

Soft Skills or Hard Cash?

The impact of training and wage subsidy programs on female youth employment in Jordan^{*}

Matthew Groh, *World Bank*
Nandini Krishnan, *World Bank*
David McKenzie, *World Bank, BREAD, CEPR, and IZA.*
Tara Vishwanath, *World Bank*

Abstract

Throughout the Middle East, unemployment rates of educated youth have been persistently high and female labor force participation, low. We study the impact of a randomized experiment in Jordan designed to assist female community college graduates find employment. One randomly chosen group of graduates was given a voucher that would pay an employer a subsidy equivalent to the minimum wage for up to 6 months if they hired the graduate; a second group was invited to attend 45 hours of employability skills training designed to provide them with the soft skills employers say graduates often lack; a third group was offered both interventions; and the fourth group forms the control group. We find that the job voucher led to a 40 percentage point increase in employment in the short-run, but that most of this employment is not formal, and that the average effect is much smaller and no longer statistically significant 4 months after the voucher period has ended. The voucher does appear to have persistent impacts outside of the capital, where it almost doubles the employment rate of graduates, but this appears likely to largely reflect displacement effects. Soft skills training has no average impact on employment, although again there is a weakly significant impact outside of the capital. We elicit expectations of academics and development professionals to demonstrate that these findings are novel and unexpected. The results suggest that wage subsidies can help increase employment in the short-term, but are no panacea for the problems of high urban female youth unemployment.

Keywords: Wage subsidy, Soft skills, youth unemployment, Randomized Experiment, Impact Expectations, Displacement.

JEL classification codes: O12, O15, J08, J16.

^{*} We thank the World Bank's Gender Action Plan, TFESSD and PSIA trust funds for funding this project. We thank Nithin Umpathi and Abdalwahab Khatib for their assistance in implementing the project, and seminar audiences in Jordan, at the IADB, Paris School of Economics, the University of Virginia, and the World Bank for useful comments. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors and should not be attributed to the World Bank or any affiliated organization.

1. Introduction

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 enormous challenge facing young women who want to work in the Middle East, and is a stark example of the lack of employment opportunities for educated youth occurring worldwide.

Firms typically give two main reasons for being reluctant to hire young women. First, they are often reluctant to hire youth because they lack work experience, require on-the-job training, and it may be costly to assess how good a worker they are. On top of this, they have doubts about how committed young women are to pursuing careers, and whether they are as flexible in working hours and travel as males. Secondly, firms often complain that formal schooling, at best, teaches only the technical skills workers need, but that many youth are lacking in the soft skills needed for success in the workplace – such as how to interact with customers, work in teams, act professionally, and even how to properly represent themselves in job interviews in the first place.

We conducted a randomized experiment in Jordan intended to overcome both these constraints and assist female community college graduates to find employment. Community college graduates are typical of many of the relatively educated unemployed in the Middle East, with concerns expressed about the quality and relevance of their classroom training, and relatively high unemployment rates. The Jordan New Opportunities for Women (Jordan NOW) pilot randomly allocated almost an entire cohort of female graduating students into four groups: a treatment group which received a job voucher which would pay an employer a subsidy in amount equal to the minimum wage if he or she hired the worker, valid for up to six months; a treatment group which was invited to attend a 45-hour employability skills training course designed to provide key soft skills demanded by employers; a treatment group which received both the voucher and the training; and a control group.

Follow-up surveys were conducted 6 and 14 months after the start of the interventions. The wage subsidies led to extremely large short-term gains in employment, particularly outside of Amman, with a 25 percentage point increase in employment in Amman and 50 percent increase outside of it. However, by the last survey, the voucher no longer had any impact on employment likelihoods in Amman, and the treatment effect had shrunk to 8.5 percentage points outside Amman and appears as if

it may be due to displacement of jobs for the control group. The soft skills training had no short-term impact on employment, but had a weakly significant 6.1 percent impact outside of Amman at endline, although this again may reflect displacement. Despite the lack of employment impact, the soft skills training did lead to improvements in life outlook and a reduction in depression, suggesting it may have benefits outside of the labor market.

Wage subsidies have long been used to help disadvantaged groups find jobs in developed countries, and there have been several randomized experiments to measure their impacts in the U.S. (Burtless, 1985; Dubin and Rivers, 1993), which have found disappointing impacts, which the authors attribute to potential stigma effects.¹ Several non-experimental studies have found some positive impacts (e.g. Katz, 1998), although an overview of different wage subsidy evaluations by Betcherman et al. (2004) concluded that such programs have largely not been effective in developed countries. Wage subsidy programs for youth have been used in a number of transition countries such as Poland and Slovakia, and there appears to be renewed policy interest in developing countries, with examples such as Morocco's *Idmaj* program and Tunisia's *SIVP* program², and South Africa about to launch a program. Despite this policy enthusiasm, there is very little evidence on the effectiveness of such programs in developing countries, the one exception being an experiment by Galasso et al. (2004) in Argentina. They found that job vouchers to the unemployed lead to a 6 percentage point increase in wage employment 18 months later, although this impact largely occurred in informal and temporary jobs.

In these existing evaluations, the rather limited effects have in part come from low usage rates of the job vouchers, preventing existing studies from seeing whether providing access to subsidized short-term employment can lead to lasting jobs. Since voucher take-up was relatively high in our study, we are able to examine this issue, as well as further contribute to the literature by providing evidence on the effectiveness of wage subsidies in a context where female skilled youth unemployment rates are very high.

¹ Larger positive short-term effects on employment were found by Card and Hyslop (2005) for the self-sufficiency project in Canada, which gave welfare recipients subsidies to offset the loss of welfare benefits if they found full-time work – however, in that case the firms were not told of the subsidy and the subsidy didn't directly lower the cost to firms of hiring workers. Like us, they find the positive effects dissipate when the subsidy ends.

² SIVP = Stage d'Initiation à la Vie Professionnelle. This is a subsidized internship where beneficiaries receive 150 TND monthly and employers have full coverage of social security and training costs. The subsidy targets recent university graduates who have been looking for a job for six months, a group especially affected by unemployment (Almeida et al. 2012).

There is even less empirical evidence on the effectiveness of soft skills training programs, but youth employment programs in Latin America in particular have increasingly included a component that focuses on these skills (World Bank, 2010). For example, the *entra 21* program implemented in 18 Latin American countries includes a soft skills training component. Although there is no rigorous evaluation of this intervention, employers report that participants who took part in this program have greater ability to work in teams and take on responsibility than do their other employees (*entra 21*, 2009). The Dominican Republic's *Juventud y Empleo* program also teaches soft skills along with providing work experience, with early evidence from Ibarra et al. (2012) finding this increased the formality of work, but didn't lead to any increase in overall employment in the short-run. Our paper provides the first experimental evidence on the effectiveness of soft skills training alone in a developing country context.

