



Tip Sheet # 2: Adjust for Varying Randomization Fractions During Analysis

Zachary Epstein and Jacob Alex Klerman, September 2023

In an experimental evaluation of an RESEA program, your evaluator will randomly assign UI claimants to the evaluation's study groups—a control group and one or more treatment groups. Many RESEA programs operate on a weekly cycle, identifying and scheduling weekly cohorts of eligible claimants at each local American Job Center (AJC). Therefore, evaluations of RESEA programs could enroll and randomly assign weekly cohorts, assigning some fraction of that week's eligible cohort to the treatment and control groups.

A key choice in such experimental designs is the **randomization fraction**: the fraction of RESEA-eligible claimants randomly assigned to the treatment group, rather than to the control. In some experimental designs, the evaluator will keep that fraction consistent over time and across local AJCs, perhaps assigning half of the cohort to the treatment group and the other half to the control group. However, the evaluator may vary the randomization fraction over time or across offices, perhaps due to large fluctuations in weekly numbers of UI claimants or differences in program capacity across offices. Doing so is not inherently a problem, but the details of the analysis need to account for that variation.

This tip sheet is part of the *RESEA Evaluation Technical Assistance (EvalTA) Analysis Tip Sheet Series* designed to help state staff better understand: 1) the main activities a RESEA evaluator will carry out, and 2) help staff identify and avoid common analysis pitfalls when working with an evaluator. This tip sheet discusses varying randomization fractions, why analysis needs to account for them, and analytic strategies to do so. Though technical in some sections, the discussion focuses on the intuition for varying randomization fractions.¹ We provide references to resources with more technical discussion for evaluators to review.

TIP #2: ANALYSES NEED TO ACCOUNT FOR VARYING RANDOMIZATION FRACTIONS

To understand *Tip #2*, you must first understand the definition of the randomization fraction and why that fraction might vary.

Box 1: Terminology

Experimental Evaluation: Also called a random assignment study or randomized controlled trial (RCT), an evaluation whereby eligible individuals are randomly assigned to one of—at minimum—two study groups: treatment or control. Comparing outcomes between the study groups estimates the impact of the program or intervention under evaluation.

Treatment Group: Those randomly assigned RESEA-eligible claimants who are enrolled in the intervention under evaluation (e.g., the RESEA program or component thereof).

Control Group: Those randomly assigned RESEA-eligible claimants who are *not* enrolled in the intervention under evaluation (e.g., the RESEA program or component thereof).

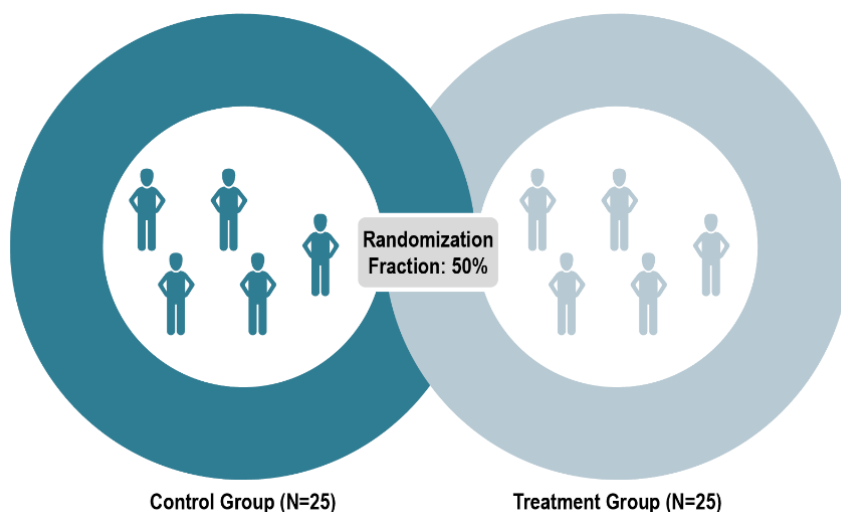
Randomization fraction: In an experimental evaluation, the fraction of RESEA-eligible claimants randomly assigned to each study group.

