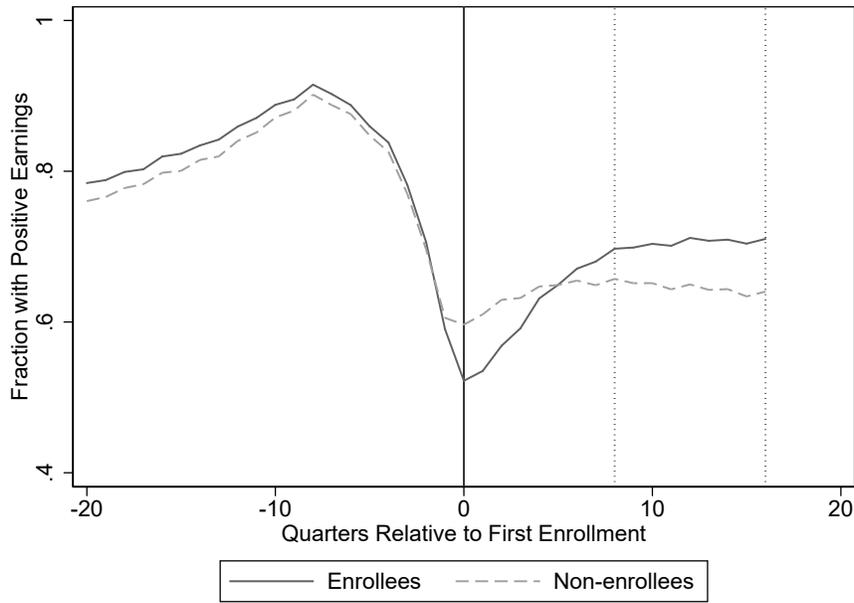
(B) Reduction in Variance



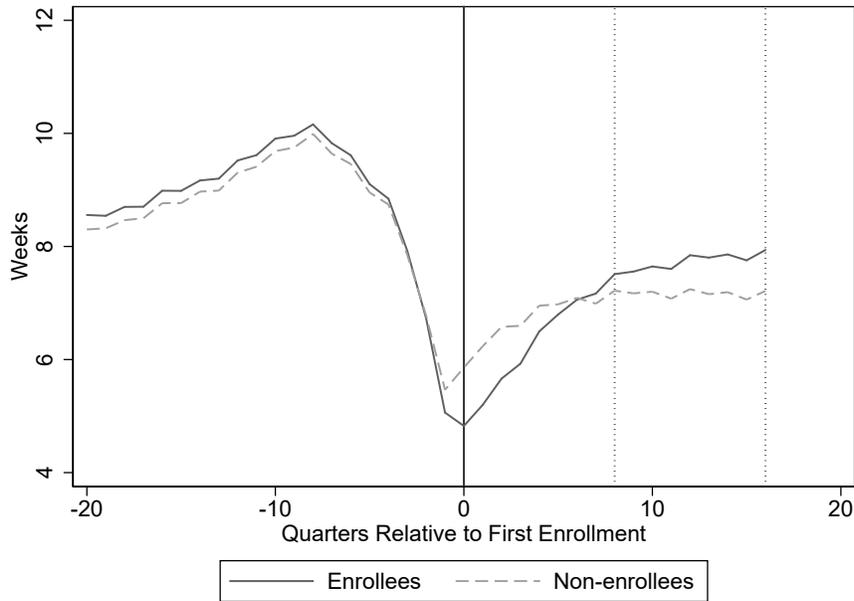
Notes: Panel A shows how the estimated enrollment effect (3-4 years after enrolling) varies by the number of matched neighbors for men who did not previously work in manufacturing and who filed a UI claim in the first quarter of 2009. Panel B shows the percent reduction in the variance of the estimated effect, relative to using one neighbor to match. $N = 58,690$ UI claims, corresponding to 58,690 unique individuals.

Figure A.5: Employment of Enrollees and Matched Non-enrollees

(A) Probability of Having Any Positive Earnings

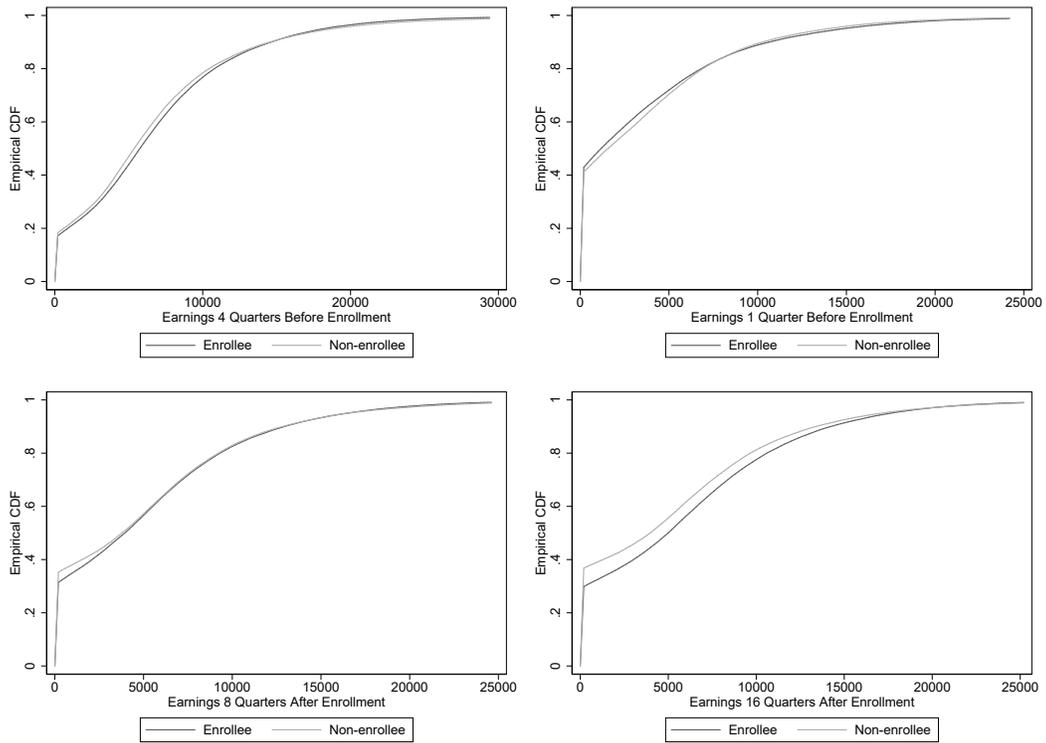


(B) Weeks Worked Per Quarter



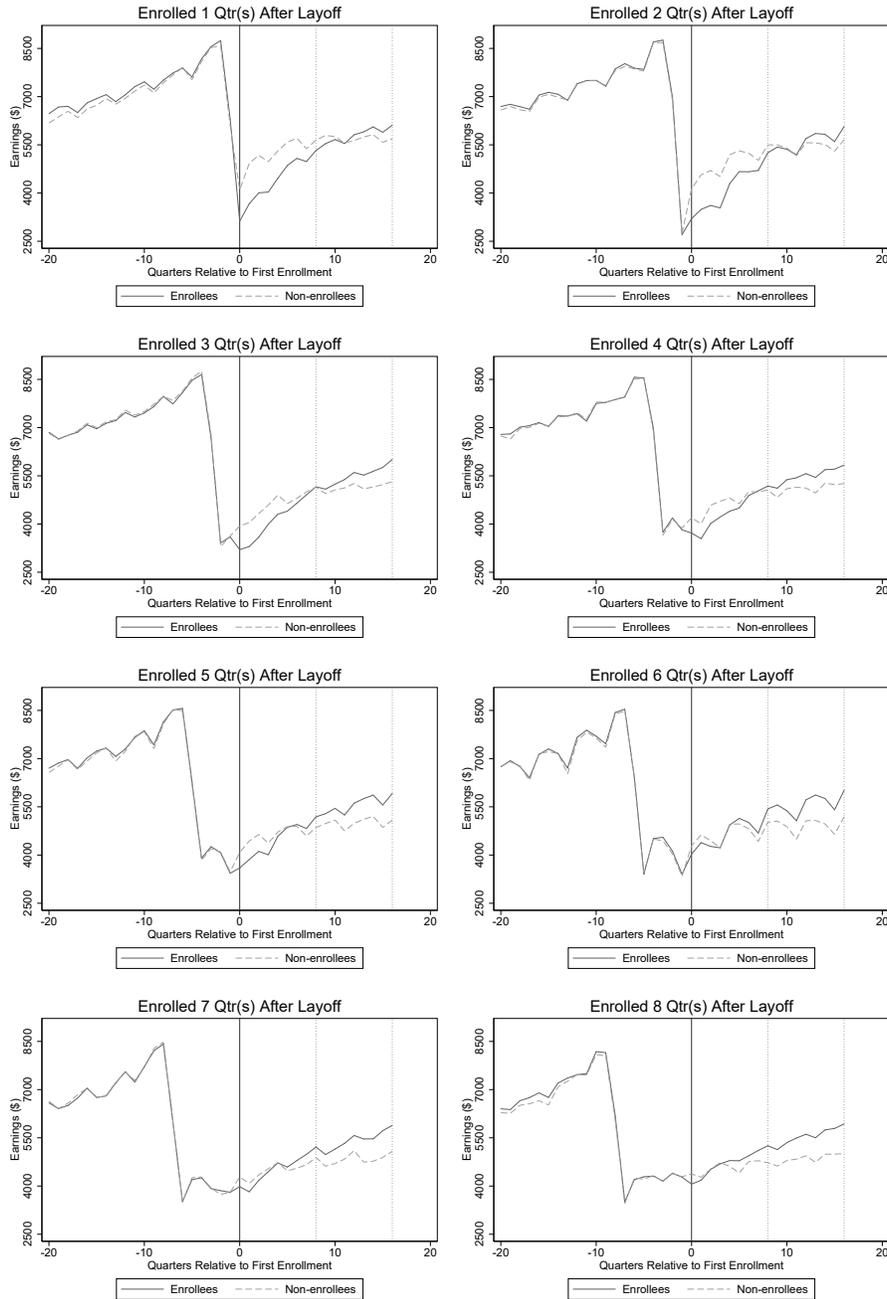
Notes: These figures plot the fraction of enrollee and matched non-enrollee UI claimants with positive earnings in each quarter (Panel A) and the quarterly average number of weeks worked (Panel B). The solid vertical line denotes the quarter of first enrollment, and the vertical dashed lines denote eight and 16 quarters after first enrollment. $N = 175,326$, corresponding to 167,422 unique individuals.

Figure A.6: Distributions of Enrollee and Matched Non-enrollee Earnings



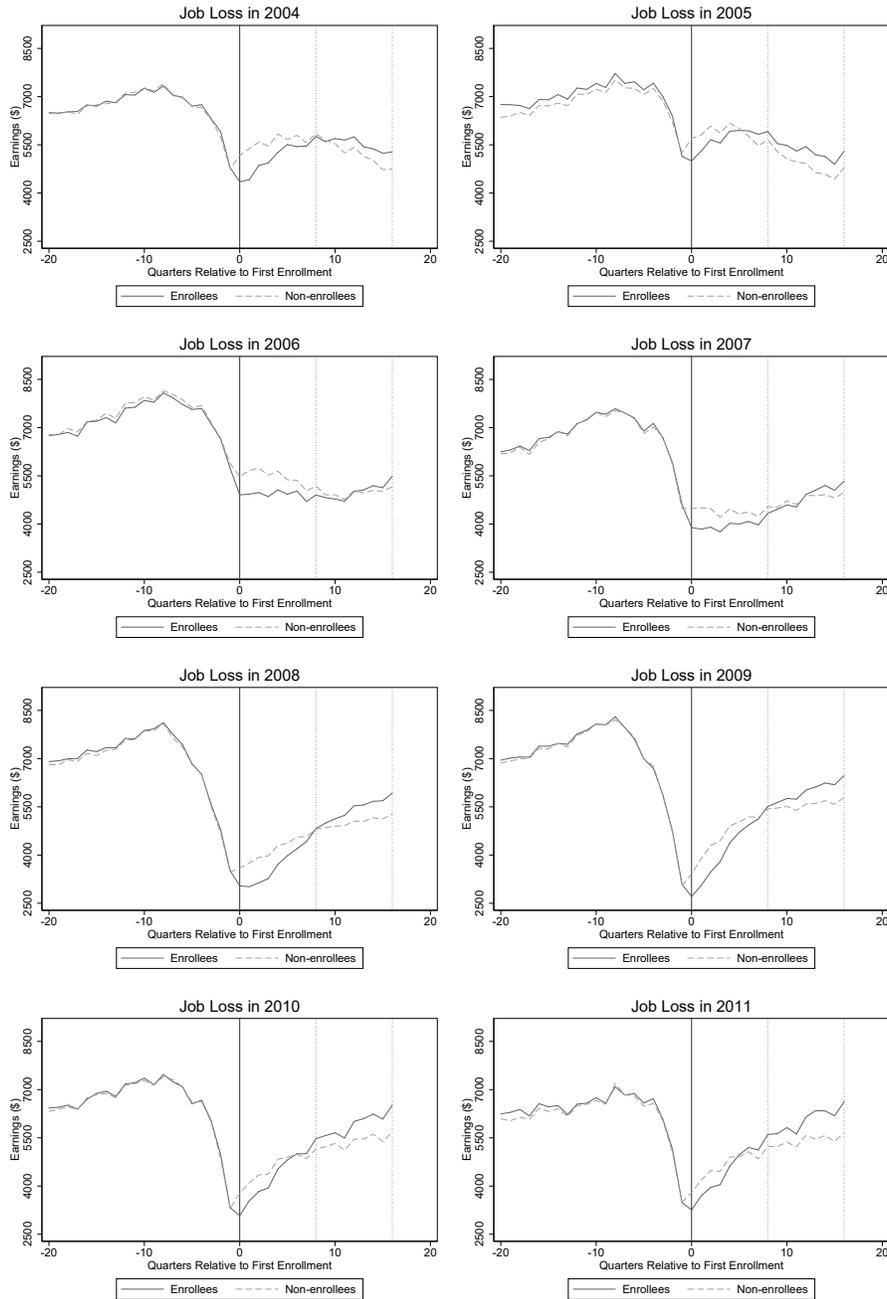
Notes: These figures show the empirical cumulative distribution functions of earnings for enrollees and matched non-enrollees four quarters before enrollment, one quarter before enrollment, eight quarters after enrollment, and 16 quarters after enrollment. $N = 175,326$, corresponding to 167,422 unique individuals, for each graph.

