OPPORTUNITY INSIGHTS

Digital Twins:

A New Tool to Evaluate Workforce Training and Education Programs

THE OPPORTUNITY INSIGHTS TEAM

Measure the effects of workforce training and education programs on earnings with the same accuracy as randomized trials, without the cost.

Randomized controlled trials (RCTs) are currently the gold standard for evidence on the impacts of programs and policies. For example, the What Works Clearinghouse—the U.S. Department of Education's repository of evidence used to guide funding decisions on educational interventions—reserves its "strong" evidence rating for randomized controlled trials. Although well-designed randomized trials provide high-quality evidence of program impact, they are expensive and challenging to implement. As a result, very few programs have been evaluated experimentally. For example, of the more than 5,000 workforce training programs in the United States, fewer than 50 (less than 1%) have been evaluated via RCTs.

In a new paper (Chetty, Fogel, Katz, Noray, Porter, Reisinger 2026), we propose a new "Digital Twins" method to estimate the impacts of workforce training and other career advancement programs by comparing the earnings outcomes of program participants to those of similar non-participants. We identify comparable non-participants ("Digital Twins") using modern machine learning methods to predict what each participant would have earned had they not participated in the program.

We test our Digital Twins estimates against 37 RCTs of workforce training programs that were conducted between 2000 and 2019. We find that our approach produces estimates very similar to those from RCTs, with comparable rates of error relative to what was identified in the randomized trials. As a result, Digital Twins can provide information on the impacts of many programs that is essentially as good as what one would get from a randomized trial—without the cost of having to run one.

EXECUTIVE SUMMARY

- Digital Twins is a new method of estimating the impacts of workforce training and other career advancement programs that provides evidence as accurate as that from randomized controlled trials (RCTs).
- The method applies machine learning to anonymized Census and tax records covering all American adults to identify "digital twins" for program participants: people with similar backgrounds and work histories who did not enroll.
- Previous methods that compared participants with similar matched non-participants generally failed to reproduce RCT results. Digital Twins closely matches RCTs because it draws on far larger samples, much richer data, and modern machine learning.
- Because Digital Twins is cheaper and faster than running an RCT, it can be used to evaluate programs at scale, creating a continuous feedback loop for improving them.

Applying our method to a broader set of sites and enrollment years beyond those evaluated experimentally, we find that program effects on earnings vary substantially across locations and over time. The variability of treatment effects across sites and over time increases the gains from using site- and enrollment-year-specific Digital Twins estimates relative to estimates from RCTs, which are typically conducted at a single point in time in a subset of sites. In fact, Digital Twins estimates can provide *more* accurate evidence for a particular site and year than typical randomized experiments.

Importantly, the Digital Twins method evaluates programs by the earnings *gain* that they generate. Judging programs simply by the earnings *levels* of their participants may penalize programs that enroll lower-income participants and reward programs that enroll people who would have done well anyway. Estimating earnings gains from programs (value-added) is essential for determining which programs should be scaled and which programs need to be improved.

We are now expanding the Digital Twins tool to evaluate a large set of workforce training and career advancement programs across the United States. Our vision is to equip policymakers, practitioners, students, philanthropists, and investors with precise data on which programs deliver the greatest returns so they can direct scarce resources in ways that allow more people to achieve the American Dream of upward mobility.

What Sets Digital Twins Apart?

Unlike RCTs, which typically provide just one static estimate of treatment effects that is then assumed to persist in future years and in new (untested) locations, the Digital Twins approach allows for low-cost, dynamic, and near real-time measurement of cohort-specific program impacts.

Our machine learning method produces more accurate estimates of program impact than previous approaches using techniques such as regression, propensity score matching, and other methods.[1] These gains in accuracy relative to prior work come from using much larger samples and richer information—especially on participants' earnings and employment histories—coupled with a flexible model that can capture more complex relationships between those characteristics and future earnings.

HOW WE BUILD DIGITAL TWINS

For each participant in a given program, we construct a digital twin using detailed information from linked participation and Census and tax records covering all American adults (270 million people today). We identify non-participants who have similar characteristics as participants before program enrollment across multiple dimensions: prior earnings and employment, industry of employment, age, sex, family structure, and neighborhood characteristics.

In principle, we might look for someone who matches the participant exactly. In practice, as we add more characteristics and measure them in greater detail, exact matches become increasingly difficult to find. For example, someone earning $21,000 at age 24 may be very similar to someone earning $22,000 at age 25; requiring an exact match would discard this useful comparison. We therefore look for "near twins." We use a machine learning method called gradient-boosted trees to learn which similarities matter for predicting earnings and combine information from these near twins into a single synthetic digital twin for each participant. That digital twin provides the prediction of what the participant would have earned without the program.