¹ These issues were also discussed briefly as Question #9 in Common Pitfalls in Designing RESEA Causal Evaluations; Clarkwest and Wheaton, 2023

We will use a two-arm whole program to illustrate, though the same logic applies to component evaluations, including those that have multiple treatment arms. For a state that wants to evaluate its RESEA program as a whole, evaluators randomly assign some UI claimants to the RESEA program model (i.e., treatment group) and other UI claimants to a control group that is not offered RESEA services.²

The Whole Program Evaluation's design may call for varying the fraction of claimants randomly assigned to the treatment group over time or across offices. There can be good reasons to do so. For example, if the number of new claims vary by week, but the number of RESEA slots remains constant, then evaluators vary the randomization fraction each week to align the number of UI claimants selected for RESEA with the number of available appointments at the local AJC at which they would receive services³.

Exhibit 1. Randomization Fraction in a Normal Week



Some states may allow for weekly variation in the randomization fraction to exactly match office-level capacity in each week. Other states may keep the randomization fraction the same in most weeks while allowing for occasional variation in response to large swings in the number of RESEA-eligible claimants in a given week or time of year. For example, during seasonal spikes in claim activity after the new

year holiday, some AJCs may be overwhelmed with customers if the state does not reduce the fraction selected for RESEA. Under either approach, many states will likely need some adjustment to the randomization fraction over the course of an evaluation's enrollment period.

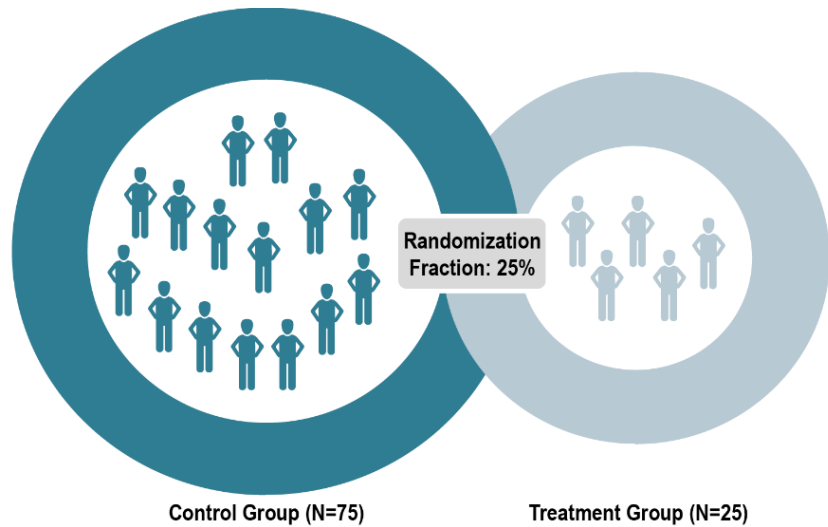
To illustrate, consider an AJC with 25 available slots for new RESEA customers each week. In a **normal week (NW)**, the state expects a cohort of 50 RESEA-eligible new UI claimants. In this case, the evaluation could set the randomization fraction to 50 percent, assigning 25 claimants to the RESEA treatment group and the other 25 to the control group (Exhibit 1).

² For more on experimental designs and random assignment, see the [RESEA Evaluation Toolkit](#) (Mills de la Rosa, et al., 2021).

³ Similarly, in any given week, a state may vary the randomization fraction across offices if local offices vary in the proportion of the eligible UI caseload that they are able to serve. So randomization can vary over time, across offices, or both.

Now consider a week in which the number of RESEA-eligible claims increases substantially, perhaps due to a mass layoff event. In the **mass layoff week (MLW)**, the state expects 100 RESEA-eligible new UI claimants. The same AJC can still serve 25 new claimants. So, the state sets the randomization fraction to 25 percent. As in a normal week, 25 claimants will be selected for RESEA (i.e., up to the AJC's capacity), and the other 75 claimants will be assigned to the control group.