Figure A.7: Earnings of Enrollees and Matched Non-enrollees, By Enrollment Timing



Notes: Each graph shows the average quarterly earnings of UI claimants who enroll a certain number of quarters after filing a UI claim and their matched non-enrollees. The solid vertical lines denote the quarter of first enrollment, and the vertical dashed lines denote eight and 16 quarters after first enrollment. There are $N = 29,038$; 34,042; 28,880; 23,220; 18,790; 16,204; 13,894; 11,258 UI claims in each graph, corresponding to 28,663; 33,540; 28,469; 22,909; 18,608; 16,046; 13,743; 11,162 unique individuals.

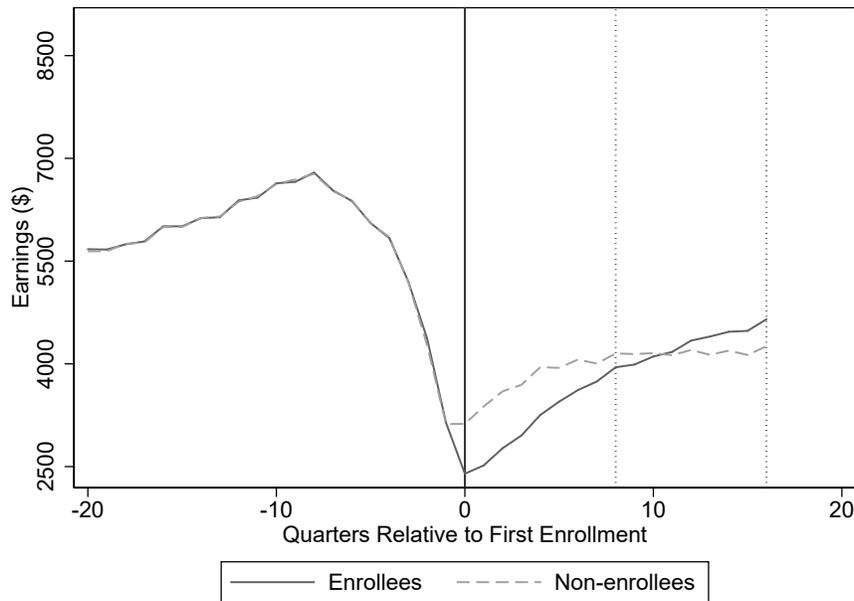
Figure A.8: Earnings of Enrollees and Matched Non-enrollees, By Year of Job Loss



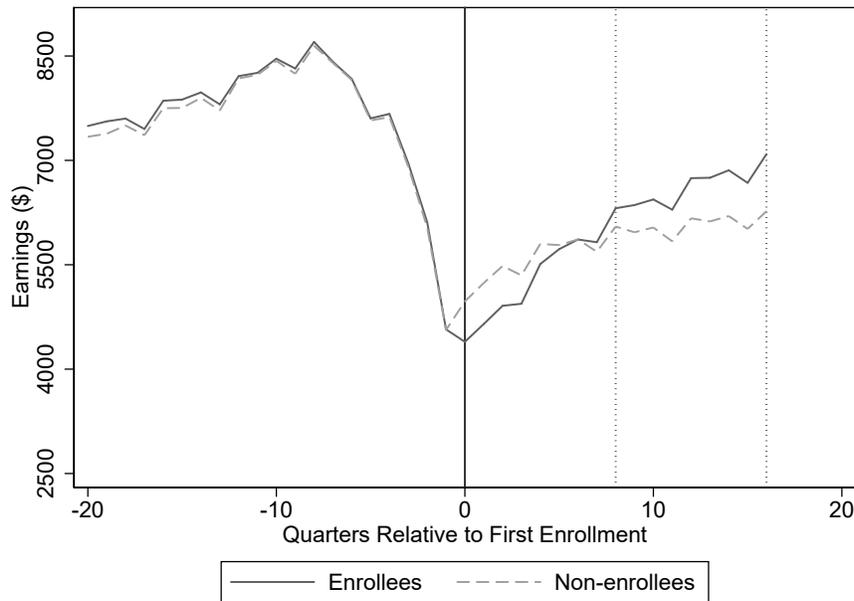
Notes: Each graph shows the average quarterly earnings of enrollees and matched non-enrollees who filed a UI claim in a specific year. The solid vertical lines denote the quarter of first enrollment, and the vertical dashed lines denote eight and 16 quarters after first enrollment. There are $N = 15,206$; 11,566; 16,790; 18,274; 34,134; 42,212; 23,744; 13,400 UI claims in each graph, corresponding to 14,840; 11,291; 16,443; 17,852; 33,237; 41,055; 23,047; 13,027 unique individuals.

Figure A.9: Earnings of Enrollees and Matched Non-enrollees, By Gender

(A) Women

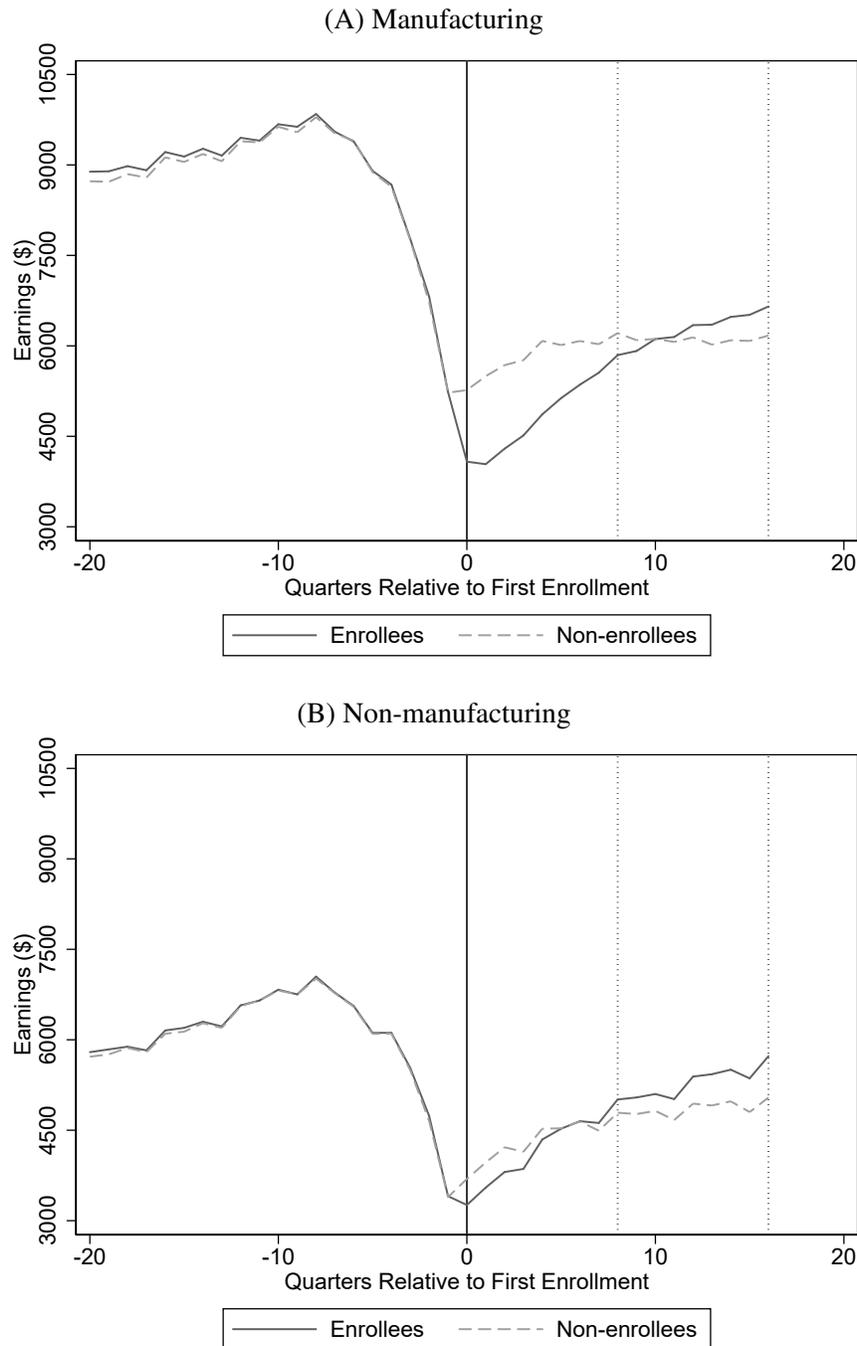


(B) Men



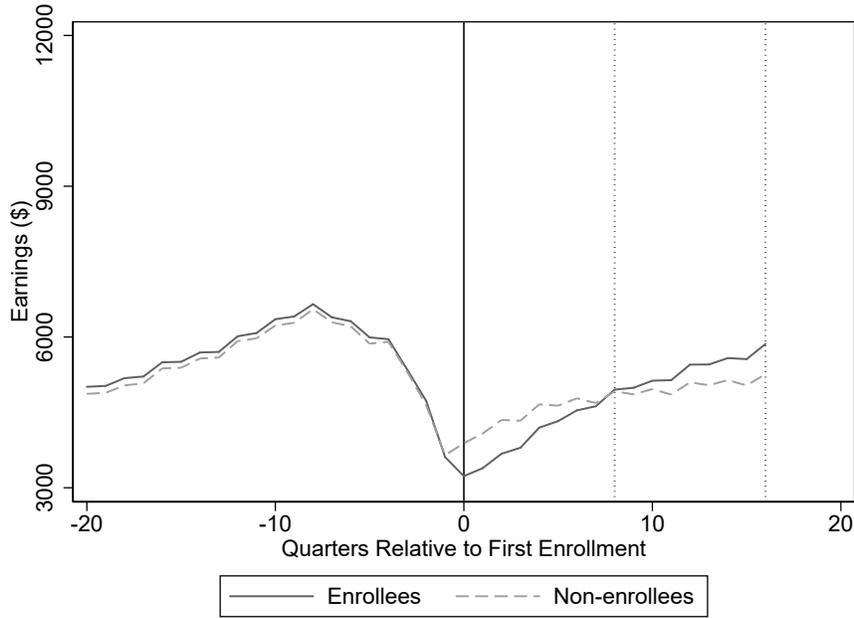
Notes: The upper (lower) graph shows the average quarterly earnings of female (male) enrollee and matched non-enrollee UI claimants. The solid vertical lines denote the quarter of first enrollment, and the vertical dashed lines denote eight and sixteen quarters after first enrollment. There are $N = 78,990$ ($96,336$) UI claims in the upper (lower) graph, corresponding to $74,810$ ($92,612$) unique individuals.

Figure A.10: Earnings of Enrollees and Matched Non-enrollees, Manufacturing vs. Non-manufacturing

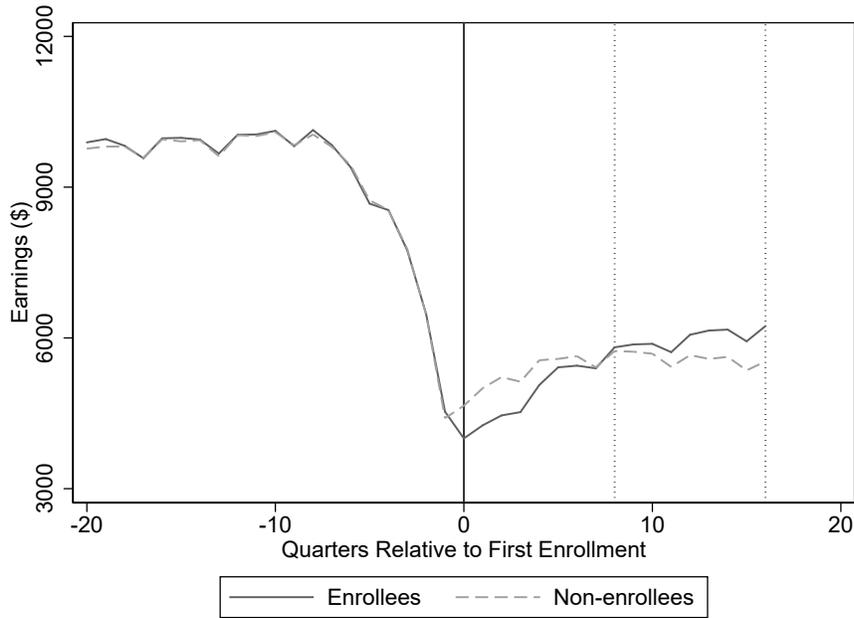


Notes: The upper (lower) graph shows the average quarterly earnings of enrollees who previously worked in a manufacturing (non-manufacturing) sector and matched non-enrollee UI claimants. The solid vertical lines denote the quarter of first enrollment, and the vertical dashed lines denote eight and 16 quarters after first enrollment. There are $N = 49,828$ (125,498) UI claims in the upper (lower) graph, corresponding to 47,580 (120,324) unique individuals.