A common refrain following the presentation of experimental results or results from other rigorous impact evaluations is often that the findings "will surprise no one who had opened an economics textbook" (Scheiber, 2007). A further contribution of this study is to provide a way to quantify the extent to which the findings are novel or unexpected by means of an audience expectations elicitation exercise. The study was presented in seminars in academic settings and to development professionals, and described online on a popular development economics blog, and point expectations of the treatment impacts were elicited before the results were revealed. The results reveal considerable heterogeneity in expected outcomes from our interventions, and that the short-term impact of the voucher was much larger than almost everyone participating expected, while the long-term impact of the voucher and the impact of the soft-skills training were somewhat lower than people expected. As well as highlighting the unexpected nature of the findings in this study, we believe the approach pioneered here could be useful for other experimental studies to employ.

The remainder of this paper is structured as follows: Section 2 describes the context, our sample, and the details of the intervention; Section 3 the experimental design, data collection, and intervention take-up; Section 4 the midline and endline experimental impacts; Section 5 whether these impacts are in line with expectations or not; Section 6 then discusses possible explanations for these impacts; and Section 7 concludes.

2. Context, Sample, and the Intervention

2.1 Labor Market Context

Two striking features of labor markets throughout the Middle East are bulging youth populations with very high unemployment rates, particularly among relatively educated youth; and low labor force participation rates among women. Jordan is typical in this regard. Youth aged 15-29 are the largest demographic group in the country, making up 30 percent of the total population. In 2011 the unemployment rate for 20-24 year old females was 47.6 percent, compared to 23.1 percent for 20-24 year old males, and only 10.5 percent of 20-24 year old females are actually employed, compared to 49.1 percent of 20-24 year old males. Women with post-secondary education are more likely to be unemployed than women with secondary or primary education, largely because the latter do not participate in the labor force at all (Jordanian Department of Statistics, 2011).

Young women face twin constraints in accessing jobs. First, firms are often reluctant to hire youth, regardless of gender, since they lack job experience, are of untested quality, and may lack soft skills such as reliability, a strong work ethic, and knowledge of how to work and communicate effectively in a workplace (IYF, 2010). If hiring, training, and firing workers is costly, firms may be reluctant to take a chance on someone untested in the labor market. Second, young women face additional barriers because of gender. Employers often express clear preferences for male workers, based on the belief that women are less committed to their jobs and may leave if they get married or have children, that men are more flexible in work hours and ability to work over-time, and that women might experience more difficulties interacting with customers in some occupations due to culture (Jordanian Enterprise Survey, 2006; Peebles et al, 2007). When faced with these constraints and a lack of networks and role models who have obtained jobs, many young women may lack confidence to look for jobs in the first place.

2.2 Educational System Context and Focus on Public Community College Students

Students in the Jordanian educational system go through a common core curriculum that ends with the tenth grade, followed by two years of specialization where students choose between an academic track (focusing on either sciences or arts) and a vocational track. Both tracks end upon the completion of twelfth grade with a General Secondary Education Examination (the *Tawjihi*), which if passed concludes secondary education. Students who take the academic track in Arts or Science can then gain entrance to university provided they achieve a competitive score on the *Tawjihi*. Alternatively,

those taking part in the Vocational track, those who get a low Tawjihi score, or those with limited financial means can enroll in a two year community college – either as a terminal qualification with skills in a particular field, or as a second chance at university admissions (Kanaan and Hanania, 2009).

Analysis of labor force survey data from 2007 shows that, among 25 to 30 year olds in Jordan, 67 percent of females and 72 percent of males have at most secondary education. Amongst the third of women this age with higher levels of education, one-third has a diploma from a community college, and two-thirds have a university degree. There are 14 public community colleges in Jordan, with total enrollment of approximately 11,895 students (7,072 female, 4,823 male) in 2007/08. These public community colleges have significantly lower tuition than universities and private community colleges.³ However, as in much of the Middle East, there are concerns about both the quality and usefulness of some of the training being offered, leading to concerns about the employability of community college graduates, particularly in a labor market with limited jobs in which they may be outcompeted by university students. This context led the Government of Jordan to request assistance from the World Bank in conducting a pilot program to try to increase employment of female graduates.

To conduct this pilot we choose the 8 public community colleges with the largest female enrolment numbers, together comprising over 85 percent of all female public community college enrolment. They consist of four colleges in Central Jordan (Amman University College, Princess Alia University College, Al-Salt College, Zarqa University College) and four located in Northern and Southern Jordan (Al-Huson University College for Engineering, Irbid University College, Ajloun University College, and Al-Karak University College). For simplicity's sake since Amman is the capital of Jordan and many people from adjacent governorates commute to Amman for work, we will herein refer to Central Jordan as Amman and Northern and Southern Jordan as outside Amman.

Baseline surveys were conducted in July 2010 for all second-year students in these colleges, giving data on 1755 female students, before students had taken their final examinations. In August 2010 this was then merged with administrative data on examination results, which revealed that 1395 had passed their examinations. We randomly selected 1350 of these graduates to be our experimental sample.

³ There were approximately 14,500 students enrolled in 22 private community colleges in 2007/08 according to Government statistics.

The typical graduate is 20 to 22 years old, is unmarried, and has never worked before. Only 13.8 percent were married at baseline, and 16.3 percent have ever worked. Only 7 percent of these young women's mothers are currently employed, whereas 57 percent of their fathers are currently employed. At the time of the baseline survey, which took place only weeks before final examination results were available, only 8 percent of students had already found a full-time job for work after graduation.

Table 1 shows the main courses of study undertaken by the experimental sample at the overall program of study level, as well as at the level of specialization. We see the majority of students are taking courses in administration and finance (43%), which covers specializations such as accounting, electronic administration and management information systems; courses in medical assistance (24%), which covers mainly nursing and pharmacy specializations; and educational programs (10%), which covers those aiming to be teachers.

2.3 Pilot Interventions

We piloted two policy interventions, each designed to overcome some of the barriers to firms hiring young female graduates: wage subsidy vouchers and employability skills training. Wage subsidies have a long history of use by policy makers as part of their active labor market policies to generate employment for the disadvantaged (e.g. Kaldor, 1936, Layard and Nickell, 1980, and Katz, 1998). It is argued that short-term subsidies may have long-term effects by raising the productivity of youth through work (Bell et al., 1999), and may encourage employers to take a chance on hiring inexperienced, untested workers (World Bank, 2006). There is also now growing evidence that non-cognitive or soft skills are important for employment and a range of other life outcomes (e.g. Bowles et al, 2001; Heckman et al., 2006; Heckman and Kautz, 2012). Interventions that aim to teach employability skills may enhance employment prospects by giving youth better skills and confidence for looking for jobs and by making them more productive in their first months in the job by reducing the amount of time firms need to spend training them on the basics of working in a business environment.