Exhibit 2. Randomization Fraction in a Mass Layoff Week

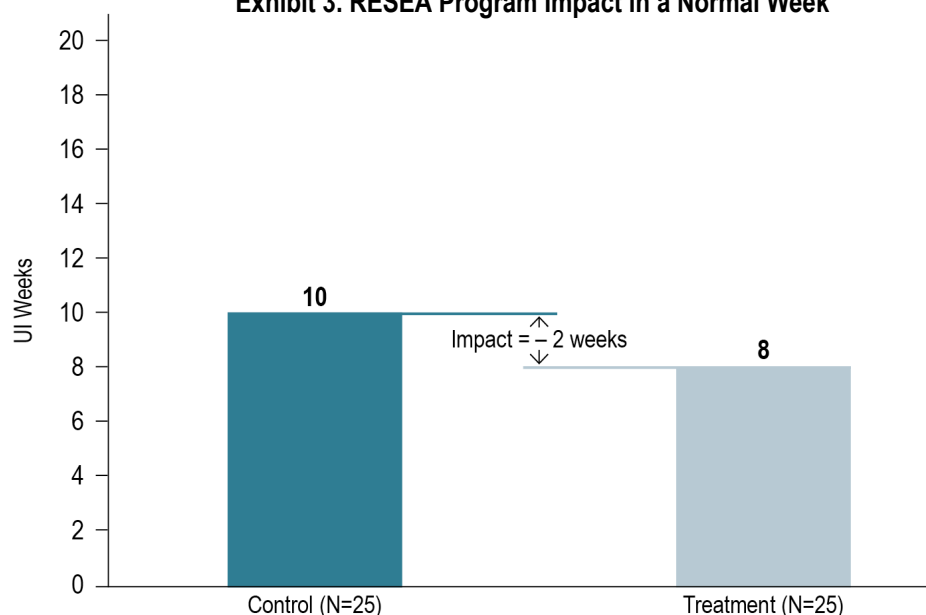


THE PITFALL: HAVING UNBALANCED TREATMENT AND COMPARISON GROUPS FAILURE TO ACCOUNT FOR VARYING RANDOMIZATION FRACTIONS

This section develops a simple example to explain why analysis must take account of varying randomization fractions and how failing to do so will yield an incorrect estimate of your program's impact.

The exhibit below illustrates an RESEA program's impact. The vertical bars show the average UI duration for the treatment and control groups, and the difference in the height of those bars represents the program's impact (i.e., the differences in average outcomes between the two study groups). The width of each bar represents the number of claimants included in that calculation of average duration. For example, in Exhibit 3, which illustrates impacts in the normal week example described above, the two vertical bars are of equal width, since both groups included 25 RESEA-eligible claimants.

Exhibit 3. RESEA Program Impact in a Normal Week



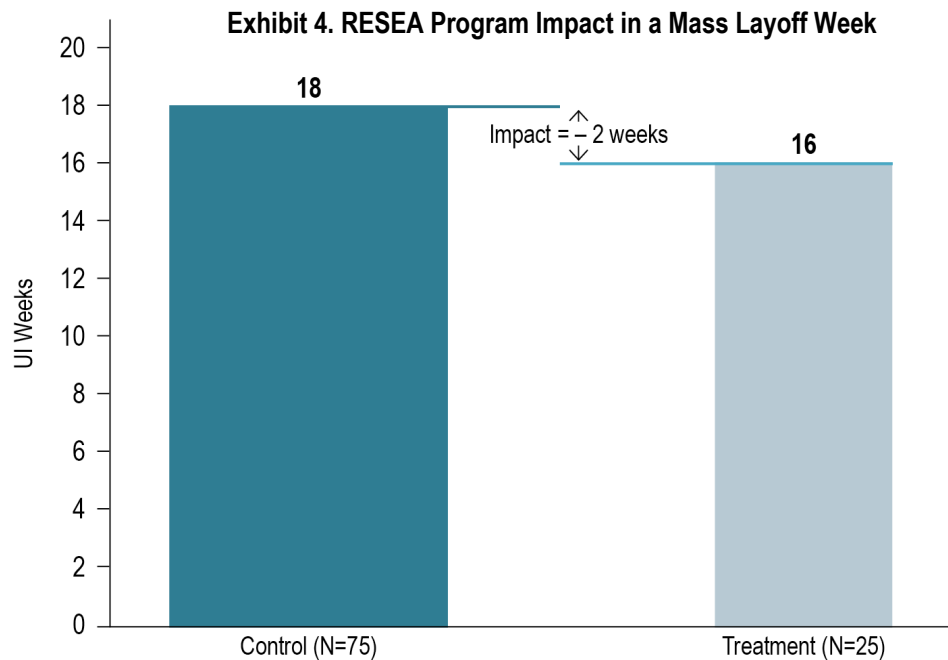
Suppose that in the normal week, those not selected for RESEA collect UI for an average of ten weeks. Suppose also that those selected for RESEA collect UI for an average of eight weeks. The impact of being selected for RESEA is the difference in UI weeks between those offered RESEA (10 weeks) and those not selected for RESEA (8 weeks); that is, a decrease of 2 weeks.