Figure A.11: Earnings of Enrollees and Matched Non-enrollees, By Age
 (A) Under Age 40

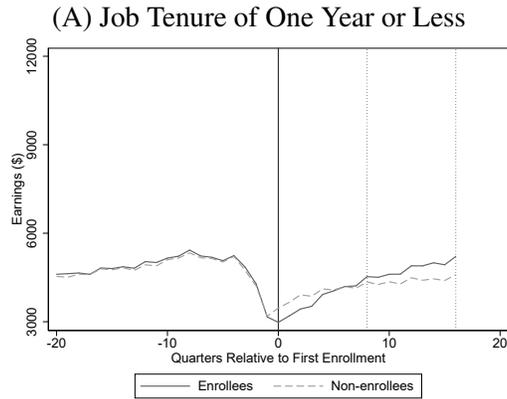


(B) Age 40 or Over

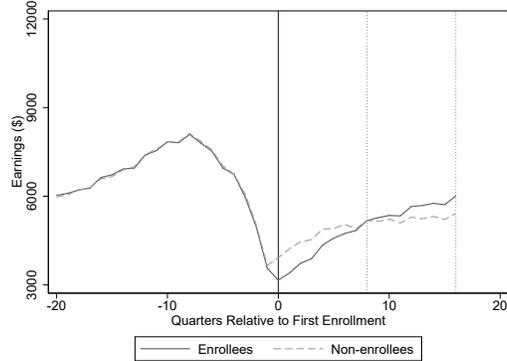


Notes: The upper (lower) graph shows the average quarterly earnings of enrollees under age 40 (age 40 or over) and matched non-enrollee UI claimants. The solid vertical lines denote the quarter of first enrollment, and the vertical dashed lines denote eight and 16 quarters after first enrollment. There are $N = 115,618$ (59,850) UI claims in the upper (lower) graph, corresponding to 109,541 (58,183) unique individuals.

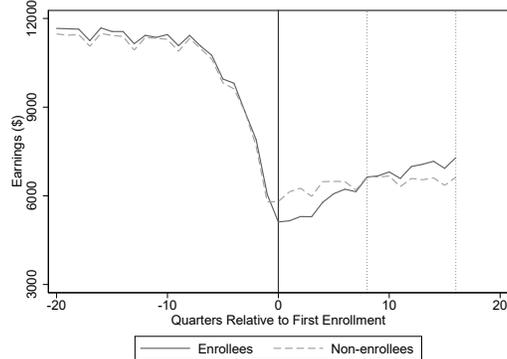
Figure A.12: Earnings of Enrollees and Matched Non-enrollees, By Tenure



(B) Job Tenure More Than One Year, Less Than Or Equal to Six Years



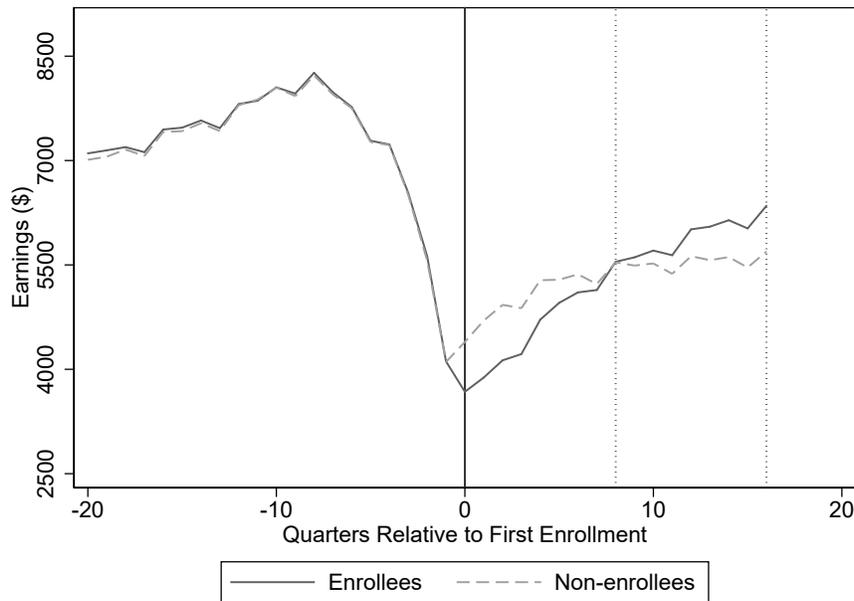
(C) Job Tenure More Than Six Years



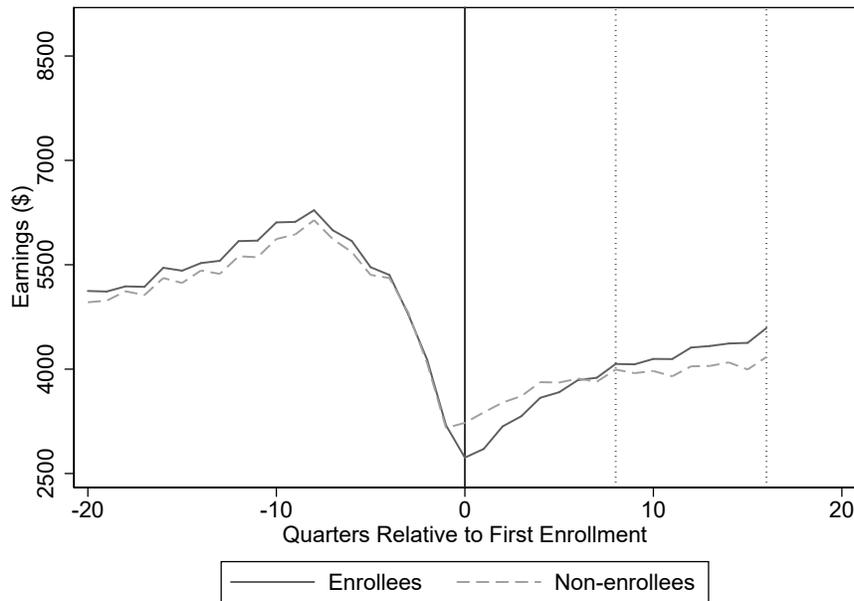
Notes: These graphs shows the average quarterly earnings of enrollees in different job tenure categories and matched non-enrollee UI claimants. The solid vertical lines denote the quarter of first enrollment, and the vertical dashed lines denote eight and 16 quarters after first enrollment. There are $N = 60,974$; $79,054$; and $35,440$ UI claims in the each graph, respectively, corresponding to $58,620$; $76,295$; and $34,146$ unique individuals.

Figure A.13: Earnings of Enrollees and Matched Non-enrollees, By Race

(A) White



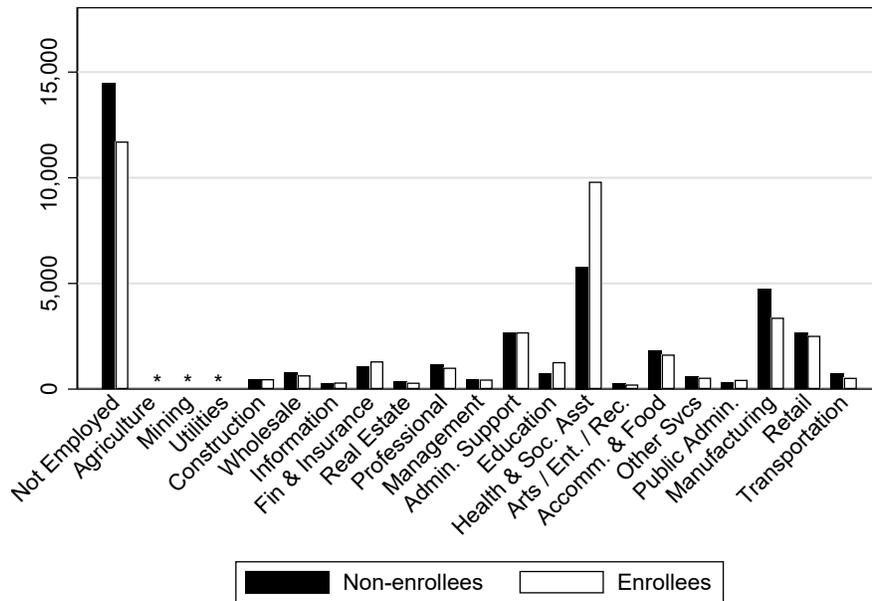
(B) Black



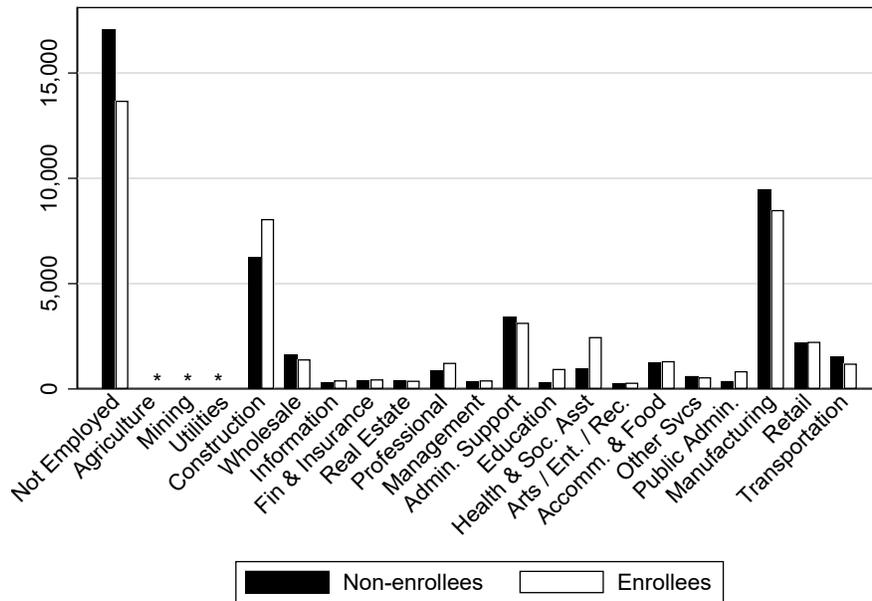
Notes: The upper (lower) graph shows the average quarterly earnings of white (Black) enrollees and matched non-enrollee UI claimants. The solid vertical lines denote the quarter of first enrollment, and the vertical dashed lines denote eight and 16 quarters after first enrollment. There are $N = 131,850$ (30,814) UI claims in the upper (lower) graph, corresponding to 126,178 (29,090) unique individuals.

Figure A.14: Industries of Enrollees and Matched Non-enrollees, 16th Quarter Post-Enrollment

(A) Women



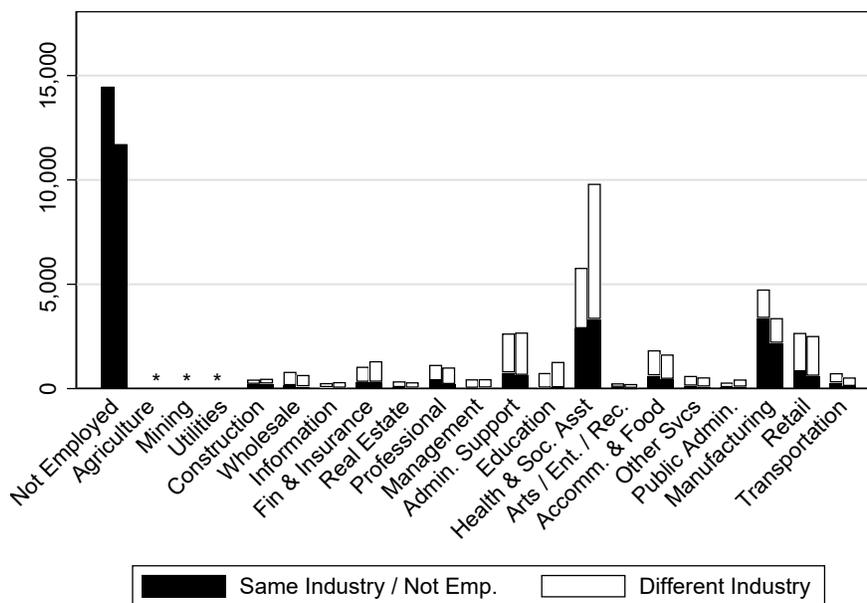
(B) Men



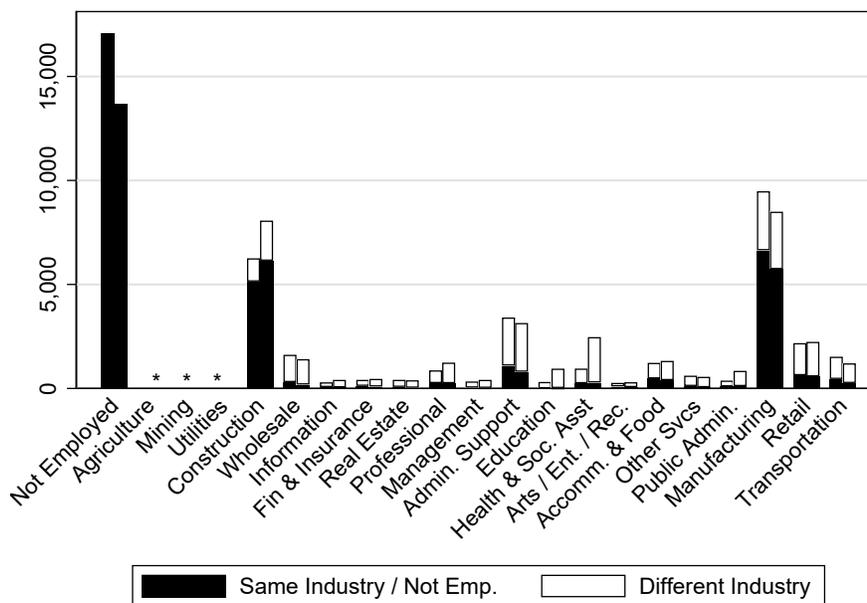
Notes: These figures plot the number of enrollee and matched non-enrollee UI claimants employed in each sector (two-digit NAICS) or not employed. Agriculture, Mining, and Utilities sectors have fewer than 250 workers in each enrollee/non-enrollee cell and are not plotted.