Our pilot program was marketed to participants as the Jordan New Opportunities for Women (Jordan NOW) pilot. The details of our interventions are as follows:

Job Vouchers or Wage Subsidies

Graduates receiving this intervention were given a job voucher that they could take to a firm while searching for jobs. The voucher had the graduate's name on it, and was non-transferable. The job

voucher paid the employer an amount equal to the mandatory minimum monthly wage of 150 JD (USD 210) per month for a maximum of six months if they hired the worker, acting as a wage subsidy. To be eligible to use the voucher a firm had to provide proof of registration, have a bank account to receive payment in, and provide an offer letter with the graduate's name and specification of work duties. The salary agreed for employment must be at least the minimum wage of 150 JD per month. We did not require registration of workers in the social security system for eligibility, so employers were subject to the existing law on this, which in principle requires workers to be registered after three months with the firm. After the start of employment, both the firm and graduate were required to confirm their employment with the program administrator each month, with periodic monitoring and random visits made to ensure reimbursement claims were legitimate.

The voucher was valid for a maximum of six months within an eleven month period starting October 3, 2010 and ending August 31, 2011. Should employment terminate before the end of the 6 months, the voucher remains with the graduate, who could then use the remaining months on it with a different firm. In addition to a letter explaining the program which students could give firms, the program was advertised through the Chamber of Commerce and through newspaper advertisements, and an official government website and information helpline was provided, in order to further the legitimacy of the voucher and provide more information to firms as needed.

Due to budget constraints, we could afford to fund a maximum of 450 graduates for the full 6 months each. Graduates receiving this intervention were informed that vouchers would be honored on a first-come-first-served basis and that they should therefore take every effort to use their vouchers as soon as possible to find a job, and that they would be notified if this cap was reached. In practice this cap did not bind.

Employability Skills Training

Graduates receiving this intervention were invited to free intensive training on soft skills identified by employers as important areas for new graduates to have. The training course was 45 hours over a 9 day period (5 hours per day), with a maximum of 30 participants in each training group. Training took place during September and October 2010. The training was provided by Business Development Center (BDC)⁴, a Jordanian NGO established in 2005 which has widespread local name recognition and a good reputation for skills training, having implemented USAID, UNCTAD, and a wide variety of local

⁴ <http://www.bdc.org.jo/>

training programs. Training sessions took place in 17 sessions offered throughout 6 governorates to maximize access. Training facilities and training content were identical across all 17 sessions. To minimize the effect of social and cultural restrictions on mobility, sessions were held during daylight hours at locally known and trusted institutions such as the Chambers of Industry and local universities.

BDC designed a course which covered effective communication and business writing skills (e.g. making a presentation, writing business reports, different types of correspondence), team-building and team work skills (e.g. characteristics of a successful team, how to work in different roles within a team), time management, positive thinking and how to use this in business situations, excellence in providing customer service, and C.V. and interviewing skills. Sessions were based on active participation and cooperative learning rather than lectures, with games, visual learning experiences, group exercises, and active demonstrations used to teach and illustrate concepts.

The cost of the training was approximately \$150,000, which was based on up to 600 graduates attending – leading to a cost of \$250 per assigned graduate, and given that only 373 attended, an effective cost of \$400 per attendee.

3. Experimental Design, Data Collection, and Take-up

Randomization into treatment and control groups was done via computer. Students were first stratified into 16 strata on the basis of geographic region Amman (Amman, Salt and Zarqa), and outside Amman (Irbid, Ajloun and Karak), whether their Tawjihi examination score at the end of high school was above the sample median or not, whether they indicated at baseline that they planned to work full-time and thought it was likely or somewhat likely that they would have a job within 6 months of graduating, and whether or not she is usually permitted to travel to the market alone (a measure of empowerment). Within each strata, 22.2% of the students were allocated to receive the job voucher program only, 22.2% allocated to receive the employability skills training only, 22.2% allocated to receive both the job voucher program and the employability skills training program, and 33.3% allocated to the control group. This resulted in 300 in each treatment group, and 449 in the control group, for a final experimental sample of 1349.⁵ It turned out that two of the individuals assigned to receive a job voucher (one from the voucher only group, and one from the voucher and training group) were actually males who had incorrectly been recorded as female in the baseline questionnaire. These were dropped from

⁵ Rounding within strata resulted in 449 rather than 450 being allocated to the control group in total.

the program, given a final breakdown of 299 in job voucher only, 300 in training only, 299 receiving training and voucher, and 449 in the control group.

The choice of variables to stratify on was based on two considerations. First, stratifying on the basis of variables which we believe would influence the main outcomes of interest (employment) can improve the power of an experiment to detect a given sized treatment effect (Bruhn and McKenzie, 2009). Since few graduates were employed at baseline, we did not stratify on employment status itself, but instead on key variables we thought might predict employment. We hypothesized that the likelihood of getting a job will be higher for students around Amman than in other areas, for those with higher academic ability (proxied by the Tawhiji score), for those with the desire and confidence to work, and for those with greater empowerment or freedom of movement (proxied by whether they are allowed to travel to the market alone).

Second, stratifying on these variables prevents against chance imbalances in these characteristics, and serves as a means for specifying ex ante our interest in examining the heterogeneity of treatment effects according to these characteristics. A priori it was difficult to predict which direction this heterogeneity would act in. For example, we expected graduates in Amman to be more likely to find work in the absence of an intervention since the majority of private sector activity is concentrated around the capital, and because families are more traditional outside of Amman and thus perhaps more reluctant to allow their daughters to work. However, it was unclear whether the interventions would then work better in Amman because there would be fewer other constraints on finding work, or whether they would have less effect there if it is the case that anyone who wants to work should be able to find a job, whereas outside of Amman where it is more challenging to find a job, only those who receive assistance might be able to find one. Likewise it was unclear whether the interventions would act as a complement or a substitute for higher academic ability, higher desire to work, or higher empowerment.

3.1 Baseline Information by Treatment Status

Table 2 provides summary statistics on the experimental sample by treatment status. As one would expect given computerized randomization, the characteristics look similar across the different treatment groups. At baseline, the graduates express high levels of desire to work, with 93 percent saying they plan to work after they graduate, 91 percent say they would like to work outside the house after they are married, and 82 percent say they think it is very likely or somewhat likely that they will

have a job within 6 months of graduating. Our stratifying variable is whether they desire to work, think it is likely they will be working, and plan on working full-time, with 59 percent of graduates falling into this high likelihood of working full-time category.

