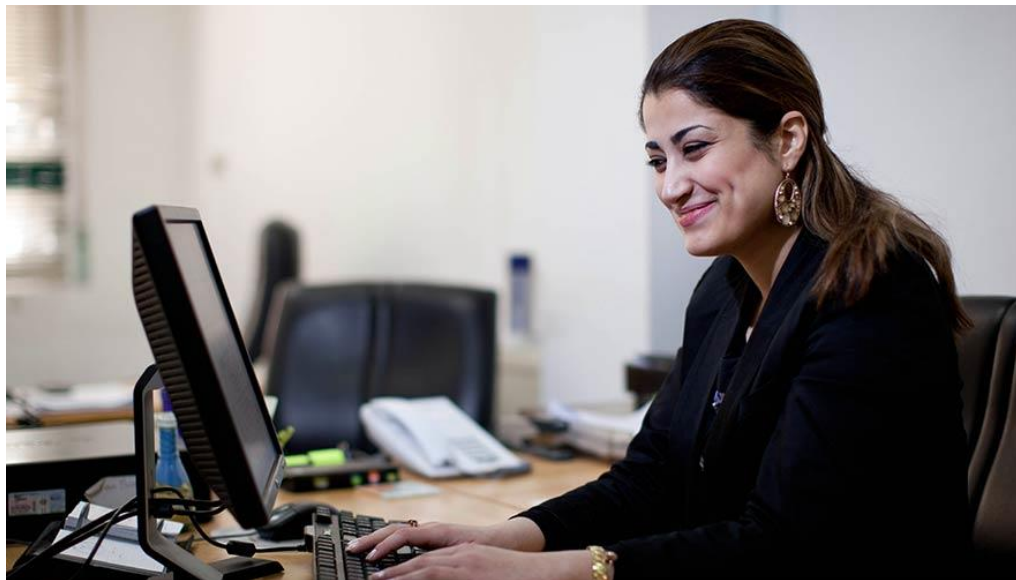


## CASE STUDY 4: SKILLS TRAINING AND WAGE SUBSIDY PROGRAMS IN JORDAN

Addressing threats to experimental integrity

---



This case study is based on “Soft Skills or Hard Cash? The Impact of Training and Wage Subsidy Programs on Female Youth Employment in Jordan” by Matthew Groh, Nandini Krishnan, David McKenzie, and Tara Vishwanath.

J-PAL thanks the authors for allowing us to use their paper

## KEY VOCABULARY

---

**Counterfactual:** what would have happened to the participants of a program had they not participated. The counterfactual cannot be observed from the treatment group; can only be inferred from the comparison group.

**Equivalence:** groups are statistically identical, indicated by similarity on baseline characteristics, both observable and unobservable. Ensured by randomization.

**Attrition:** individuals selected for a study drop out of the treatment or comparison group over the course of the study, before the final outcomes are measured.

**Attrition Bias:** statistical bias, which occurs when both: (a) individuals drop out of either the treatment or the comparison group, and (b) who drops out is correlated with which group they are assigned.

**Partial Compliance:** individuals do not comply with their assignment (to treatment or comparison). Also termed "diffusion" or "contamination."

**Intention to Treat:** measured impact of a program that compares outcomes from all individuals assigned to the treatment group to those assigned to the control group (regardless of whether they actually availed of the treatment). Often termed, "Average Treatment Effect" (ATE).

**Local Average Treatment Effect (LATE):** the estimated impact of a program on participants who participated in the program solely because they were assigned to the treatment group (also called the Complier Average Causal Effect). This is different from the Intention to Treat estimate when there is partial compliance. A special case of the Local Average Treatment Effect, called the Treatment on the Treated estimate, occurs when partial compliance only occurs in the Treatment Group.

**Externality:** an indirect cost or benefit incurred by individuals who did not directly receive the treatment. Also termed "spillover."

## INTRODUCTION

---

The Jordan New Opportunities for Women (Jordan NOW) tested two interventions to overcome some of the barriers to firms hiring young female graduates: wage subsidy vouchers and employability skills training. A randomized evaluation was designed to test the relative effectiveness of each intervention separately and both combined together.

Randomization ensures that the treatment and comparison groups are comparable at the beginning of a study. However it cannot ensure that people within each group comply with the treatment to which they were assigned; nor can it ensure that the groups remain comparable until the end of the program when the post-intervention outcomes are measured. . Life also goes on after the randomization: other events besides the program happen between initial randomization and the endline. These events can potentially reintroduce selection bias, diminishing the validity of the impact estimates, and are threats to the integrity of the experiment.

How can common threats to experimental integrity be managed?