Figure A.15: Enrollees and Matched Non-enrollees Who Switched Industry, 16th Quarter Post-Enrollment

(A) Women

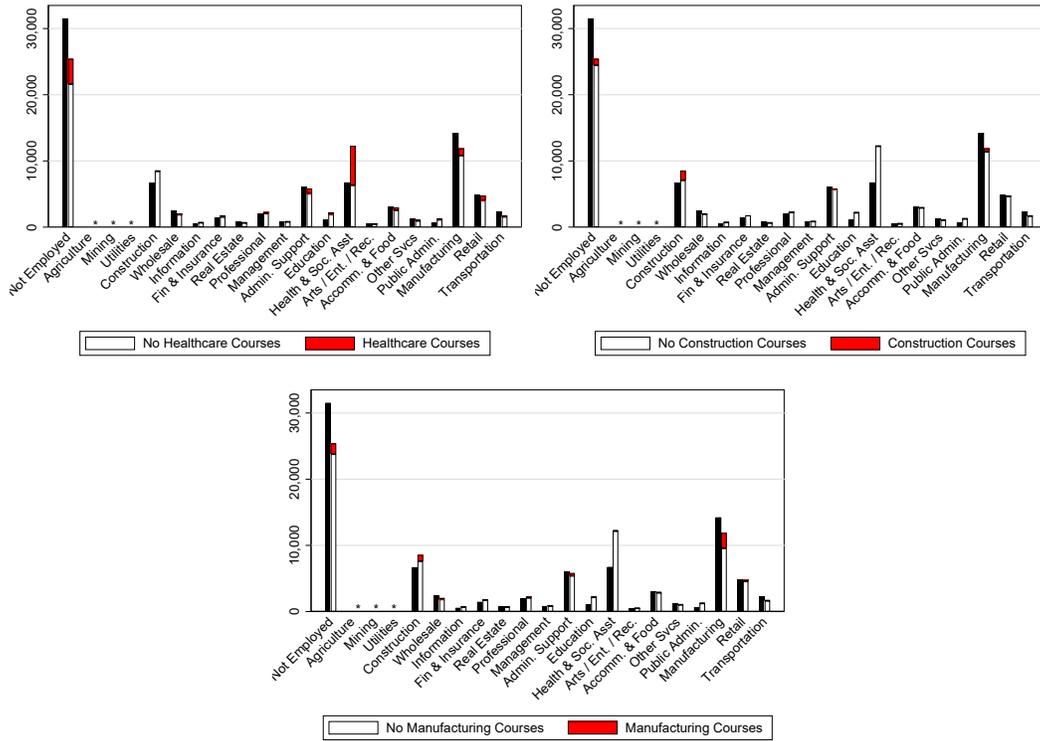


(B) Men



Notes: These figures plot the number of enrollees (right bar of each pair) and matched non-enrollees (left bar) that are employed in a different sector (two-digit NAICS) than their pre-layoff sector. Agriculture, Mining, and Utilities sectors have fewer than 250 workers in each enrollee/non-enrollee cell and are not plotted. If there are 10 or fewer enrollees/non-enrollees in a sector who have switched sectors, it is not plotted.

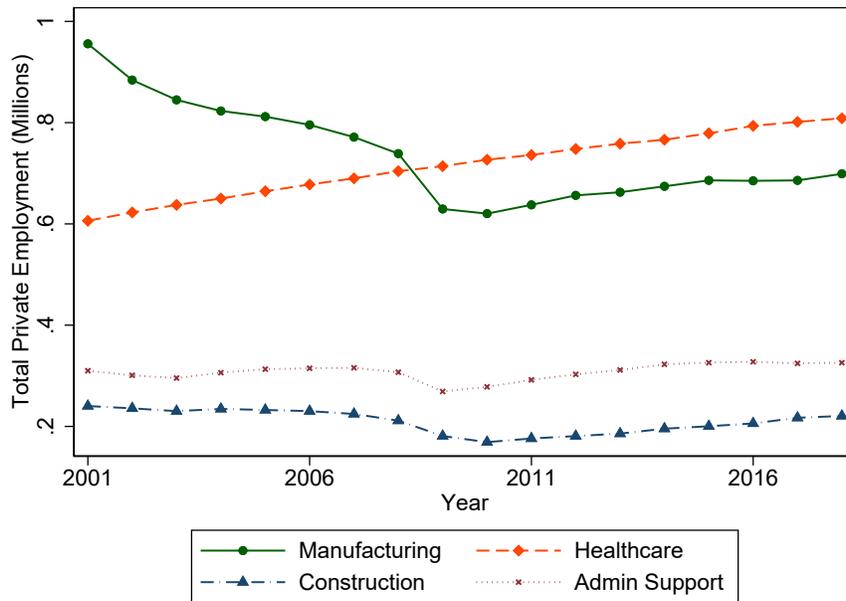
Figure A.16: Types of Courses Taken, by Industry of Employment at 16th Quarter Post-Enrollment



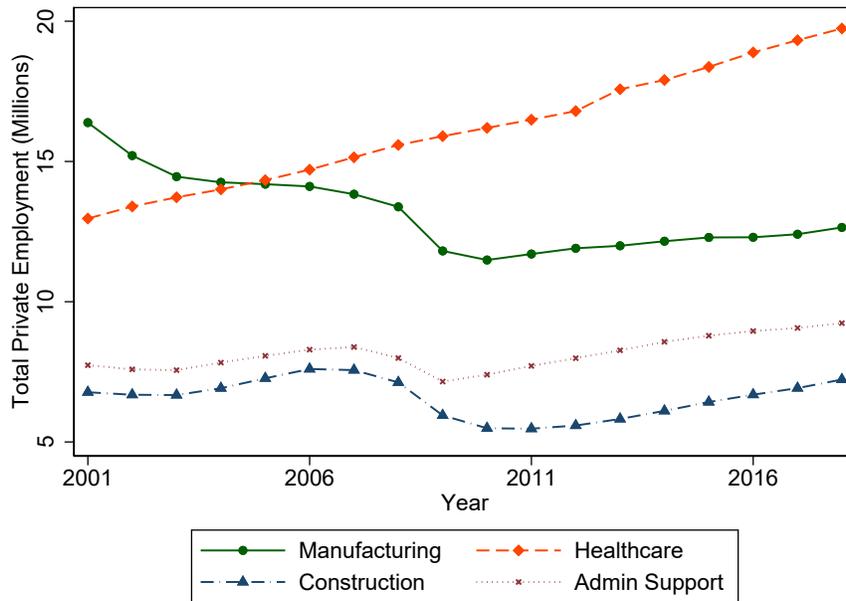
Notes: In each graph, the right-hand side bar of each pair plots the number of enrollees whose courses are associated with a certain industry. The left bar shows the number of matched non-enrollees in each sector. Agriculture, Mining, and Utilities sectors have fewer than 250 workers in each enrollee/non-enrollee cell and are not plotted. If there are 10 or fewer enrollees in a sector who have taken health, construction, or manufacturing courses, it is not plotted.

Figure A.17: Industry Employment Growth, Ohio and U.S.

(A) Ohio

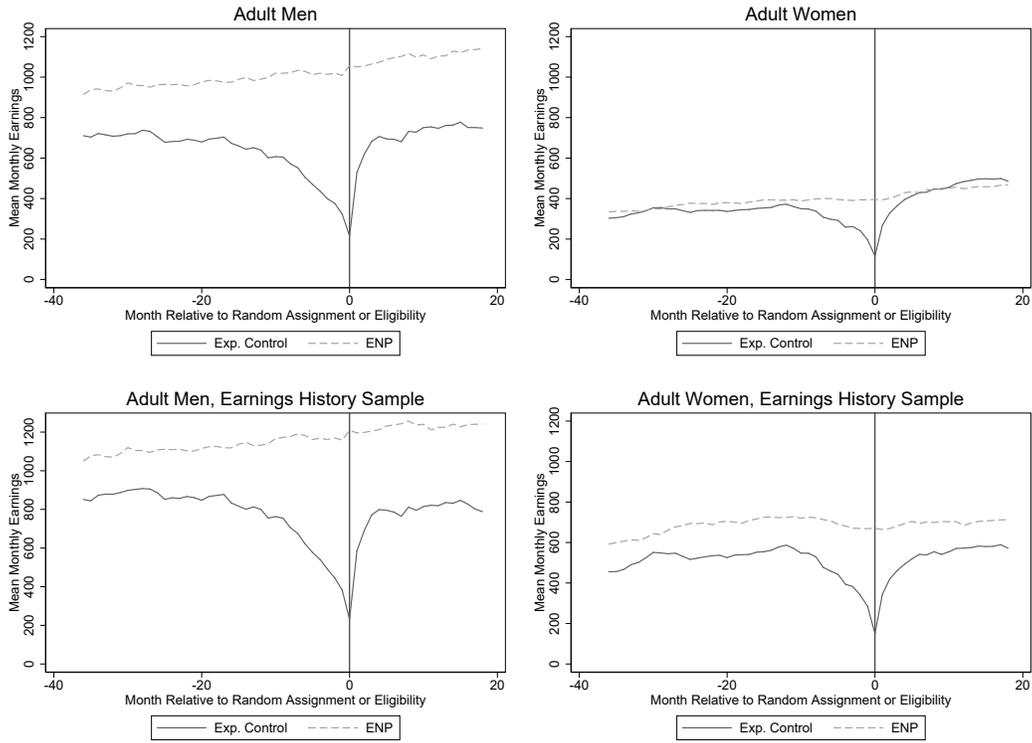


(B) U.S.



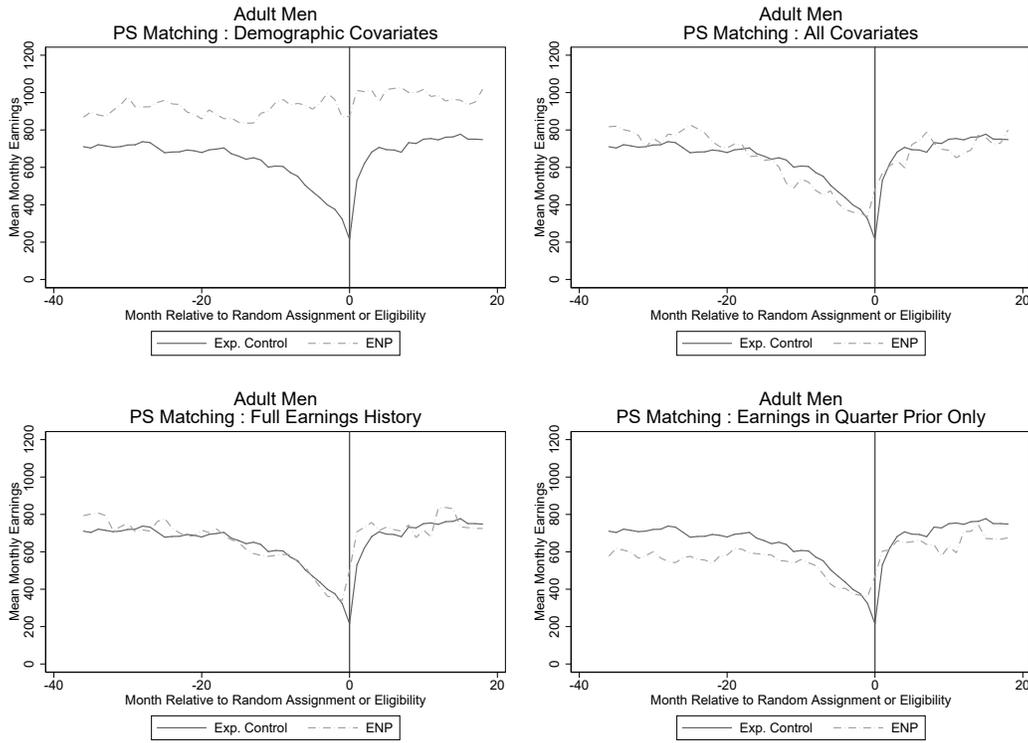
Notes: The upper figure plots the total private employment in the four largest sectors in Ohio over time. The lower figure plots the total private employment in the U.S. for the same industries.

Figure A.18: Earnings of NJS Validation Samples



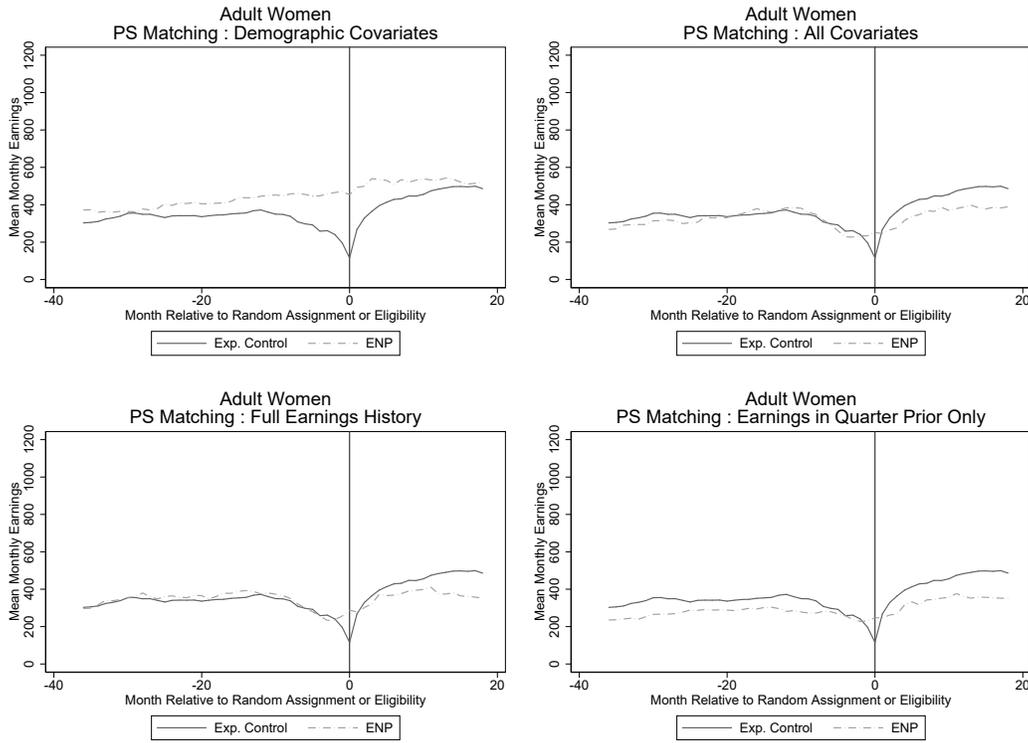
Notes: These plots show the mean monthly earnings of the experimental controls and eligible non-participants (ENPs) for each of the four samples used in our validation analysis.