One hypothesis for low levels of female employment is that patriarchal norms limit the extent to which women work. We have noted that 91 percent say they would like to work outside the home after marriage, and the baseline also shows that 91 percent think they will be allowed to do so. Most women say that the main reason for obtaining an education is to increase one's earnings (45%) or to find a job with better working conditions (32%), with only 3% saying the main reason is to improve marriage prospects. However, there is evidence that some women do face mobility restrictions, with only 51 percent saying they are allowed to go to the market alone. Table 2 also shows that 81 percent of the graduates say they would prefer a public sector job to a private sector job, which is consistent with evidence from elsewhere in the Middle East.

3.2 Data collection

A midline follow-up survey was conducted in April 2011, six months after graduates initially received their vouchers and/or attended training, and while the job voucher usage window was still open. The midline succeeded in interviewing 1237 of the 1347 (92%). In December 2011 a more extensive endline survey was conducted, with more extensive efforts to track down graduates. This resulted in interviewing 1249 of the 1347 (93%) and collecting data by proxy from 38 of the 1347 graduates' parents (3%). Both rounds collected data on employment outcomes, and in addition, the endline collected information on a range of well-being measures, including mental health, subjective well-being, and empowerment.

In addition to the survey data, an agreement with the Social Security Corporation of Jordan enabled us to obtain administrative data on formal employment for 1282 graduates (95% of the sample), with this data not available for those individuals for whom we did not have the social security number. Finally, we also supplement our analysis with data taken from an October 2011 survey of 368 of the firms who had employed these graduates at the midline (whether or not they had used the voucher to do so), which is approximately 100 percent of all firms who were employing a graduate with a voucher and more than 67% of all firms who were employing a graduate without a voucher.

The attrition rates are low, but for both the surveys, vary by treatment status (appendix table 1). Attrition rates in the survey are lowest (3% midline, 1% endline) for the job voucher treatments, since

more recent contact information was available for them as a result of the process of monitoring their voucher usage and take-up, and highest for the control group (11% midline, 7% endline), since there was no contact with them outside of the surveys. Almost all the attrition was due to inability to contact the individuals due to cell phone numbers changing, rather than direct refusals. In contrast, there is no difference in attrition rates by treatment status in the social security data. In appendix 1 we show that our main results are robust to bounding approaches to deal with selection.

3.3 Take-up of the Interventions

The employability skills training course was completed by 373 of the 599 graduates (62%) assigned to the training only or training and voucher treatments. Only 5 students attended part of the course but did not finish. Qualitative feedback from participants immediately after the course was universally extremely positive, with many participants in particular saying it had given them more confidence and positivity, as well as noting their appreciation for learning practical topics not taught in college. Correlates of take-up are shown in Table 3 and discussed in Appendix 3.

Take-up of the voucher is equivalent to finding a job with an employer that met the voucher requirements and was willing and able to use the voucher. In total 301 of the 598 graduates assigned to receive a voucher ended up using it for at least one month (50.3%), although this varied greatly with location, with only 35% of those offered the voucher using it around greater Amman, compared to 65% outside of Amman. Three-quarters of those who used the voucher used it for the full 6 months, 9 percent used it for 5 months, 6 percent for 4 months, and only 10 percent for 3 months or fewer. In the midline survey, 85 percent of those employed with a voucher say they earn 150 JD per month, which is the amount of the voucher, 1.9 percent say they were paid less than this, and the highest earnings is 320 JD per month. The most common occupations for those using the vouchers were teachers (often in nursery schools), comprising 33% of all those who had used the voucher; secretaries, clerks or administrative assistants, 17%; nurses or medical assistants, 10%; data entry workers 9%; and pharmacists 8%.

3.4 Correlates of employment in the control group

The remaining columns of Table 3 examine the correlates of being currently employed for the control group at the time of the midline (columns 5 and 6) and endline (columns 7 and 8) surveys. The purpose of this analysis is to see whether the correlates of take-up are similar or different to those of working in the absence of an intervention. We see a strong contrast to the job voucher take-up results in

terms of location effect: only 10-11 percent of the control group is employed at the time of the midline and endline surveys outside of Amman, whereas 28 percent are employed in Amman at the time of the midline, and 38 percent at the time of the endline. The much lower take-up of job vouchers inside Amman therefore doesn't reflect lower employment rates in Amman.

The control group is less likely to be employed if they said at baseline that they didn't think they would be working full-time after graduation, and if they were married at baseline. Students who took administrative and financial courses are less likely to be employed, which is in line with the job voucher take-up results. Household wealth, tawjihi score, and baseline mobility do not predict employment among the control group after conditioning on the other variables in these regressions.

4. Results

To evaluate the impact of being offered one or both of the interventions, we estimate the following equation for graduate i via OLS:

$$Outcome_i = \alpha + \beta_1 Voucher_i + \beta_2 Training_i + \beta_3 Both_i + \sum_{s=1}^{16} \delta_s d_{i,s} + \varepsilon_i \quad (1)$$

where $Voucher_i$ is a dummy variable taking the value one if graduate i was assigned to receive a job voucher (either alone or with training), $Training_i$ is a dummy variable taking the value one if graduate i was assigned to be invited to the training course, and $Both_i$ is a dummy variable denoting the individual was offered both treatments. Individuals who got both treatments therefore have value one on all three of the $Voucher$, $Training$, and $Both$ variables. The $d_{i,s}$ are the randomization strata dummies. The coefficient β_1 then measures the intention-to-treat (ITT) effect of the job voucher, β_2 the ITT of the training, and β_3 whether there is an interaction effect between the two treatments. Since we are explicitly interested in whether the impacts differ in the midline survey - when the vouchers are still in effect - from the endline survey after they have ended, we estimate equation (1) separately for the midline and endline data.

We focus on ITT impacts which give the average effect of being offered training or the voucher, rather than the effect of actually attending training or taking up the voucher. We choose not to estimate the treatment effect on the treated because, especially in the case of the job voucher, it seems plausible that being offered the treatment may have impacts on employment outcomes even if the treatment is not actually used. Indeed Galasso et al. (2004) find evidence of this in a wage subsidy experiment in Argentina, and they suggest that one main effect of vouchers in their experiment was to encourage

workers to exert more effort finding work and to give them more confidence approaching employers, even though actual take-up of the vouchers was low. Similarly, one could imagine that the offer of training or a voucher may affect job search behavior even if the treatments are not taken up.

In addition to average impacts, we examine heterogeneity in treatment response with regard to the four variables used to stratify the randomization. To do this we estimate the following equation for interaction $Interact_i$:

$$\begin{aligned}
 Outcome_i = & \alpha + \beta_1 Voucher_i + \beta_2 Training_i + \beta_3 Both_i + \\
 & \gamma_1 Voucher_i * Interact_i + \gamma_2 Training_i * Interact_i + \gamma_3 Both_i * Interact_i \\
 & + \sum_{s=1}^{16} \delta_s d_{i,s} + \varepsilon_i \quad (2)
 \end{aligned}$$

where the strata dummies capture the level effects of the interaction term, and $Interact_i$ is either a dummy for being in Amman, for not expecting to work full-time at baseline, for having above median tawjihi grade, or for being allowed to travel to the market alone at baseline (denoted “empowered”).