Now consider the impact estimate in the mass layoff week (Exhibit 5). When a mass layoff happens, we assume there are fewer jobs available per job seeker. As a result, duration on UI jumps sharply for all claimants: to 20 weeks for those assigned to the no-service control group and 18 weeks for those assigned to the RESEA treatment group. The impact of being selected for RESEA is calculated in the same way, taking the difference in UI duration between the two groups. In this example, the impact in the mass layoff week is the same as that of a normal week: a decrease of 2 weeks ($18 - 20 = -2$).⁴

In both weeks—the normal and mass layoff weeks—selection for the RESEA program reduces UI duration by two weeks. That is the true impact of RESEA. We can see it visually, and when we compute the average impact mathematically, we know the result should be -2.

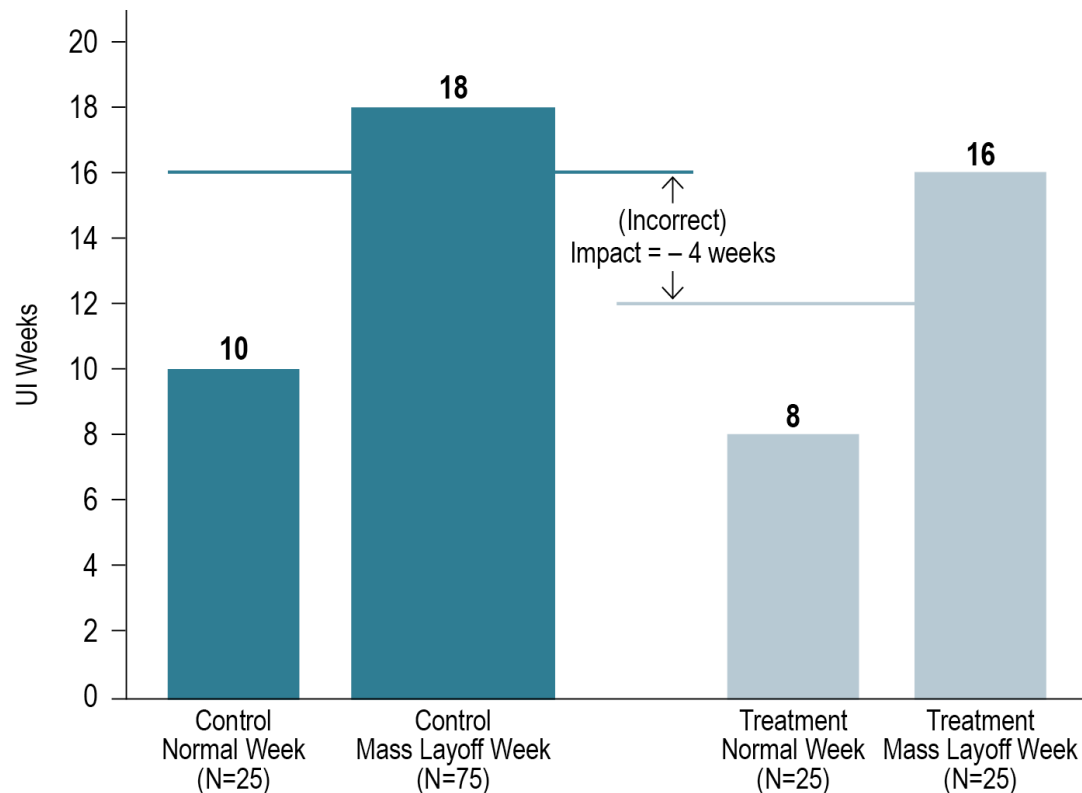
However, when analyzing the two weeks together, the computation must take into account the fact