Figure A.19: Earnings of NJS Men (Full Sample), Propensity-Score Matched



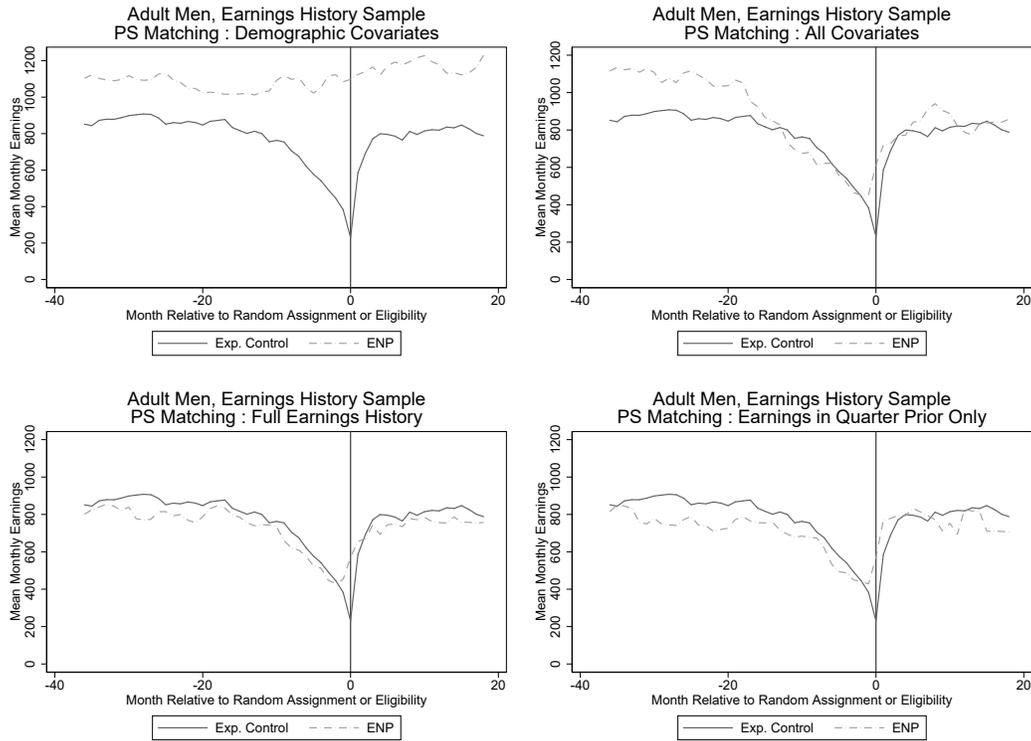
Notes: These plots show the mean monthly earnings of the experimental controls and matched eligible non-participants (ENPs) in the NJS Adult Men sample used our validation analysis. Matching is done via estimated propensity score, using one neighbor per experimental control observation. The propensity score is estimated using a logit model containing demographic characteristics only (upper left), demographic characteristics and 12 quarters of earnings (upper right), 12 quarters of earnings only (lower left), and one quarter of earnings only (lower right).

Figure A.20: Earnings of NJS Women (Full Sample), Propensity-Score Matched



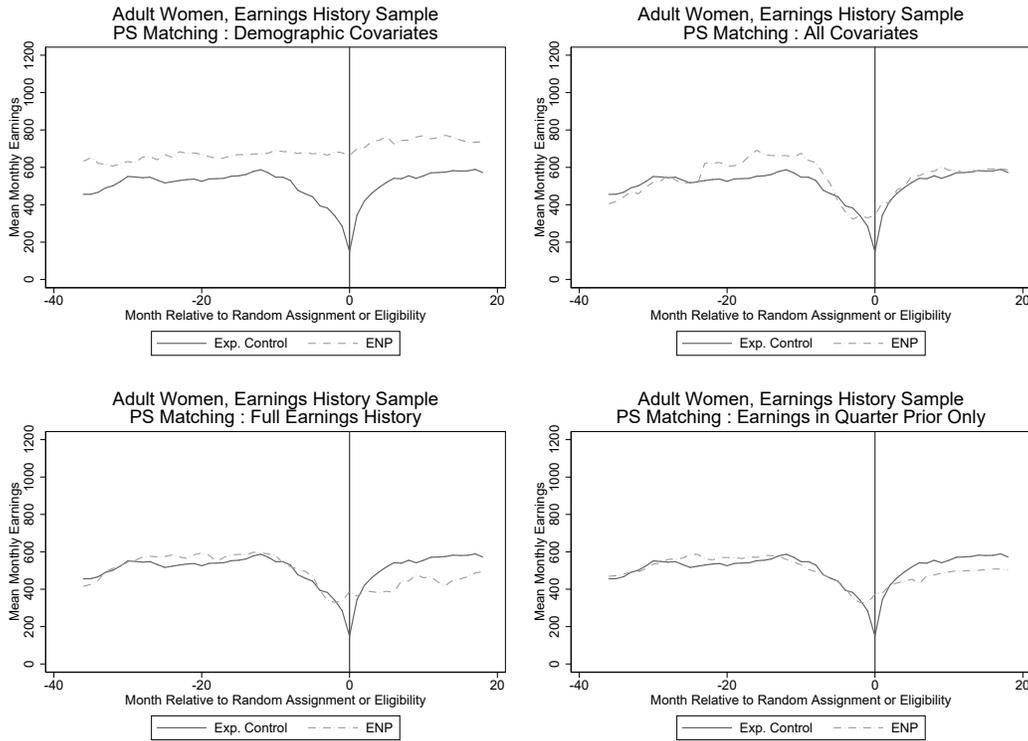
Notes: These plots show the mean monthly earnings of the experimental controls and matched eligible non-participants (ENPs) in the NJS Adult Women sample used in our validation analysis. Matching is done via estimated propensity score, using one neighbor per experimental control observation. The propensity score is estimated using a logit model containing demographic characteristics only (upper left), demographic characteristics and 12 quarters of earnings (upper right), 12 quarters of earnings only (lower left), and one quarter of earnings only (lower right).

Figure A.21: Earnings of NJS Men (Earnings History Sample), Propensity-Score Matched



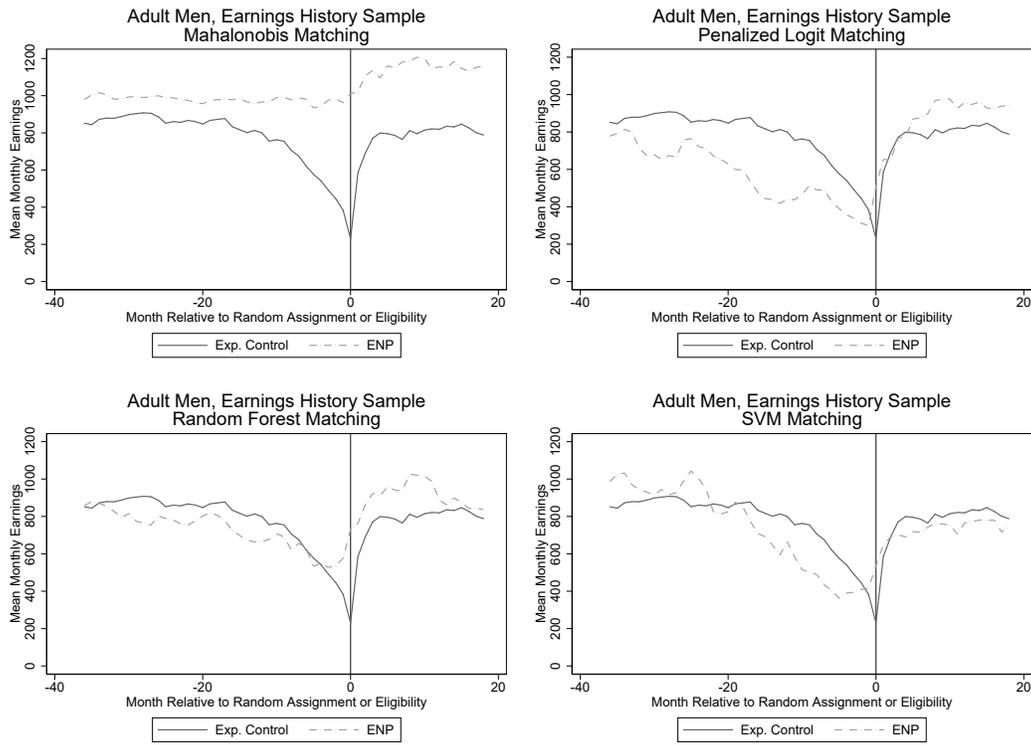
Notes: These plots show the mean monthly earnings of the experimental controls and matched eligible non-participants (ENPs) in the NJS Adult Men with Earnings History sample used in our validation analysis. Matching is done via estimated propensity score, using one neighbor per experimental control observation. The propensity score is estimated using a logit model containing demographic characteristics only (upper left), demographic characteristics and 12 quarters of earnings (upper right), 12 quarters of earnings only (lower left), and one quarter of earnings only (lower right).

Figure A.22: Earnings of NJS Women (Earnings History Sample), Propensity-Score Matched



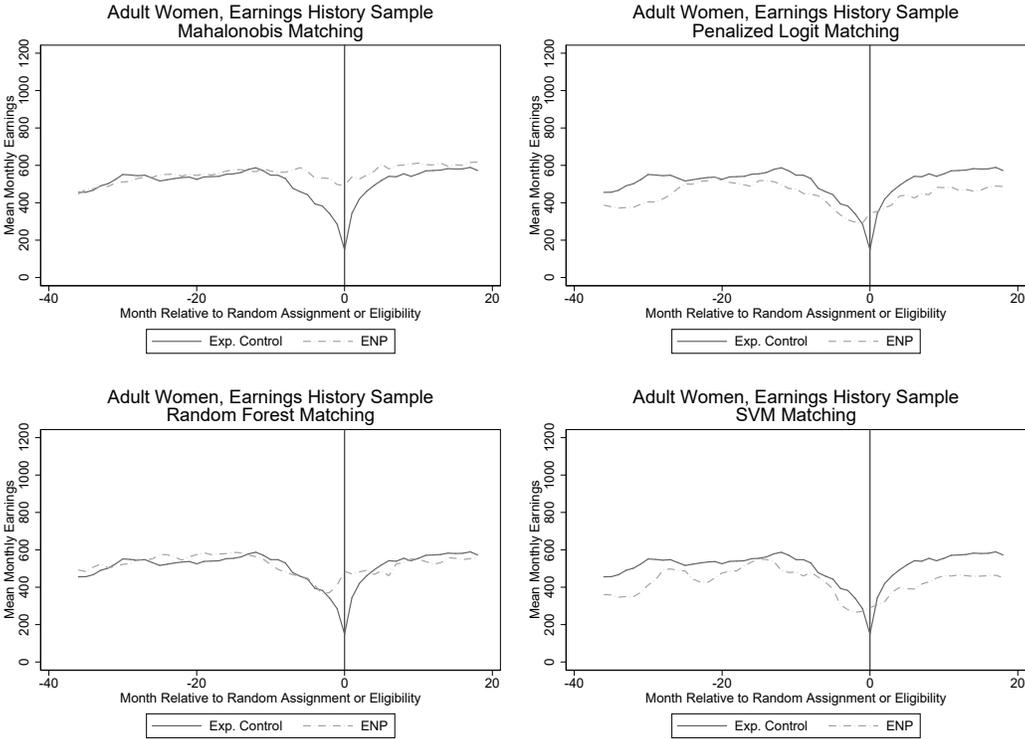
Notes: These plots show the mean monthly earnings of the experimental controls and matched eligible non-participants (ENPs) in the NJS Adult Women with Earnings History sample used in our validation analysis. Matching is done via estimated propensity score, using one neighbor per experimental control observation. The propensity score is estimated using a logit model containing demographic characteristics only (upper left), demographic characteristics and 12 quarters of earnings (upper right), 12 quarters of earnings only (lower left), and one quarter of earnings only (lower right).