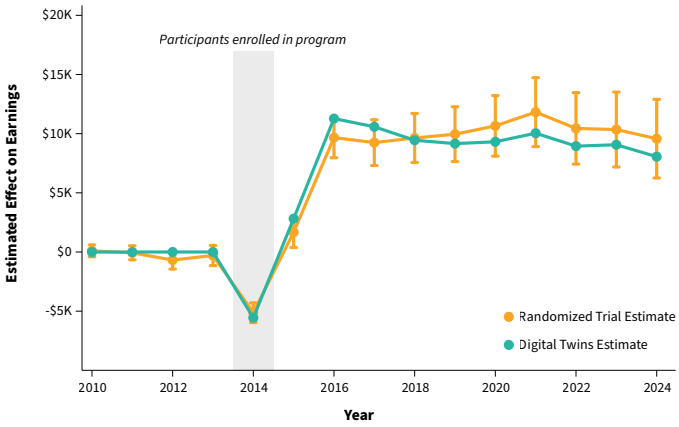
EXPERIMENTAL VALIDATION

Ex-ante, there is no guarantee that the Digital Twins approach will produce evidence that accurately identifies programs' causal effects. For example, program participants might be more motivated than their digital twins to find jobs, which could lead them to have higher earnings even absent any effect of the program. Because we do not directly observe motivation in our data, such factors could lead to biased conclusions about program impacts. It is therefore critical to test whether the Digital Twins approach works by comparing its results to evidence from randomized trials. We validate our Digital Twins estimates against 37 RCTs of workforce training programs conducted between 2000 and 2019 across the United States.

Figure 1 illustrates this comparison for one of the programs we evaluate: Year Up United (YUU). YUU was evaluated in a randomized trial in 2014. The orange line in the figure plots the estimated impact of YUU on participants' earnings from the randomized trial. Before 2014, there was no difference in earnings between program participants and the randomized control group, as expected since the intervention had not yet occurred. In the year of enrollment, earnings fell among program participants relative to their peers because they enrolled in YUU training and thus were working less. Following program enrollment, earnings rose by $9,700 per year (in 2025 dollars, adjusting for inflation)—a gain that was sustained for the subsequent decade.

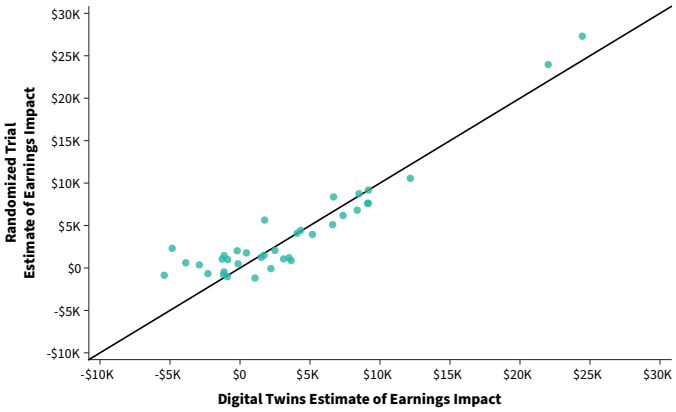
The green line in Figure 1 shows that we obtain virtually identical estimates of earnings impacts using the Digital Twins approach. This result provides a first demonstration that the Digital Twins method accurately recovers what one would have obtained from running a randomized trial.

Figure 1: Digital Twins Estimates Closely Track Estimates from Randomized Trial of Year Up United



This figure plots the estimated effect of enrollment in Year Up United's Core program based on evidence from the 2014 PACE randomized trial (orange series) and the Digital Twins methodology (green series). All earnings values are inflated to 2025 dollars. Enrollment occurred between January 2013 and August 2014. The figure aligns cohorts by years relative to enrollment, with the enrollment year displayed as 2014. Download figure.

Figure 2: Digital Twins Estimates Are Similar to RCT Estimates Across 37 Trials



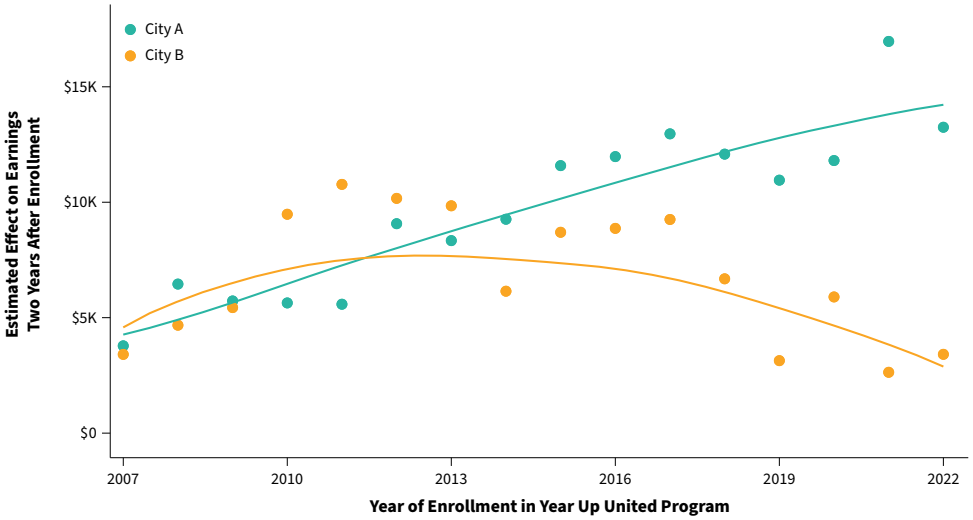
Each point compares the randomized trial and Digital Twins estimates of a program's effect on average annual wage earnings two to five years after enrollment. Points on the diagonal are cases where the two approaches agree exactly. Download figure.