that the randomization fractions differed between the two weeks (50% in the normal week and 25% in the mass layoff week). Let's consider what happens if the analysis ignores those varying fractions. In Exhibit 5, the vertical bars show the UI duration outcomes for claimants assigned to the Treatment and Control groups in the normal and mass layoff weeks, as shown above in Exhibits 3 and 4. The horizontal lines show the average UI duration for the Treatment and Control groups, calculated without accounting for varying randomization fractions.



⁴ For simplicity, we assume the impact is the same across the NW and MLW. The principles in this Tip Sheet would hold even if the impacts differed between weeks.

Exhibit 5. Impact Estimate Using Simple Average, Not Accounting for Varying Fractions



Among those selected for RESEA (Treatment), the simple average UI duration is 12 weeks.⁵ And for those not selected for RESEA (Control), the simple average UI duration is 16 weeks.⁶ If we calculate the impact simply by taking the difference in average outcomes between the two groups, we get a difference of -4 weeks. But we know that is incorrect, since we established earlier that the true impact of RESEA is a two-week decrease in UI duration.

So what leads to that erroneous estimate? Intuitively (and visually), we can see that the problem lies in the composition of the Treatment and Control groups. Due to the differences in randomization fractions between weeks, the Control group has a larger share of claimants from the Mass Layoff week than does the Treatment Group. Relative to the treatment group, the control group is weighted more heavily toward claimants from the mass layoff week. So even though we randomly assigned claimants within each individual week—thus producing sets of Treatment and Comparison group members who were similar to each other *within* each week—when we combine across weeks, the Treatment and Comparison groups are

⁵ Computation of the 12-weeks average UI duration for Treatment:

$$\begin{aligned}
 12 \text{ weeks} &= \frac{(8 \text{ weeks} * 25 \text{ claimants in NW}) + (16 \text{ weeks} * 25 \text{ claimants in MLW})}{25 \text{ claimants in NW} + 25 \text{ claimants in MLW}} \\
 &= \frac{200 \text{ weeks} + 400 \text{ weeks}}{50 \text{ claimants}}
 \end{aligned}$$

⁶ Computation of the 17-weeks average UI duration for Comparison:

$$\begin{aligned}
 19 \text{ weeks} &= \frac{(10 \text{ weeks} * 25 \text{ claimants in NW}) + (18 \text{ weeks} * 75 \text{ claimants in MLW})}{25 \text{ claimants in NW} + 75 \text{ claimants in MLW}} \\
 &= \frac{250 \text{ weeks} + 1,350 \text{ weeks}}{100 \text{ claimants}}
 \end{aligned}$$

not similar. Analyses can fully correct for those differences be corrected for, but simple Treatment-Control comparisons will give biased impact estimates.

THE FIX: DESIGN ANALYSES TO TAKE INTO ACCOUNT VARYING RANDOMIZATION FRACTIONS

Having established that variation in the randomization fraction introduces differences between the Treatment and Control group, this section discusses the approach to accounting for that variation in the analysis to adjust for those differences.

The intuition for how to adjust the analysis for varying randomization fractions is as follows. **Identify cohorts of claimants in your data with common randomization fractions (e.g., normal weeks and mass layoff weeks). Then, estimate impact separately for groupings of cohorts with common randomization fractions and calculate the average impact across those groupings.**

In practice, there are two ways to implement this adjustment.