Figure A.23: Earnings of NJS Men (Earnings History Sample), Selected Alternative Matching Methods



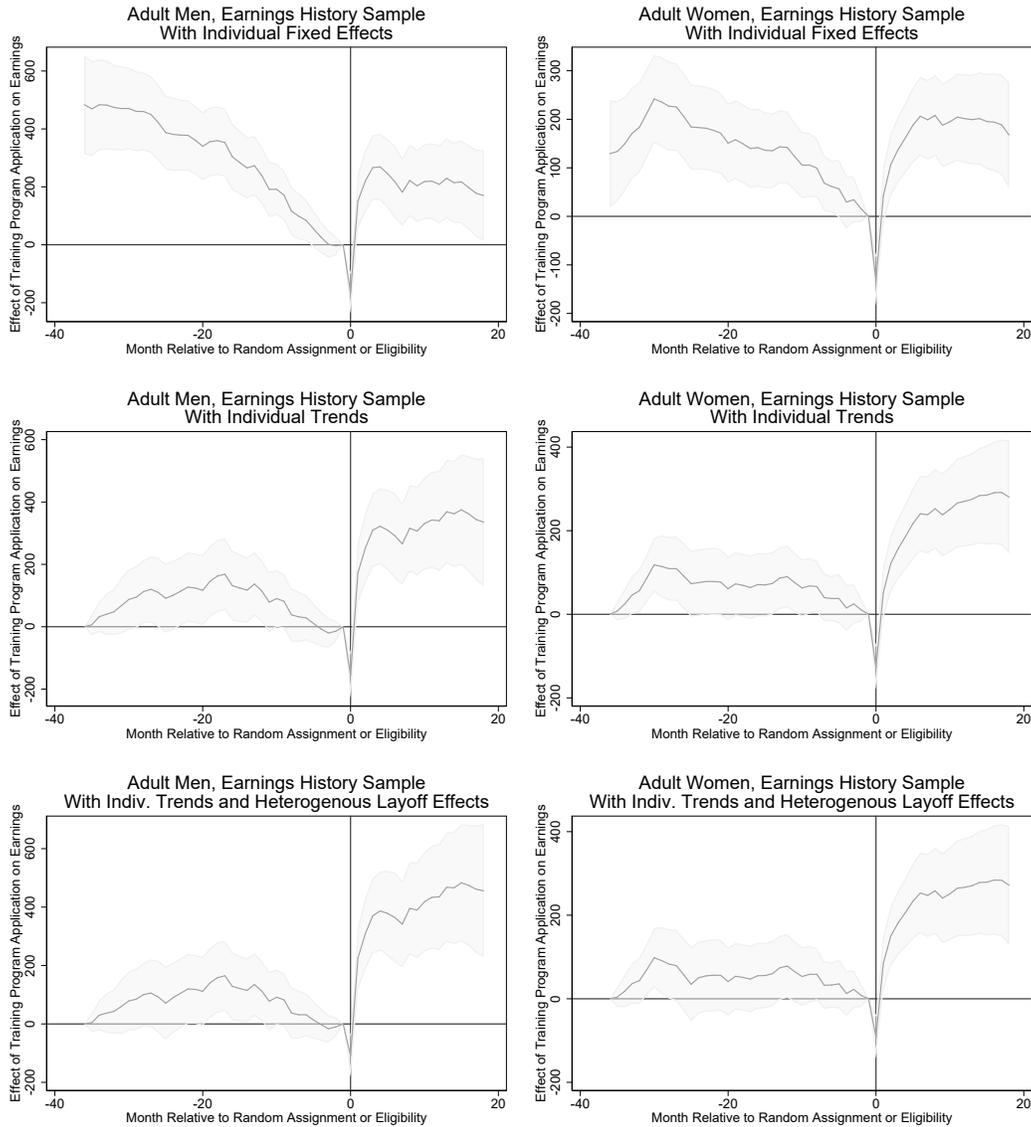
Notes: These plots show the mean monthly earnings of the experimental controls and matched eligible non-participants (ENPs) in the NJS Adult Men with Earnings History sample used in our validation analysis. The samples are matched using Mahalanobis distance (upper left), penalized logit propensity score matching (upper right), random forest matching (lower left), and support vector machine matching (lower right), and all models contain demographic characteristics and 12 quarters of pre-period earnings. Each experimental control observation is matched to one neighbor.

Figure A.24: Earnings of NJS Women (Earnings History Sample), Selected Alternative Matching Methods



Notes: These plots show the mean monthly earnings of the experimental controls and matched eligible non-participants (ENPs) in the NJS Adult Women with Earnings History sample used in our validation analysis. The samples are matched using Mahalanobis distance (upper left), penalized logit propensity score matching (upper right), random forest matching (lower left), and support vector machine matching (lower right), and all models contain demographic characteristics and 12 quarters of pre-period earnings. Each experimental control observation is matched to one neighbor.

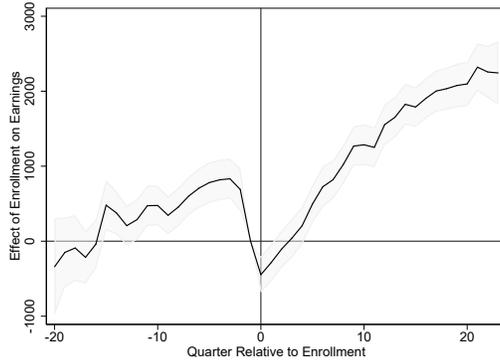
Figure A.25: “Event Study” Graphs for Fixed Effect Models (NJS Samples)



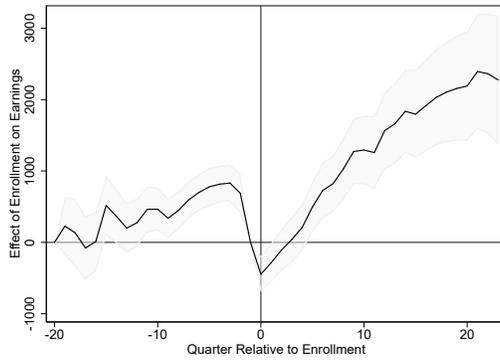
Notes: These figures show specification checks of fixed effect models using NJS validation analysis samples. Each graph plots the estimated earnings “effect” of belonging to the experimental control group relative to the ENP group. The first column corresponds to the NJS Adult Men with Earnings History sample and the second column corresponds to the Adult Women with Earnings History sample. The specifications in the first row include individual fixed effects, month fixed effects, and indicators for time relative to layoff. The specifications in the second row add individual time trends. The specifications in the third row add heterogeneous layoff effects. Shaded regions are 95 percent confidence intervals, where standard errors are clustered at the individual level.

Figure A.26: “Event Study” Graphs for Fixed Effect Models (Ohio Sample)

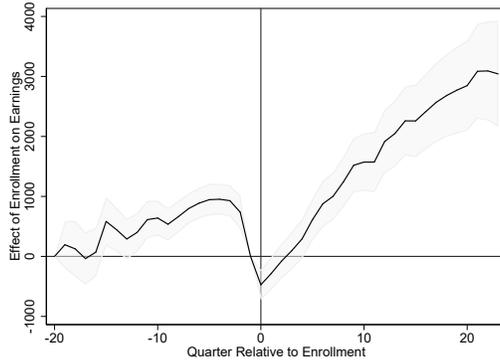
(A) With Individual Fixed Effects



(B) With Individual Trends

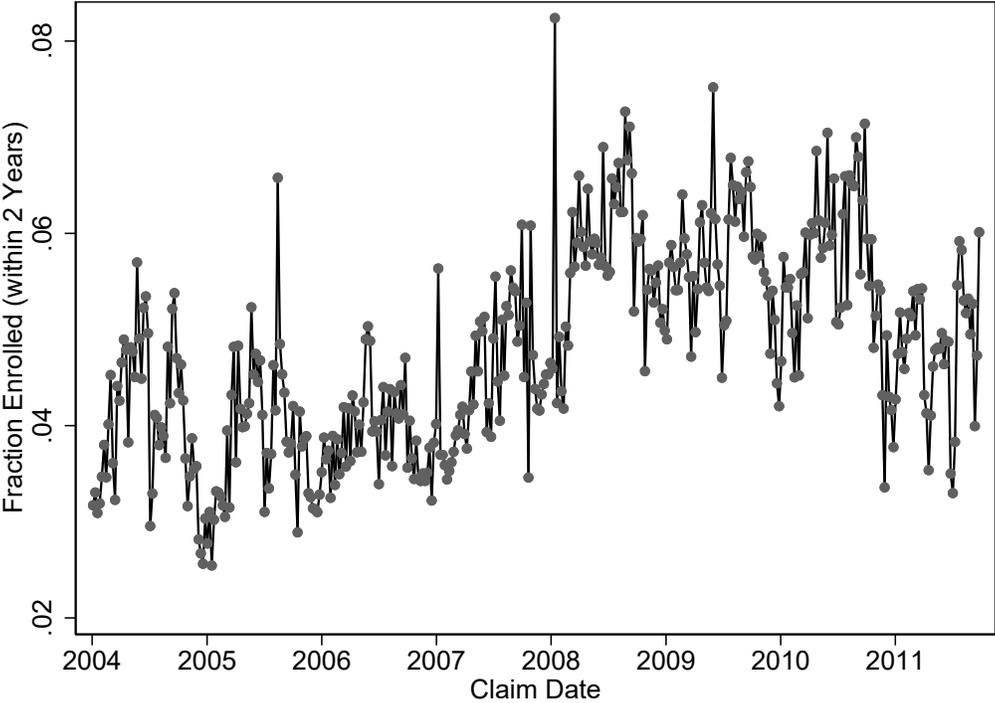


(C) With Individual Trends and Heterogeneous Layoff Effects



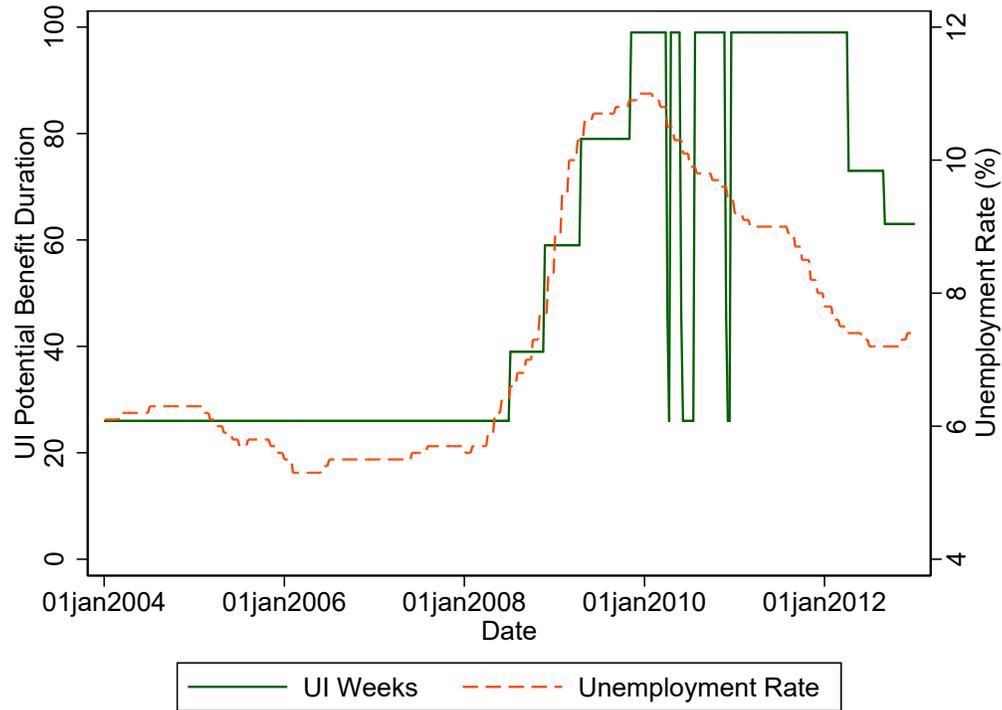
Notes: These figures show specification checks of fixed effect models using five percent of our main analysis sample. Each graph plots the estimated “effect” of enrollment on earnings. The model in Panel A includes individual fixed effects, quarter fixed effects, and indicators for time relative to layoff. The model in Panel B adds individual time trends. The model in Panel C adds heterogeneous layoff effects. Shaded regions are 95 percent confidence intervals, where standard errors are clustered at the UI claim level. $N = 99,478$ UI claims (corresponding to 96,874 unique individuals).

Figure A.27: Enrollment by UI Claim Date



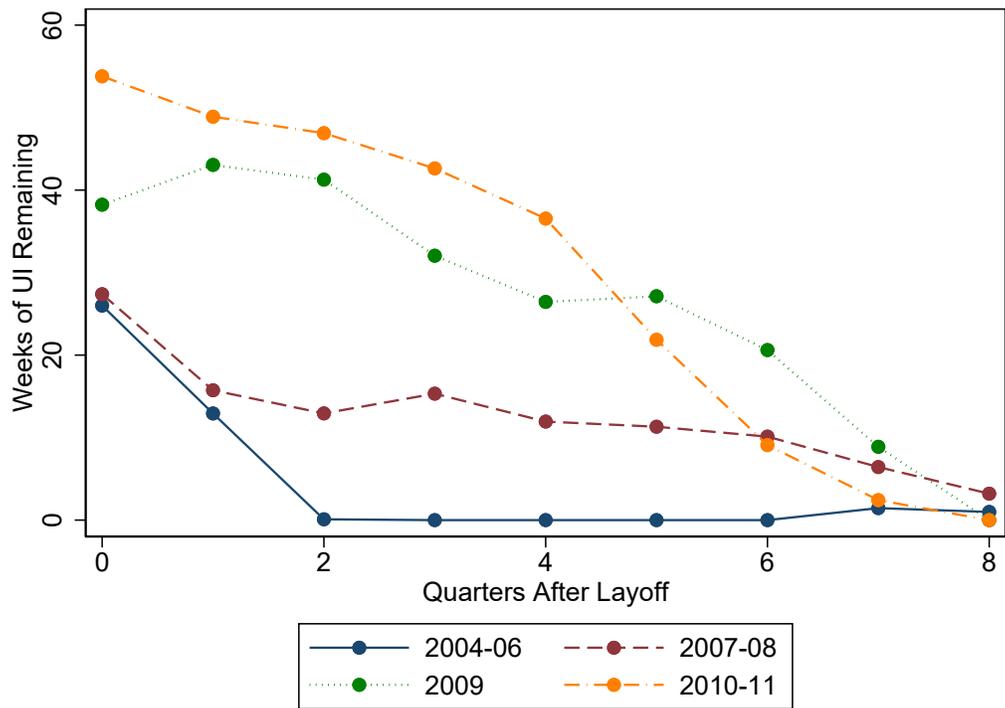
Notes: This figure plots the fraction of UI claimants who enroll within two years of claiming UI, by the date of the UI claim. $N = 1,994,777$, corresponding to 1,335,958 unique individuals.