4.1 Average impacts on different dimensions of employment

Table 4 reports the results of estimating equation (1) for different employment outcomes. Panel A reports the results at the time of the midline survey, and Panel B the results at the time of endline survey. We begin in column 1 by looking at the impact on labor force participation, defined as either working or actively looking for work. At the time of the midline, labor force participation was high for everybody, with 77 percent of the control group participating. The treatments had no additional impact on labor force participation at this time. In contrast, by the time of the endline, labor force participation had fallen to 48 percent of the control group, reflecting graduates stopping actively looking for work. The job voucher treatment has a statistically significant 10 percentage point increase in labor force participation, while the training has an insignificant 5 percentage point increase (although we can also not reject the effect is equal to the voucher effect), and there is no interaction effect.

Columns 2 through 4 then look at employment. Column 2 looks at whether individuals are currently employed or have worked for cash in the last month. In panel A we see a large and strongly significant impact of the job vouchers on this measure at midline. The 39.5 percentage point increase in employment more than triples the employment rate of 17.8 percent in the control group. In contrast, training only has a small (3 percentage point) and statistically insignificant impact on employment, and

there is no interaction effect between training and the job voucher. However, this impact does not persist once the voucher period has expired, with panel B showing that the job voucher group was only an insignificant 2.8 percentage points more likely to be employed at endline, with still no significant impact of training.

A key question is whether the jobs created by the voucher reflected genuine work opportunities in their field or study, or makeshift less-skilled jobs in which graduates had few chances to practice and enhance their skills? The vast majority of young women hired with the voucher were hired as teachers, nurses and pharmacists, accountants, and secretaries or business administrators.⁶ Spot checks by the implementation unit verified that those hired were actually working in these jobs doing typical tasks for someone in these positions. Despite the baseline preference for public sector work, almost all jobs were in the private sector, with median firm size of ten workers, of which 60 percent on average were women. The jobs created thus were in the field of study, in relatively small private firms that already had previously hired women. This was very similar to the types of firms and jobs that the control and training group who were employed were working in.

Columns 3 and 4 look at formal employment, defined by whether individuals report being employed and registered for social security in the survey (column 3), and in the social security administrative data (column 4). We see that almost all of the employment created by the voucher at the time of the midline is not formal employment, with only a 4.5 percentage point increase in formal employment according to survey reports, and no significant increase according to administrative records. There is no significant impact on formal employment at endline either. Outside of the voucher group, formal employment accounts for just over half of all employment - 12.6 percent of the control group were formally employed according to the endline survey.

Although the job voucher group is no more likely to be working at endline, they have more job experience as result of having had worked. Columns 5 and 6 show they are 27 percentage points more likely to have ever worked than the control group at endline, and have accumulated an average of 2.4 months more job experience.

Finally we look at the intensity of work and the earnings from it. Column 7 looks at weekly hours worked, coding hours as zero for those not working. Hours worked are significantly higher for the job

⁶ Consistent with this, we find no significant difference in treatment effect by the three main fields of study (business administration, medical assistance, and education).

voucher group at midline, but not significantly different at endline. The impact on work income is examined in column 8. The job voucher group earns 64 JD more per month than the control group at midline, which has the effect of more than tripling the 25 JD per month the control group earns. However, by the endline the difference has fallen to only 6 JD per month and is not statistically significant. To see whether the higher income of the voucher group in the midline is just coming from them being more likely to work, or also from them earning higher wages conditional on working, in column 9 we examine the impact on wages conditional on working. Since working is an outcome which is itself affected by the treatment, randomization does not guarantee that the treatments are uncorrelated with other determinants of conditional wages, and so this should be considered exploratory and suggestive evidence only. We find that the job voucher appears to have increased by 23 JD the wages conditional on working at the midline, but that there is no significant difference at endline. Also noticeable is that the mean wage for the control group at the midline is 141 JD per month, less than the minimum wage of 150 JD.

We also examined whether any of the interventions lead to changes in the characteristics of the firms in which workers are employed at endline. We see no difference by treatment status in the number of workers or number of female workers at the firm, or in whether the firm is owned by a woman.⁷ Nevertheless, conditional on having had a job, graduates assigned to the job voucher group report statistically higher levels of job satisfaction.

4.2 Heterogeneity in Employment Impact

Table 5 reports the results of estimating equation (2) to examine how the impact of the different interventions on employment varies with the stratifying variables. Again panel A shows impacts at midline, and panel B at endline. Employment status is again defined as being currently employed or having worked for cash in the past month, regardless of formal status.

Column 1 shows how the treatment effect varies with geographic location. Consistent with the much low take-up of the voucher in Amman than outside of Amman, we see that the voucher had much larger impacts on midline employment outside of Amman. Graduates assigned to receive the voucher outside Amman experienced a 50.4 percentage point increase in the likelihood of being employed at midline, compared to a 25 percentage point increase in Amman. Given the control group had much lower employment rates outside of Amman, this is equivalent to the job voucher group having six times

⁷ Results available upon request.

the employment rate of the control group outside Amman, and double the employment rate in Amman. Moreover, the endline results show that the voucher treatment continued to have a significant impact on employment outside Amman even once the subsidy period had finished, as well as some evidence that the employability skills training also had an effect at endline outside Amman. Only 11 percent of the control group are employed outside Amman at endline, with an increase of 8.5 percentage points for the job voucher treatment and 6.1 percentage points for the training treatment (and a negative but insignificant interaction effect). In contrast, the point estimates suggest 4-5 percentage point lower employment rates for the voucher and training treatment groups in Amman than for the control group, and we can't reject the null of no treatment effects in Amman.

Figure 1a and 1b use information from the surveys on the start and end dates of employment to construct and plot monthly employment rates by treatment status to graphically illustrate these impacts. We see the voucher only and voucher plus training lines track each other, rising rapidly at first, and then falling as the graduates hit the six month limits on use of these vouchers. The rise is steeper outside of Amman. The difference in control group behavior inside Amman versus outside is also noticeable – employment rates continue to rise over time inside Amman, but level off at 11 percent outside Amman and hover around this rate for at least one year.

In contrast, columns 2 through 4 of Table 5 show no significant interaction effects with the other stratification variables, with the coefficients generally small in magnitude. Moreover, the sign of the coefficients often varies between the midline and endline, showing no consistent tendency for the interventions to have differential impacts according to baseline expressed likelihood of working fulltime, tawjihi score, or ability to travel to the market alone. This lack of interaction effect with treatment is not because these stratifying variables do not have an impact on employment prospects – the control group means show that employment rates are higher for those with more academic ability, those who face fewer mobility restrictions, and those who at baseline want to work full-time and think they will do so. The treatments just do not significantly interact with these determinants of employment.