We similarly compare the Digital Twins and RCT estimates for each of the other programs evaluated in randomized trials. We find that the Digital Twins estimates closely match the RCT estimates across these programs, with a correlation of 0.93 (Figure 2). While the Digital Twins estimates are not identical to the estimates from the trials, they exhibit very little error. In particular, the Digital Twins estimator can clearly distinguish programs that have no earnings impact from those that raise annual earnings by roughly $5,000, and from exceptional programs whose impacts exceed $10,000.

APPLYING DIGITAL TWINS TO SETTINGS WITHOUT EXPERIMENTAL EVIDENCE

The power of the experimentally validated Digital Twins method is that it can be applied to any program on an ongoing basis, yielding a continuous feedback loop on program performance. Year Up United illustrates this approach. As shown above, the Digital Twins estimate closely matches the result of Year Up United's PACE RCT for those participants. But estimated earnings gains vary substantially across sites and enrollment years (Figure 3). In some cities, Year Up United's impacts have grown substantially over time; in others, they have not. This type of granular information provides more actionable information and allows programs to recalibrate their approaches—cohort-specific information allows programs and funders to identify more precisely which models should be scaled and which need to be adjusted to maximize impacts.

Figure 3: Year Up United's Impact on Earnings Differs Across Cities and Enrollment Years



Each point is a Digital Twins estimate of Year Up United's effect on annual wage earnings two years after enrollment, by city and year in which participants enrolled in the program. We use earnings two years after enrollment so that recent enrollment years can be included. Effects two years after enrollment are strongly correlated with longer-run effects across programs. Download figure.

The variation evident in Figure 3 points to another important benefit of the Digital Twins approach relative to evidence from typical RCTs. Because RCTs are costly to implement, they are typically conducted infrequently, and historical estimates from some sites are used as measures of likely impact in other time periods and places. But when impacts differ so much across contexts even within a single program, those historical predictions can be inaccurate for a new site. A Digital Twins estimate for the relevant site and enrollment year can thus offer a much more accurate assessment of impact than a historical RCT estimate.

IMPLICATIONS FOR PROGRAM EVALUATION

Having developed and experimentally validated the Digital Twins estimator, we can now apply it to a broad range of programs to measure their value-added. We are currently working with training providers, state agencies, and higher education institutions across the U.S.—e.g., Western Governors University, the Massachusetts Department of Higher Education, Dallas College, and Texas State Technical College—to construct such estimates and use them to draw lessons about what makes career advancement programs most effective.

While our approach provides a new source of information on program impacts, we emphasize that RCTs remain essential for validating and recalibrating Digital Twins estimates. Our validation covers workforce training programs serving working-age adults with prior work histories and open or lightly screened enrollment. Applying our approach to other populations and policy areas will require experimental validation in those settings.

INTERESTED IN APPLYING THE DIGITAL TWINS APPROACH TO EVALUATE YOUR PROGRAM?

We welcome inquiries from interested training programs, higher education institutions, and others to discuss applications of our Digital Twins method to your programs. In general, this method is most appropriate for programs that have at least 200 participants per year over multiple (3+) years or an equivalent total number over fewer years. Given fixed costs of data processing and analysis, it is often most efficient to obtain data from institutions that cover many programs (e.g., community college systems, state agencies, etc.). Programs need only share information on enrollees (e.g., Social Security numbers, names and dates of birth, or dates of birth and ZIP code) that will permit linkage of their records to Census and tax data. Opportunity Insights uses secure linkage protocols that protect individual identities, with all identifying information stripped from datasets before analysis and release of aggregate statistics only following a disclosure review process by the Census Bureau to ensure that no individual information is inadvertently revealed.

If you are interested in learning more or exploring potential data partnerships with OI, please contact data@opportunityinsights.org

1. See, for example, LaLonde, Robert J. 1986. "Evaluating the Econometric Evaluations of Training Programs with Experimental Data." American Economic Review 76 (4): 604–620; and Imbens, Guido W., and Yiqing Xu. 2025. "Comparing Experimental and Nonexperimental Methods: What Lessons Have We Learned Four Decades after LaLonde (1986)?" Journal of Economic Perspectives 39(4): 173–202.

Want to learn more?

Read the Paper ↗ See Presentation Slides ↗ Download Figures ↗

All materials are freely available for use with citation

Based at Harvard University, **Opportunity Insights** identifies barriers to economic opportunity and develops scalable solutions that will empower families throughout the United States to rise out of poverty. opportunityinsights.org