Figure A.28: Ohio UI Extensions, 2004-2012



Notes: This figure shows the statutory UI benefit duration in weeks (left axis) and the unemployment rate (right axis) in Ohio.

Figure A.29: Simulated Weeks of UI Remaining to Workers of Different UI Claim Cohorts



Notes: This figure shows the simulated number of weeks of UI benefits remaining to various cohorts of UI claimants in five percent of our analysis sample. $N = 99,478$ UI claims (corresponding to 96,874 unique individuals).

Table A.1: Pre-enrollment Earnings Differences, By Subgroup

Subgroup	Quarterly Earnings				No. of Enrollees (5)
	1-2 Yrs Pre-Enrollment		3-4 Yrs Pre-Enrollment		
	TOT Estimate (1)	t-stat (2)	TOT Estimate (3)	t-stat (4)	
All	37 (23)	1.61	19 (27)	0.71	87663
Quarters From Layoff to Enrollment					
1	122 (65)	1.86	52 (69)	0.76	14519
2	39 (54)	0.73	54 (63)	0.85	17021
3	-32 (55)	-0.58	-54 (68)	-0.79	14440
4	44 (60)	0.74	8 (74)	0.11	11610
5	12 (66)	0.19	51 (84)	0.61	9395
6	67 (70)	0.95	9 (81)	0.12	8102
7	18 (77)	0.23	-28 (80)	-0.35	6947
8	-2 (91)	-0.02	54 (84)	0.64	5629
Layoff Year					
2004	79 (72)	1.10	-13 (78)	-0.17	7603
2005	97 (91)	1.06	172 (97)	1.77	5783
2006	-78 (88)	-0.89	-93 (88)	-1.06	8395
2007	68 (76)	0.89	42 (80)	0.52	9137
2008	51 (49)	1.04	61 (59)	1.03	17067
2009	5 (46)	0.10	7 (57)	0.12	21106
2010	59 (64)	0.92	-11 (76)	-0.15	11872
2011	69 (81)	0.85	18 (98)	0.18	6700
Male	45 (34)	1.32	29 (39)	0.74	48168
Female	28 (30)	0.92	6 (34)	0.18	39495
Manufacturing	39 (51)	0.76	22 (54)	0.41	24914
Non-manuf.	36 (25)	1.45	18 (30)	0.58	62749
Age <40	37 (23)	1.57	106 (26)	4.12	57809
Age >=40	27 (47)	0.57	5 (54)	0.09	29925
Tenure					
<=1 Year	46 (32)	1.42	60 (37)	1.65	30487
1-6 Years	-70 (33)	-2.14	-52 (38)	-1.37	39527
>6 Years	170 (63)	2.68	113 (67)	1.69	17720
White	27 (27)	0.99	26 (31)	0.82	65925
Black	27 (47)	0.58	135 (50)	2.68	15407

Notes: This table presents balance tests for enrollees and matched non-enrollees within subgroups. Columns (1) and (3) show the difference between enrollees and matched non-enrollees in the two years prior to enrollment and three to four years prior to enrollment, respectively. Standard errors (in parentheses) and t-statistics for the mean pairwise difference between enrollees and matched-non-enrollees are reported.

Table A.2: Characteristics of WIA Participant Subsample

Demographic and Pre-Layoff Job Characteristics	
Female	49%
Race	
White	82%
Black	12%
Other	2%
Unknown	4%
Prior Industry	
Manufacturing	50%
Construction	4%
Admin. Support & Waste	8%
Healthcare and Social Assistance	6%
Retail Trade	7%
Accommodation and Food Services	2%
Wholesale Trade	4%
Transportation	6%
Tenure at Recent Employer	
<=1 year	26%
>1 years to <=3 years	28%
>3 years to <=6 years	17%
>6 years	29%
Age	40.58
Cty Unempl. Rate at Layoff	8.57%
Earnings	
1 year before layoff	37872 [21202]
2 years before layoff	36168 [21794]
3 years before layoff	35332 [22803]
Enrollment Characteristics	
Time from job loss to enrollment (quarters)	3.3
Terms/Quarters Enrolled	5.3
Type of Institution Attended	
Technical Center	41%
Community College	56%
Four-year Institution	11%
Types of Courses	
Taken at least one occupational course	73%
Avg proportion of courses occupational	0.8
Credential	58%
Bachelors	2%
Associate	18%
Less Than Associate	38%
Observations	9,765

Notes: This table presents descriptive characteristics for the subgroup of enrollees who received WIA training services. Type of Institution Attended, Terms/Quarters Enrolled, Types of Courses, and Credential are calculated within four years of first enrollment. Enrollees may attend more than one type of institution over the four-year period. “Less than Associate” credentials include less than two-year awards from HEI and any credential from OTC. Standard deviations are in brackets.

Table A.3: Summary Statistics for NJS Validation Samples

	Adult Men, All			Adult Women, All			Adult Men, Earnings History			Adult Women, Earnings History		
	ENP	Exp Control		ENP	Exp Control		ENP	Exp Control		ENP	Exp Control	
Corpus Christi, TX	42.0%	17.4%		32.2%	22.1%		42.6%	17.9%		33.0%	21.9%	
Fort Wayne, IN	33.3%	53.8%		29.2%	41.0%		35.3%	56.4%		36.9%	48.0%	
Jersey City, NJ	12.6%	14.4%		18.7%	22.3%		11.1%	14.5%		12.6%	21.9%	
Providence, RI	12.0%	14.4%		19.9%	14.5%		11.1%	11.3%		17.5%	8.2%	
White	39.3%	53.5%		38.7%	39.3%		41.2%	55.8%		44.4%	43.5%	
Black	12.3%	26.8%		18.8%	30.2%		11.8%	26.0%		18.7%	30.3%	
Hispanic	42.9%	16.0%		38.1%	28.4%		42.9%	16.2%		33.5%	24.5%	
Other Race	5.4%	3.7%		4.4%	2.0%		4.2%	2.0%		3.4%	1.6%	
Age 22-29	32.1%	43.5%		36.5%	44.6%		32.5%	44.8%		33.0%	47.5%	
Age 30-39	40.5%	40.3%		39.7%	37.7%		40.8%	40.8%		44.2%	36.1%	
Age 40-49	21.9%	13.5%		18.1%	14.9%		21.5%	11.8%		18.9%	13.2%	
Age 50-54	5.4%	2.7%		5.7%	2.8%		5.2%	2.6%		3.9%	3.2%	
Age	34.39	31.97		33.76	31.98		34.28	31.63		33.81	31.60	
Education												
Less than 10 Yrs	31.8%	22.4%		33.1%	22.6%		29.4%	18.5%		26.5%	17.4%	
10-11 Yrs	18.0%	22.0%		19.4%	20.7%		19.0%	24.0%		16.3%	20.6%	
12 Yrs	28.8%	35.5%		34.2%	40.2%		31.8%	36.1%		41.3%	43.8%	
13-15 Yrs	13.2%	15.6%		10.6%	14.0%		13.8%	17.1%		12.6%	15.0%	
>15 Yrs	8.1%	4.6%		2.7%	2.5%		5.9%	4.3%		3.4%	3.2%	
Currently Married	62.2%	29.5%		42.2%	17.5%		63.7%	31.2%		45.1%	19.3%	
Has a child under age 6	23.4%	11.4%		21.9%	17.0%		24.9%	12.1%		20.1%	16.6%	
Pre-period Earnings	11539	8352		4393	4079		13264	10368		8066	6327	
[std. dev.]	[9388]	[7943]		[6490]	[5153]		[8885]	[7718]		[7032]	[5342]	
Observations	333	437		770	605		289	346		412	379	

Notes: This table presents descriptive statistics for the four samples used in our validation analysis. “ENP” denotes eligible non-participants and “Exp Control” denotes experimental control observations. Pre-period earnings are annual nominal averages over the period of 13-36 months prior to random assignment (for experimental controls) or eligibility determination (for ENPs). We exclude prior months 1-12 for comparability to the annual pre-layoff earnings in our Ohio analysis.

Table A.4: Matching Validation Exercise Results for Full Adult Samples

	Demographic Covariates				Demographic Covariates- Full Earnings History				Full Earnings History				Earnings in quarter prior			
	Neighbours	Bias	S.E. (1)	t-stat	Bias	S.E. (2)	t-stat	Bias	S.E. (3)	t-stat	Bias	S.E. (4)	t-stat	Bias	S.E. (4)	t-stat
Panel A: Adult Men (N=770)																
Simple Difference (no covariates)		-386	56	-6.90												
Ordinary Least Squares		-310	63	-4.94												
Logit-Based Propensity Score Matching (L-PSM)	1	-295	76	-3.87	-37	62	-0.60	-81	55	-1.48	-55	57	-0.96	78	60	1.03
Logit-Based Propensity Score Matching (L-PSM)	5	-366	69	-5.30	3	76	0.03	-21	101	-0.21	101	76	1.69	101	60	1.69
Difference-in-Differences Matching	1	63	68	0.93	-31	71	-0.44	14	69	0.20	14	69	-0.01	74	61	0.11
Difference-in-Differences Matching	5	17	67	0.26	-37	47	-0.79	-13	67	-0.36	-13	67	0.11	7	71	0.98
Mahalanobis Matching	1	-351	71	-4.98	-266	79	-3.35	-103	100	-1.03	70	71	1.69	70	71	1.69
Mahalanobis Matching	5	-382	59	-6.45	-276	61	-4.50	-51	81	-0.63	99	58	1.69	99	58	1.69
Inverse Variance Matching	1	-357	70	-5.10	-112	80	-1.40	-76	97	-0.78	70	97	0.98	70	97	0.98
Inverse Variance Matching	5	-367	58	-6.28	-102	62	-1.66	-35	76	-0.46	99	58	1.69	99	58	1.69
Inverse Probability Weighting	1	-335	67	-5.01	37	95	0.39	65	62	0.39	65	62	1.10	63	63	1.10
L-PSM Using Imbens-Rubin Covariate Selector	1	-323	82	-3.93	77	68	1.14	-18	94	-0.19	78	76	1.03	78	76	1.03
L-PSM Using Imbens-Rubin Covariate Selector	5	-293	66	-4.43	0	85	0.00	52	61	0.85	100	60	1.68	100	60	1.68
Penalized Logit-Based Propensity Score Matching	1	-369	86	-4.28	130	100	1.30	62	87	0.71	63	71	0.89	63	71	0.89
Penalized Logit-Based Propensity Score Matching	5	-339	80	-4.23	56	112	0.50	29	76	0.38	103	58	1.76	103	58	1.76
GenMatch	1	-267	79	-3.39	8	87	0.09	-10	80	-0.12	65	67	0.96	65	67	0.96
GenMatch	5	-316	71	-4.48	-48	72	-0.67	10	81	0.12	99	58	1.69	99	58	1.69
Classification Tree, Pruned	1	-291	67	-4.31	-90	93	-0.96	18	63	0.28	18	63	0.28	18	63	0.28
Classification Tree, Pruned	5	-291	67	-4.31	8	91	0.09	18	63	0.28	18	63	0.28	18	63	0.28
Classification Tree, Bagged	1	-421	78	-5.40	13	78	0.16	-40	98	-0.41	-70	83	-0.85	83	-0.85	-0.85
Classification Tree, Bagged	5	-340	69	-4.91	-8	72	-0.11	-114	90	-1.26	-10	61	-0.16	-10	61	-0.16
Classification Tree, Random Forest	1	-311	83	-3.72	71	86	0.82	82	88	-1.32	-103	71	-1.45	-103	71	-1.45
Classification Tree, Random Forest	5	-315	74	-4.17	-29	83	-0.35	-134	77	-1.75	-25	60	-0.43	-25	60	-0.43
Classification Tree, Boosted	1	-270	74	-3.63	31	122	0.26	124	73	1.72	74	66	1.11	74	66	1.11
Classification Tree, Boosted	5	-305	67	-4.55	31	114	0.27	95	66	1.43	72	65	1.12	72	65	1.12
Support Vector Machine	1	-318	84	-3.77	424	187	2.26	110	100	1.10	81	76	1.07	81	76	1.07
Support Vector Machine	5	-345	80	-4.32	141	251	0.56	150	80	1.87	60	68	0.89	60	68	0.89
Panel B: Adult Women (N=1375)																
Simple Difference (no covariates)		-5	31	-0.15												
Ordinary Least Squares		-46	32	-1.45												
Logit-Based Propensity Score Matching (L-PSM)	1	-54	41	-1.31	47	29	1.63	74	27	2.71	94	28	3.36	94	28	3.36
Logit-Based Propensity Score Matching (L-PSM)	5	-25	38	-0.66	79	39	2.03	73	44	1.67	115	28	4.09	115	28	4.09
Difference-in-Differences Matching	1	53	42	1.26	86	35	2.49	88	44	2.02	68	31	2.17	68	31	2.17
Difference-in-Differences Matching	5	60	36	1.66	61	36	1.72	93	31	2.99	55	29	1.91	55	29	1.91
Mahalanobis Matching	1	-46	45	-1.02	88	60	1.47	119	31	3.79	115	28	4.09	115	28	4.09
Mahalanobis Matching	5	-53	38	-1.42	85	40	2.11	110	28	3.87	118	27	4.38	118	27	4.38
Inverse Variance Matching	1	-45	43	-1.04	47	42	1.12	92	35	2.36	115	28	4.12	115	28	4.12
Inverse Variance Matching	5	-51	38	-1.35	52	33	1.56	84	29	2.99	116	27	4.27	116	27	4.27
Inverse Probability Weighting	1	-58	36	-1.61	67	32	2.12	87	29	2.99	116	27	4.27	116	27	4.27
L-PSM Using Imbens-Rubin Covariate Selector	1	-59	42	-1.40	92	35	2.61	80	34	2.36	115	28	4.09	115	28	4.09
L-PSM Using Imbens-Rubin Covariate Selector	5	-83	39	-2.14	95	25	3.79	88	29	2.98	118	27	4.32	118	27	4.32
Penalized Logit-Based Propensity Score Matching	1	-49	37	-1.33	97	39	2.51	-18	77	-0.23	-5	30	-0.15	-5	30	-0.15
Penalized Logit-Based Propensity Score Matching	5	-53	37	-1.41	103	33	3.13	89	39	2.25	-5	30	-0.15	-5	30	-0.15
GenMatch	1	-71	39	-1.80	42	41	1.03	109	34	3.26	112	27	4.10	112	27	4.10
GenMatch	5	-68	38	-1.77	47	36	1.31	109	34	3.26	112	27	4.10	112	27	4.10
Classification Tree, Pruned	1	-51	35	-1.45	95	32	1.03	31	31	1.09	118	27	4.38	118	27	4.38
Classification Tree, Pruned	5	-51	35	-1.45	95	32	1.03	31	31	1.09	118	27	4.38	118	27	4.38
Classification Tree, Bagged	1	-45	47	-0.96	88	58	1.52	81	30	2.68	101	26	3.83	101	26	3.83
Classification Tree, Bagged	5	-45	47	-0.96	88	58	1.52	81	30	2.68	101	26	3.83	101	26	3.83
Classification Tree, Random Forest	1	-31	43	-0.71	79	39	2.03	62	35	1.76	57	34	1.68	57	34	1.68
Classification Tree, Random Forest	5	-31	43	-0.71	79	39	2.03	62	35	1.76	57	34	1.68	57	34	1.68
Classification Tree, Boosted	1	-41	38	-1.07	66	33	2.02	48	31	1.55	66	31	2.13	66	31	2.13
Classification Tree, Boosted	5	-41	38	-1.07	66	33	2.02	48	31	1.55	66	31	2.13	66	31	2.13
Support Vector Machine	1	-22	36	-0.61	111	32	3.49	88	37	2.39	128	27	4.88	128	27	4.88
Support Vector Machine	5	-55	40	-1.37	116	38	3.10	115	48	2.41	-3	32	-0.08	-3	32	-0.08
Support Vector Machine	5	-37	39	-0.96	87	33	2.65	88	54	1.63	6	29	0.21	6	29	0.21