Table 4 showed a significant impact on labor force participation at endline, despite the lack of significant employment effect. When we examine this effect by location, it appears to be coming entirely from outside of Amman – the impact is 17.6 percentage points ($p < 0.001$) outside of Amman, and 0.3 percentage points in Amman ($p = 0.95$).

4.3 Impact on Well-being, Empowerment and Attitudes

Jobs can be more than just a source of income, with employment status associated with improved subjective well-being and increases in female empowerment (World Bank, 2011, 2012). In Table 6 we examine the impacts of the interventions on different measures of well-being, empowerment, and attitudes, all measured at the time of the endline survey.

Column 1 examines the impact on current subjective well-being, measured on the Cantril self-anchoring striving scale (Cantril, 1965), a measure that has been used by Gallup around the world. Respondents are asked to imagine a ladder with 11 rungs, number from 0 at the bottom to 10 at the top, where the top represents the best possible life for them, and the bottom the worst possible life. Kahneman and Deaton (2010) refer to this as “life evaluation”. The control group mean (s.d.) is 5.0 (2.4) on this scale, and being assigned to the job voucher intervention results in a significant 0.58 unit increase, while training has an insignificant 0.28 increase. There is a significant and large negative interaction effect for receiving both treatments, so that graduates assigned to both treatments have no better current life evaluation than the control group.

Graduates were also asked at endline to assess which step on the Cantril ladder they believe they will be on in five years time, with this forward-looking measure reflecting the degree of optimism they have about their futures. Overall the graduates show a high degree of optimism, with the mean of 8.1 (s.d. of 1.8) for the control group three rungs higher than they assess their current position to be. Column 2 shows that the training intervention leads to additional optimism: graduates assigned to this group think they will be 0.27 steps higher than what the control group thinks. In contrast, there is a negative and insignificant impact of the job voucher on this measure.

Mental health is a distinct concept of well-being from happiness, and has been shown to have different associations with individual characteristics and with life events (Das et al, 2008). We measure mental health using the Mental Health Inventory (MHI-5) of Veit and Ware (1983). This is a five item scale with a maximum score of 25 and minimum score of 5, with higher scores indicating better mental health in terms of the experience of psychological well-being and the absence of psychological distress in the past month. While there is no universal cutoff, several studies have used a cutoff of less than 17 as an indicator of major depression (e.g. Urban Institute, 1999; Yamazaki et al., 2005). Twenty percent of the control group have scores below this threshold. Columns 3 and 4 look at the impact on the overall index, and on the binary classification of depression respectively. We find that the training intervention

results in a significant 0.58 increase in the MHI-5, and a 4.8 percentage point reduction in the likelihood of having major depression. The job voucher has no significant impact on either measure. These results are consistent with the ladder of life in the future questions, and is consistent with the qualitative feedback right after the training intervention, which suggested that it had really lead to strong positive attitudes.

Columns 5 and 6 of Table 6 examine whether these changes in subjective well-being have also brought about changes in attitudes towards women's role in home and society, and changes in empowerment, measured by ability to go to different places by themselves. The data appendix defines each variable in more detail. We see that there is no significant impact on attitudes towards the role of women: at endline only 51 percent disagree that a girl must obey her brother's opinion even if he is younger than her, and only 65 percent agree that boys should do as much domestic work as boys, with the interventions not changing these attitudes. Part of the reason for the lack of effect might be that there is already strong agreement on attitudes related to work – 97 percent think women should be allowed to work outside the home. The last column examines the impact on the number of locations out of 6 that a graduate is allowed to go to by herself. The mean in the control group is 5.2, with those who only get the job voucher being permitted to go alone to 0.54 fewer locations. Training has a small negative but insignificant impact. The interaction between job voucher and training is strongly positive though, so that graduates assigned to receive both treatments are permitted to go alone to 0.30 more locations than the control group.

Finally, we note that while at baseline only 13.7 percent of the experimental sample were married, this had increased to 31.6 percent by the endline, only 18 months later, with a further 9.4 percent engaged at endline. Only 9.6 percent of married graduates in our sample are working at endline, compared to 29.9 percent of those who are engaged and 32.5 percent of those who are single. There is thus a strong association between getting married and not working, which makes it of interest to see whether any of the interventions had any impact on the likelihood of being married by the endline. Column 7 shows that the treatment effects are all small and statistically insignificant, so that there is no impact on marriage of the interventions.

We examined heterogeneity in these well-being impacts by geographic location. Both the job voucher ($p < 0.01$) and training interventions ($p < 0.10$) have stronger impacts on current subjective well-being inside of Amman, whereas there is no significant geographic heterogeneity in treatment impacts on position on the life-ladder 5 years from now, on mental health, or on the number of locations the

graduate is allowed to go to alone. The job voucher has significantly more impact on empowered attitudes inside of Amman, although the effect size is relatively modest (0.27) and only significant at the 10 percent level.

5 Are these impacts unexpected?

Ex post it is relatively easy for policymakers and academics to view the result of almost any impact evaluation and claim that the results are exactly what they would have expected. To understand whether this study merely confirms conventional wisdom or generates new and unexpected results, we therefore undertook an audience expectations elicitation exercise. During the first four presentations of this research (two in academic departments, and two in international organizations)⁸, no paper was distributed or available in advance. Instead, the audience was presented with 25 slides detailing the motivation, existing evidence, context, design, and implementation of this study. We then distributed paper surveys throughout the audience and asked them to provide their point estimates of the midline and endline ITT impacts for each treatment group: voucher, training, and both combined. We also asked the audience whether they thought the voucher treatment impacts would be larger, the same, or smaller by each of our four stratifying variables. To complement the audience expectations elicitation exercises, we posted a description of the study on the World Bank's Development Impact blog⁹ with a link to an online survey to capture reader expectations.¹⁰ The advantage of seminar presentations is that they offer time to answer any questions, explain the intervention in detail, and get high response rates (very few seminar attendees refused to fill in the one-page expectations sheet). In contrast, online elicitation opens up the process to a wider audience, but offers less of a chance to explain the intervention, and has low and self-selected response rates.¹¹ Finally, a simplified version of the

⁸ This comprises presentations in May and June 2012 at the Applied Micro workshop at the University of Virginia; the development seminar at the Paris School of Economics; the applied micro seminar at the World Bank; and a seminar with the labor group of the InterAmerican Development Bank. The World Bank and IADB seminars contained a mix of researchers and development professionals engaged in the implementation of programs. We thank the seminar audiences for their participation in this exercise.