- **Fixed Effects:** Include a set of binary indicators (“dummy variables”) for every weekly cohort—or office/week combination, if fractions vary by office, too—with the same randomization fraction.⁷ In our example, the dummy variables account for the changes in claimants’ outcomes attributed to the changing randomization fractions (i.e., the mass layoff event), independent of the RESEA program.
- **Weighting:** Before doing the analysis, reweight the individual-level data so that, for each randomization fraction, the treatment and Control groups have the same weight. This entails scaling up each individual’s outcome by the inverse of their assignment probability. In theory, the ideal weights will vary with the exact research question, but a complete discussion of these issues is beyond the scope of this Tip Sheet.

For your evaluator to make those adjustments, they will need a record of how randomization fractions varied over time. For weighting, specifically, they will need to know the randomization fractions used at the time and place that each claimant was assigned. With that information, an evaluator with strong experience in conducting impact studies will be able to implement the appropriate adjustments to maintain the valid treatment/control comparison and satisfy CLEAR’s standards for a “high” study rating.⁸

Regardless of your evaluator’s preferred approach, consider asking, “*Did the random assignment fraction vary over time or across offices?*”

If the evaluator says yes, you could confirm that they accounted for that variation in their analysis. If they do not, you may want to request that your evaluator changes the analysis plan to include dummy variables, weights, or both.

⁷ That strategy sometimes involves a very large number of dummy variables and will therefore sometimes raise some computational issues. Saunders et al. (2019) discusses these issues in more detail.

⁸ For reference, see CLEAR’s RCT Criterion #3 in the [Causal Evidence Guidelines](#): *Did the probability of assignment into treatment or control groups vary over time without appropriate adjustment?*

ADDITIONAL PUBLICATIONS

Clarkwest, Andrew and Victoria Wheat. 2023. Common Pitfalls in Designing RESEA Causal Evaluations.

Mills De La Rosa, S., Souvanna, P., Clarkwest, A., Kappil, T., Epstein, Z., Rothschild, L., Kuehn, D., Wall, A., Klerman, J., and Nightingale, D. (2021) Reemployment Services and Eligibility Assessment (RESEA) Evaluation Toolkit: Key Elements for State RESEA Programs. US Department of Labor. https://www.dol.gov/sites/dolgov/files/OASP/evaluation/pdf/RESEA_Toolkit_February2021.pdf.

Saunders, C., Dastrup, E., Epstein, Z., Walton, D., Adam, T., Klerman, J.A. and Barnow, B.S. (2019). Evaluation of Impacts of the Reemployment and Eligibility Assessment (REA) Program: Final Report Appendices. U.S. Department of Labor. <https://www.dol.gov/sites/dolgov/files/OASP/evaluation/pdf/REA%20Impact%20Study%20-%20Final%20Report%20Appendices.pdf>

ABOUT THIS SERIES

State agencies' evaluations of their RESEA interventions—whether the whole program or components thereof—are integral to expanding the evidence base of effective service delivery strategies. This tip sheet is part of a series developed to serve as a quick-reference tool for states planning, launching, and carrying out evaluations of their RESEA programs. Additional tip sheets in this series are available on the [RESEA Evaluation and Evidence Resources WorkforceGPS page](#).

Suggested citation: Epstein, Z. & Klerman, J.A. (2023). *Tip Sheet #2: Adjust for Varying Randomization Fractions During Analysis*. RESEA Evaluation Technical Assistance Analysis Tip Sheet Series. Prepared for the U.S. Department of Labor, Chief Evaluation Office. Abt Associates.

Additional Resources and Information

DOL's RESEA Helpline: RESEA@abtassoc.com

Additional RESEA evaluation technical assistance resources are available here: https://rc.workforcegps.org/resources/2019/07/30/17/32/RESEA_Evaluation_Evidence_Resources

Study reports, publications, and other materials are available here: <https://www.dol.gov/agencies/oasp/evaluation/completedstudies/RESEA-Evidence-Building-Implementation-Study>

Chief Evaluation Office
U.S. DEPARTMENT OF LABOR



This document was prepared for the U.S. Department of Labor (DOL), Chief Evaluation Office by Abt Associates under Contract # 1605DC-18-A-0037. The views expressed are those of the authors and should not be attributed to DOL, nor does mention of trade names, commercial products, or organizations imply endorsement of same by the U.S. Government.