## SKILLS TRAINING AND WAGE SUBSIDIES

---

Throughout the Middle East, unemployment rates of educated youth have been persistently high and female labor force participation low. Only 23% of female community college graduates in Jordan are employed 16 months after graduating, despite 93% saying they want to work at the time of graduation. This enormous gap between expectations and reality highlights the challenge facing young women who want to work in the Middle East.

The problems faced by young women in the labor market are twofold. Firstly, firms are often reluctant to hire younger workers, regardless of gender, since they lack experience and are of untested quality. Secondly, employers have qualms about hiring women because they believe that they are less committed to their jobs and might leave if they get married or have children.

Jordan NOW designed two interventions to get at these labor market frictions: wage subsidy vouchers (to reduce the cost of employing women) and employability skills training (to increase the benefit of employing them).

**Skills Training** may enhance employment prospects by giving youth better skills and confidence in looking for jobs. In addition, skills training could make employees more productive in their first months on the job by reducing the amount of time that *firms* need to devote to training new hires on the basics of working in a business environment.

**Wage subsidies** can be used to generate employment for the disadvantaged. Employers see females as having a higher probability of leaving early, which lowers any estimated returns from training them and from the experience females accumulate over their tenure with the employer. If the expected benefit is lower, wage subsidies can keep the expected return of investing in female employees positive by offsetting some of the costs of employing them.

## EVALUATION DESIGN

---

Female students from community colleges who had passed their second year exams were chosen for the purpose of this

program. They were randomized into three different programs and a control group:

**Treatment 1:** Skills Training

**Treatment 2:** Wage Subsidies

**Treatment 3:** Skills Training *and* Wage Subsidies

**Control group:** No Program

Upon randomization, graduates receiving the wage subsidy program were given a non-transferrable job voucher that they could take to a firm while searching for jobs. The voucher paid the employer the minimum monthly wage for a maximum of six months if they hired the worker. Graduates receiving the skills training program were invited to a free intensive training course on soft skills. The training course lasted 45 hours over a 9 day period (5 hours per day). The take up of the programs was not 100 percent (i.e. not everyone assigned to a given treatment group actually took up the program) and posed significant challenges in evaluating the effectiveness of the two programs.

### DISCUSSION TOPIC 1

#### *Threats to experimental integrity*

Randomization ensures that the groups are equivalent, and therefore comparable, at the beginning of the program. The impact is then estimated as the difference in the average outcomes of the treatment group and the average outcomes of the comparison group, both at the end of the program. To be able to say that the program caused the impact, we need to be able to say that *the program* was the only difference between the treatment and comparison groups over the course of the evaluation.

- a. What does it mean to say that the groups are equivalent at the start of the program?
- b. Can you check if the groups are equivalent at the beginning of the program? How?
- c. Other than the program's direct and indirect impacts, what can happen over the course of the evaluation (after conducting the random assignment) to make the groups non-equivalent?

- d. How does non-equivalence at the end threaten the integrity of the experiment?

### MANAGING ATTRITION—WHEN THE GROUPS DO NOT REMAIN EQUIVALENT

Attrition is when people drop out of the sample—treatment and/or comparison groups—over the course of the experiment. This can happen because people drop out of the study, refuse to answer some questions, or cannot be found by enumerators during the endline survey. One common example in clinical trials is when people die before the final outcomes are measured; attrition of this sort is called “experimental mortality”.

#### Discussion Topic 2

##### Managing Attrition

Suppose there are 16,000 female graduates randomized into four groups as follows:

- Treatment 1:** 4000
- Treatment 2:** 4000
- Treatment 3:** 4000
- Control group:** 4000

After you randomize, the treatment and comparison groups are equivalent, meaning jobseekers from each of the three categories are equally represented in all groups. You are looking at the employment outcomes of female graduates who received the wage subsidy, either alone or combined with skills training. Employment outcomes are scaled as follows:

- Unemployment = score of 2
- Temporary employment = score of 1
- Permanent employment = score of 0

Suppose all jobseekers that are in the two voucher treatment groups use up their wage subsidy vouchers and, because these vouchers are non-transferrable, none of the jobseekers in the control group do so. The employment

outcomes for jobseekers in each group are shown for both the pretest and posttest.

**TABLE 1**

Outcome	Pretest			Posttest		
	Voucher	Voucher + Training	Control	Voucher	Voucher + Training	Control
2	4,000	4,000	4,000	500	500	2,000
1	-	-	-	2,500	1,000	1,000
0	-	-	-	1,000	2,500	1,000
<b>Total</b>	4,000	4,000	4,000	4,000	4,000	4,000

- Using the table above, calculate the following:
  - At pretest, what is the average employment outcome for each group?
  - At posttest, what is the average employment outcome for each group?
  - What is the impact of each program?