⁹ <http://blogs.worldbank.org/impactevaluations/are-our-blog-readers-better-predictors-of-impact-results-than-seminar-audiences-evaluating-programs>

¹⁰ To our knowledge this is the first paper to systematically collect audience expectations for an impact evaluation. Dean Karlan and Annie Duflo concurrently collected qualitative expectations during an online poll at the Stanford Social Innovation Review

http://www.ssireview.org/articles/entry/can_management_consulting_help_small_firms_grow.

¹¹ The Development Impact blog post had 900 page views, along with approximately 2000 RSS and Email subscribers (of which only some would have actually read the post). From this we received 48 responses, a click-through rate consistent with click-through from blog post links to academic papers (McKenzie and Özler, 2011).

expectations was also elicited from nine Jordanian policymakers in a presentation of the results made in Jordan, who were just asked expectations in five bin ranges (<0, 0, 0 to 5%, 5% to 10%, more than 10%).

Figure 2 provides histograms of the resulting expectations, while Tables 7a and 7b summarize the resulting expectations. We see that only 4 percent of the 136 respondents gave an expected value of the voucher impact at midline that lies within the 95 confidence interval of the treatment effect, with the median expected impact of 8 percentage points less than one-quarter of the actual impact of 39.5 percentage points. In contrast the modal and median response of a 5 percentage point impact at endline is close to the 3 percent estimated impact, although there is still substantial heterogeneity in expectations with a long right tail to the distribution. The median expected impact of training was 5 percentage points in both the midline and endline, which was close to the estimated ITTs, although the mean expected impacts were 9-10 percentage points, reflecting that many respondents expected the soft skills training to have sizeable impacts and thus overestimated the impact. Audiences believed the combination of the voucher and training was likely to have larger impact than either treatment alone, but also dramatically underestimated the training impact at midline. None of the respondents had all six of their expectations lie within the 95 percent confidence intervals of the actual impacts. Furthermore, there was considerable heterogeneity among respondents in their relative rankings of the interventions: at midline (endline) 59 (47) percent thought the voucher would have a larger impact than the soft skills training, 11 (14) percent that the impact would be the same, and 30 (39) percent that the voucher would have less impact than the soft skills training. The policymakers' responses also show considerable heterogeneity in responses and that the majority also underestimated the short-term impact of the voucher.

These results therefore show that i) the midline impact of the voucher is larger than most people would expect; ii) there is considerable uncertainty as to what the impacts of such a program are likely to be; iii) that the impacts of the soft skills training are less than expected on average; and iv) the lack of complementarity between the two interventions was a surprise to the average respondent. There was also considerable heterogeneity in beliefs about how the voucher treatment would vary with randomization stratifying variables: 48 percent thought the effect would be stronger in Amman versus 36 percent outside Amman (the rest thought the impact would be the same in both places; 42 percent thought the impact would be higher for high academic ability, versus 41 percent for low academic ability; 56 percent thought the impact would be higher for the more mobile, versus 20 percent for the

less mobile; and 62 percent thought the impact would be higher for those more interested in work at baseline, versus 17 percent for those less interested in work.¹²

6 Discussion

The above analysis shows that the job voucher treatment led to a large (and unexpected) increase in employment during the period which the voucher could be used, but almost entirely in jobs not registered with social security, and that most of this employment effect disappeared by the time of the endline. The impacts of the voucher were much greater outside of Amman than inside (in contrast to what many people would expect), with some suggestion that training also had an impact outside of Amman. In this section we explore possible mechanisms that lead to these results.

6.1. Are temporary impacts due to the job voucher groups losing jobs or the other groups gaining them?

One potential explanation for the short-term impact of the wage subsidies would be that they speeded up the process of finding a job, with the training and control groups then managing to find jobs and catch up to their employment rates by the endline survey. An alternative explanation is that the reduced impact is coming from those who were employed losing their jobs when the vouchers ended. To distinguish between these explanations, in Table 8 we explore the employment dynamics, looking at the 2 x 2 transition matrices for employment and unemployment between the midline and endline.

The results show that part of the reduction in impact is due to the control and training groups being more likely to transition into employment, but the majority is due to the voucher and voucher and training groups being much more likely to transition out of employment. For the full sample, 12 percent of the control group and 13 percent of the training group found jobs between the midline and endline surveys, compared to only 5 percent of the voucher and voucher and training groups. But 37-38 percent of the training and training and voucher groups lost their jobs between midline and endline, compared to only 7-8 percent of the control and training only groups. Thus voucher group graduates losing their jobs accounts for approximately 80 percent of the closing of the gap in employment rates between

¹² In the interests of space we do not discuss in detail here differences among the different audiences. The UVA and IADB audiences underestimated more the midline impact of the voucher, while the Paris School of Economics audience overestimated more both the midline and endline impacts of the soft-skills training and the endline impact of the voucher. Blog respondents (of whom only 20 percent were World Bank staff) had similar average responses to the World Bank seminar audience.

voucher and the other groups, and catch-up of those newly getting jobs accounts for approximately 20 percent of the gap. This basic pattern holds both inside and outside Amman.

The endline survey directly asked graduates who had been employed with vouchers but were no longer in these jobs what the main reason for stopping work was. The most common reason was that the job ended because the voucher had finished, and the job was temporary in nature – this accounted for approximately 70 percent of the job exits for jobs obtained through voucher. Only 2 percent were for other employer-related reasons such as the employer firing them or the employer going bankrupt, while the remaining 28 percent were because the graduates decided to quit – mostly because they were not satisfied with some aspect of the job, including travel and hours, or because they quit for family reasons. In a minority (4%) of cases they quit because of salary disputes, with firm owners wanting to lower their wages once the subsidy ended. Among those who remained in the job, there are only 5 cases where the midline wage was 150 JD and the endline wage less than this (mean of 98 JD), but more cases where the midline wage was above 150 JD but has been lowered (e.g. from 220 JD at midline to 150 JD at endline).

When the same question was asked of firms which had hired graduates with the vouchers, but no longer employed this worker as of November 2011, firm owners said that in 51 percent of the cases it was because the worker was unaffordable without the subsidy, 10 percent because they had fired the worker, 30 percent because the worker had quit either to get another job, or to get married, and 7 percent other reasons. Firms were relatively more likely to say the reason was the graduate quitting in Amman than outside of Amman (42% vs 30%), and less likely to say the reason was that the worker was unaffordable without the subsidy in Amman. In only 36 percent of the cases, firm owners say they would have hired the worker who later lost their job without the job voucher, and the main reasons they did hire these workers were to train and test them risk free (40%), and to have an extra employee at low cost (32%).

Taken together, this evidence suggests that the main reason that the impact of the voucher was mostly temporary was that the hard cash of the wage subsidy induced firms to take a chance on hiring workers they wouldn't have otherwise hired, and that these workers then either proved not to be productive enough to earn the wages they would need to be paid, or that ex post, these workers decided that the characteristics of the job were not a good match for them.