Notes: This table shows the estimated monthly earnings difference over 18 months after random assignment between experimental controls and matched/reweighted eligible non-participants (ENPs), using estimators considered in our validation exercise, for the Adult Men (Panel A) and Adult Women (Panel B) samples. The following covariate sets were used for each estimator: demographic characteristics only (col. 1), demographic characteristics and 12 quarters of earnings (col. 2), 12 quarters of earnings only (col. 3), and one quarter of earnings only (col.4). “Bias” is the estimated difference and “SE” denotes standard error.

Table A.5: Matching Validation Exercise Results for Earnings History Samples

	Demographic Covariates				Demographic Covariates- Full Earnings History				Full Earnings History				Earnings in quarter prior			
	Neighbours	Bias	S.E. (1)	t-stat	Bias	S.E. (2)	t-stat	Bias	S.E. (3)	t-stat	Bias	S.E. (4)	t-stat	Bias	S.E. (4)	t-stat
Panel A: Adult Men (N=635)																
Simple Difference (no covariates)		-439	63	-6.94												
Ordinary Least Squares		-358	68	-5.23												
Logit-Based Propensity Score Matching (L-PSM)	1	-381	75	-5.09	-100	72	-1.38	-143	66	-2.16	-124	68	-1.81			
Logit-Based Propensity Score Matching (L-PSM)	5	-419	72	-5.78	-40	81	-0.49	47	92	0.50	30	95	0.32			
Difference-in-Differences Matching	1	-9	94	-0.09	-33	84	-0.39	21	76	0.28	48	81	0.28			
Difference-in-Differences Matching	5	-23	90	-0.25	-54	52	-1.04	-43	73	-0.58	-9	87	-0.10			
Mahalanobis Matching	1	-386	77	-5.04	-360	75	-4.79	-129	117	-1.10	20	89	0.23			
Mahalanobis Matching	5	-422	66	-6.41	-390	61	-6.38	-150	111	-1.35	52	80	0.65			
Inverse Variance Matching	1	-370	75	-4.91	-195	84	-2.32	-75	112	-0.67	20	89	0.23			
Inverse Variance Matching	5	-362	66	-5.91	-210	68	-3.10	-85	90	-0.94	52	80	0.65			
Inverse Probability Weighting	1	-368	73	-5.07	-19	99	-0.20	8	81	0.09	-14	84	-0.17			
L-PSM Using Imbens-Rubin Covariate Selector	1	-380	101	-3.75	-124	125	-1.00	-46	94	-0.49	34	95	0.36			
L-PSM Using Imbens-Rubin Covariate Selector	5	-380	89	-4.29	-78	95	-0.82	73	89	0.82	46	80	0.58			
Penalized Logit-Based Propensity Score Matching	1	-423	85	-5.01	-97	115	-0.84	40	139	-0.84	14	88	0.15			
Penalized Logit-Based Propensity Score Matching	5	-328	80	-4.12	-79	150	-0.52	40	100	-0.40	50	80	0.62			
GenMatch	1	-334	79	-4.25	-70	91	-0.77	-72	109	-0.66	7	88	0.08			
GenMatch	5	-341	72	-4.74	-43	77	-0.56	-31	100	-0.31	21	81	0.26			
Classification Tree, Pruned	1	-331	72	-4.57	-121	97	-1.25	95	95	0.09	8	80	-0.85			
Classification Tree, Pruned	5	-331	72	-4.57	-121	97	-1.25	95	95	0.09	8	80	-0.85			
Classification Tree, Bagged	1	-491	90	-5.47	-185	114	-1.62	-89	132	-0.67	-196	94	-2.08			
Classification Tree, Bagged	5	-410	80	-5.10	-91	93	-0.99	-117	101	-1.15	-89	82	-1.08			
Classification Tree, Random Forest	1	-424	87	-4.87	-123	126	-0.97	-177	113	-1.57	-157	104	-1.51			
Classification Tree, Random Forest	5	-369	73	-5.07	-122	110	-1.11	-57	80	-0.71	-96	79	-1.20			
Classification Tree, Boosted	1	-328	78	-4.23	-8	136	-0.06	64	108	0.59	3	97	0.03			
Classification Tree, Boosted	5	-379	79	-4.82	-14	134	-0.10	-13	102	-0.12	-2	95	-0.02			
Support Vector Machine	1	-365	78	-4.70	54	123	0.44	218	528	0.41	-5	102	-0.05			
Support Vector Machine	5	-409	75	-5.49	-109	107	-1.01	105	203	0.52	22	95	0.23			
Panel B: Adult Women (N=791)																
Simple Difference (no covariates)		-179	45	-3.95												
Ordinary Least Squares		-192	45	-4.22												
Logit-Based Propensity Score Matching (L-PSM)	1	-192	66	-2.91	-10	42	-0.23	15	39	0.39	19	41	0.47			
Logit-Based Propensity Score Matching (L-PSM)	5	-158	56	-2.83	-21	80	-0.26	99	63	1.57	59	43	1.39			
Difference-in-Differences Matching	1	-2	63	-0.03	45	72	0.62	115	64	0.40	66	44	1.63			
Difference-in-Differences Matching	5	9	64	0.17	52	77	0.68	79	43	1.84	34	44	1.09			
Mahalanobis Matching	1	-176	61	-2.88	-59	64	-0.92	61	48	1.28	59	43	1.39			
Mahalanobis Matching	5	-155	56	-2.75	-111	70	-1.58	65	42	1.56	67	40	1.69			
Inverse Variance Matching	1	-156	72	-2.16	-44	59	-0.74	26	55	0.48	59	43	1.39			
Inverse Variance Matching	5	-151	59	-2.55	-65	51	-1.28	51	43	1.21	67	40	1.69			
Inverse Probability Weighting	1	-195	54	-3.64	25	53	0.46	46	41	1.14	56	40	1.40			
L-PSM Using Imbens-Rubin Covariate Selector	1	-161	58	-2.78	66	57	1.15	20	53	0.37	56	44	1.28			
L-PSM Using Imbens-Rubin Covariate Selector	5	-195	57	-3.44	32	50	0.63	59	40	1.48	63	41	1.54			
Penalized Logit-Based Propensity Score Matching	1	-223	54	-4.14	82	49	1.67	94	81	1.15	53	43	1.24			
Penalized Logit-Based Propensity Score Matching	5	-212	52	-4.05	102	43	2.39	77	65	1.18	66	40	1.64			
GenMatch	1	-255	77	-3.30	39	62	0.63	88	54	1.65	57	42	1.35			
GenMatch	5	-163	55	-2.97	10	45	0.23	65	48	1.37	68	39	1.73			
Classification Tree, Pruned	1	-192	49	-3.90	56	41	1.37	76	46	1.64	45	39	1.15			
Classification Tree, Pruned	5	-192	49	-3.90	56	41	1.37	76	46	1.64	45	39	1.15			
Classification Tree, Bagged	1	-116	56	-2.10	52	53	0.98	60	55	1.10	45	39	1.15			
Classification Tree, Bagged	5	-151	48	-3.17	41	44	0.93	49	47	1.05	-20	42	-0.48			
Classification Tree, Random Forest	1	-165	67	-2.47	11	52	0.21	74	53	1.40	-12	49	-0.24			
Classification Tree, Random Forest	5	-144	54	-2.66	45	46	0.97	37	42	0.90	-6	43	-0.13			
Classification Tree, Boosted	1	-169	62	-2.72	139	58	2.41	101	46	2.18	67	40	1.67			
Classification Tree, Boosted	5	-146	57	-2.58	136	44	3.10	133	39	3.43	70	40	1.75			
Support Vector Machine	1	-143	65	-2.20	108	66	1.63	261	141	1.85	43	41	1.05			
Support Vector Machine	5	-152	58	-2.64	65	54	1.20	169	91	1.87	54	40	1.36			

Notes: This table shows the estimated monthly earnings difference over 18 months after random assignment between experimental controls and matched/reweighted eligible non-participants (ENPs), using estimators considered in our validation exercise, for the Adult Men with Earnings History (Panel A) and Adult Women with Earnings History (Panel B) samples. The following covariate sets were used for each estimator: demographic characteristics only (col. 1), demographic characteristics and 12 quarters of earnings (col. 2), 12 quarters of earnings only (col. 3), and one quarter of earnings only (col.4). “Bias” is the estimated difference and “SE” denotes standard error.

Table A.6: Fixed Effects Models Validation Exercise Results

	Bias	S.E.	t-stat
<u>A. Adult Men (N=770)</u>			
Individual FE	75	48	1.57
+Individual Trends	289	46	6.22
+Heterogenous Layoff	343	54	6.32
<u>B. Adult Women (N=1375)</u>			
Individual FE	82	25	3.23
+Individual Trends	145	26	5.62
+Heterogenous Layoff	154	29	5.26
<u>C. Adult Men with Earnings History (N=635)</u>			
Individual FE	10	59	0.17
+Individual Trends	294	58	5.09
+Heterogenous Layoff	342	64	5.35
<u>D. Adult Women with Earnings History (N=791)</u>			
Individual FE	64	38	1.69
+Individual Trends	176	39	4.47
+Heterogenous Layoff	190	43	4.43

Notes: This table shows the estimated monthly earnings effect of belonging to the experimental control group relative to the eligible non-participants (ENP) group, over 18 months after random assignment, using the NJS validation data. Each panel shows the results for different samples. Within each panel, the specification in the first row includes individual fixed effects, month fixed effects, and indicators for month relative to layoff; the second row adds individual time trends; and the third row adds heterogeneous layoff effects. “Bias” is the estimated difference in earnings and “SE” denotes the standard error, clustered by individual.

Table A.7: Effect of UI Potential Duration on Enrollment

	Dependent Variable: Enrolled			
	(1)	(2)	(3)	(4)
Benefit Duration (10 Weeks)	0.0015 (0.0002)	0.0015 (0.0002)	0.0010 (0.0002)	0.0022 (0.0003)
Benefit Duration at Layoff (10 Weeks)		-0.0003 (0.0004)		
UI Claim Fixed Effects			X	
Includes Only First Unemployment Spell				X
Dependent Variable Mean	0.0153	0.0153	0.0153	0.0207
Observations (UI Claim - Quarter)	793,448	793,448	793,448	365,664
Observations (UI Claims)	99,181	99,181	99,181	79,589

Notes: This table shows the estimated effect of UI benefit durations on enrollment over the eight quarters after filing a UI claim. Additional controls include: (all columns) indicators for quarters post layoff, year indicators, quarter-in-year indicators, quarterly state unemployment rate (quadratic); (all columns except (3)) female, 10-year age category, Black, Hispanic, indicator for having dependents, prior year wage quintile, tenure category, 2-digit prior industry, 2-digit prior occupation. Standard errors are clustered by individuals and in parentheses. There are 96,583 unique individuals in the regressions.