Suppose now that in the control group, half of the jobseekers who remain unemployed and half of those who are temporarily unemployed at the end of the year feel disillusioned and refuse to respond to the survey. The employment outcomes for jobseekers in each group are shown for both the pretest and posttest.

**TABLE 2**

Outcome	Pretest			Posttest		
	Voucher	Voucher + Training	Control	Voucher	Voucher + Training	Control
2	4,000	4,000	4,000	500	500	1,000
1	-	-	-	2,500	1,000	500
0	-	-	-	1,000	2,500	1,000
<b>Total</b>	4,000	4,000	4,000	4,000	4,000	2,500

2. Using the table above, calculate the following:
  - a. What is the impact of each program?
  - b. Is this outcome difference an accurate estimate of the impact of the program? Why or why not?
  - c. If it is not accurate, does it overestimate or underestimate the impact? Use your estimate of the impact from Discussion Topic 2 to calculate how much it over/underestimates the impact by.
  - d. How can we get a better estimate of the program's impact?
3. In Case 1, you learned about other methods to estimate program impact, such as pre-post, simple difference, differences in differences, and multivariate regression.
  - a. Does the threat of attrition only present itself in randomized evaluations?

### MANAGING PARTIAL COMPLIANCE—WHEN THE TREATMENT DOES NOT ACTUALLY GET TREATED OR THE COMPARISON GETS TREATED

Some people assigned to the treatment may in the end not actually get treated. Those randomly assigned to the skills training program may choose not to enroll. Or those assigned to the control group may decide they want skills training anyway and pay out of pocket to get this service from a private company. This is called “partial compliance”

or “diffusion” or, less benignly, “contamination.” In contrast to carefully controlled lab experiments, diffusion is ubiquitous in social programs. After all, life goes on, people will be people, and we have no control over what they decide to do over the course of the experiment. All we can do is plan our experiment and offer them programs. How, then, can we deal with the complications that arise from partial compliance?

### DISCUSSION TOPIC 3

#### *Selection Bias Due to Incomplete Take-up*

Suppose 1,000 of the 4,000 jobseekers who were offered the skills training program, either alone or combined with the wage subsidy program, were not interested in receiving skills training because they were intrinsically demotivated. Since, the 1,000 jobseekers that did not take-up the program were also not motivated to look for a job in the first place, they remained unemployed at the end of the year.

**TABLE 3**

Outcome	Pretest			Posttest		
	Training	Voucher + Training	Control	Training	Voucher + Training	Control
2	4,000	4,000	4,000	1,500	1,500	2,000
1	-	-	-	1,500	1,000	1,000
0	-	-	-	1,000	1,500	1,000
<b>Total</b>	4,000	4,000	4,000	4,000	4,000	4,000

1. Calculate the impact estimate based on the original group assignment.
  - a. This is one potential method of evaluating the impact of the program. In what ways is it useful and in what ways is it not useful?

You are interested in learning the effect of treatment on those actually treated (“treatment on the treated” (TOT) estimate).<sup>1</sup>

2. Five of your colleagues are passing by your desk; they all agree that you should calculate the effect of the treatment using only the 3,000 jobseekers that were actually treated.
  - a. Is this advice sound? Why or why not?
3. Another colleague says that it is not a good idea to drop the untreated entirely; you should use them but consider them as part of the comparison.
  - a. Is this advice sound? Why or why not?

## MANAGING SPILLOVERS—WHEN THE COMPARISON, ITSELF UNTREATED, BENEFITS FROM/ GETS HARMED BY THE TREATMENT BEING TREATED

---

People assigned to the control group may benefit or get harmed indirectly by those receiving treatment. Specifically, for such wage subsidy and job trainings programs we would be concerned that those in the treatment group will take opportunities away from individuals in the control group. Alternatively, we could imagine a situation in which spillovers are positive. Increased employment in the treatment group could improve the local economy, making it easier for control group jobseekers to find jobs. Or perhaps jobseekers in the control group had contacts in the treatment group and were now better connected to potential employers. In any of these cases, the control group would no longer represent the counterfactual—the state of the world had the program not been implemented.

---

<sup>1</sup> This is generally called the Local Average Treatment Effect (LATE) and is calculated by adjusting the Intention to Treat (ITT) estimate by the difference in proportions of people who took up the treatment in the treatment and control groups. In the specific instance where there is partial compliance in the treatment

group, but not in the control group (i.e. nobody takes up the program in the control group), the difference in proportions is the same as the compliance rate in the treatment group and the LATE is called the Treatment on the Treated (TOT) estimate.