There are two other reasons why the vouchers may not have longer term effects. First, one potential goal of the voucher was to induce employers who had not hired women before to give them a

chance, thereby overcoming prejudice and giving the graduates a chance to prove they could be productive. However, almost all of the voucher recipients were employed in typically female-dominated occupations (such as nurse, nursery school teach, or clerk), so there was no new information about the productivity of women generated for most firms. Secondly, schools and hospitals may find it harder to generate additional profits out of new workers than other sectors of the economy (e.g. because they have less control over pricing), so even productive workers may be difficult for these firms to finance.

6.2 Is this a gender or a youth effect?

Our initial survey of students at these eight community colleges also surveyed the males. However, there were only 427 male students, of which 345 passed their examinations. This small sample coupled with budget limitations lead us to restrict the experiment to females. A consequence is that we do not know whether the impacts of the treatments reflect attributes of youth, or are a feature of the gender of these youth. However, our surveys show much higher employment rates for the male graduates of these same community colleges. At midline 54 percent of male graduates inside Amman and 51 percent outside Amman were employed, which rises to 75 percent inside Amman and 61 percent outside Amman by endline. After controlling for faculty of study and specific college major, men in Amman (outside Amman) are 30 (45) percentage points more likely than women in Amman to be employed. Clearly, female youth have a particularly hard time finding jobs relative to their male peers.

6.3 Are labor laws partly responsible for the temporary nature of the job?

Article 35 of Jordan's labor law specifies a 3 month probationary period, during which an employer can terminate a worker without notification or termination remuneration. After this period, employers are required to give one month's notification, and remunerate workers one month per year of service on a pro-rated basis upon termination. The length of the job voucher period was set at 6 months with this 3 month rule in mind, the idea being that the six month subsidy may induce them to hire graduates beyond the three month promotion period and thereby bring graduates into the formal employment system. We see that in practice this does not occur – most of the added employment was not registered for social security, and only 5 percent of vouchers were used for three months. The threat of being subject to labor regulations in the future may still deter firms from keeping youth employed for long periods, but this was not something that came out in the firm surveys – and thus it seems that these labor laws were not the main constraint to firms keeping the voucher hires employed, because they were able to successfully avoid these regulations.

In contrast, 84 percent of the graduates employed with the voucher at the midline were hired at a wage of exactly 150 JD per month – which was the minimum permitted by the program. Coupled with the evidence above that a prime reason that graduates were let go by their firms once the subsidy period ended, this suggests that minimum wages which are higher than marginal products may be one important regulatory reason that the impacts were temporary and that youth unemployment rates are so high. The main incentive to avoid registering workers in social security is likely to be avoiding social security taxes (18.75% of wages) and payroll taxes (7% of wages), which together add 25% to the cost of employing a worker.

6.4. Did the interventions just change who got the jobs, or actually create new jobs?

A common concern with many active labor market policy experiments is the possibility of spillovers or general equilibrium effects. In this experiment, there are two elements of this concern. The first is a concern about interference among the experimental sample. In particular, the concern would be that the voucher, and perhaps training, groups gained jobs at the expense of the control group, so that there is no net increase in employment, just a reallocation within sample of who gets the jobs. If this were the case, while our experiment would still give an internally valid estimate of the impact of giving vouchers to some youth and not others, one could not extrapolate from this to estimate the impact of offering this program to all community college graduates. A second, related, concern is whether the jobs gained by the community college graduates are coming at the expense of other workers outside of the experimental sample who would otherwise have been hired. For example, the vouchers may induce firms to hire youth instead of an unemployed older worker. If this were the case, the experiment would still show that these policies help a disadvantaged group obtain jobs, but not whether there are costs to others in society of doing so.

In a large country with segregated labor markets one could experimentally address this issue by randomizing the intensity of the treatments, as was done by Crepón et al. (2011) in France. This was not a possibility in a small country like Jordan, and so we use a mixture of evidence to assess how important these spillovers are likely to have been.

We note first that the fact that most of the effect was temporary, and came from firms saying they hired the job voucher workers when they wouldn't have hired them otherwise suggests that most of the short-lived effect is additional (temporary) hires, rather than firms substituting hires they would have made anyway. Second, we do not see firms who let go of the job voucher workers subsequently

hiring a control group or training group worker to replace them: there are only 12 firms in our firm survey that hired graduates from both the voucher and from either the control or training only groups (almost all hospitals hiring nurses), and all of these were cases of concurrent hires rather than terminating a voucher student and replacing them with one of the other group students.

However, our intervention worked with approximately 80 percent of females graduating from public community colleges in Jordan in 2010, and thus if there are a limited number of jobs that these graduates are competing for, it seems plausible that they are likely to have been competing for some of the same jobs, resulting in displacement effects. Indeed, when the graduates were asked in the endline survey whether they think the voucher prevented women without vouchers from getting jobs because employers would only hire workers with vouchers, 12 percent of the control group in Amman, and 24 percent of the control group outside Amman agreed. This shows that the control group graduates themselves think there is some displacement outside of Amman.

Further evidence on this displacement comes from looking at the employment rates in other recent years. Table 9 uses the 2007 to 2010 Jordanian Employment and Unemployment Surveys¹³ to report the employment rates and labor force participation rates of intermediate diploma students (the group community college students fall into), and compares these to our endline employment rates. We see that the employment rate in central Jordan (Amman) for our control group is similar to that of community college graduates in recent years, whereas that outside Amman is lower.¹⁴ Taking the difference between the two locations, we see the employment gap between Amman and outside Amman for the control group is more than double in sample than it is in recent years. Coupled with the direct evidence from graduates, we view this as compelling evidence of a displacement effect – graduates outside Amman in the control group appear to have not been hired at the same rates that recent years would suggest, or at the rate that one would predict given the employment rate of the control group in Amman.

This displacement also appears to be taking place in terms of labor force participation outside of Amman. Labor force participation for intermediate diploma students outside of Amman over the 2007 to 2010 period averaged 53.5 percent, and was never more than 5 percentage points below that in

¹³ We thank the Hashemite Kingdom of Jordan Department of Statistics for supplying us with these summary statistics.

¹⁴ Note that although the control mean is higher than the Central Jordan employment rate in previous years, the confidence intervals overlap. We focus on the Central Jordan vs Northern and Southern Jordan difference.

Amman. In contrast, although labor force participation was 72.9 percent for the control group outside Amman (and 82.9 percent in Amman) at the time of the midline survey, this had fallen to 38.7 percent outside Amman by the endline survey (compared to 60.9 percent for the control group in Amman). Of this group of control group individuals outside Amman who stopped looking for work, 45 percent say it was because they are pessimistic about the chance of finding a job or believe no job is available in the area, 30 percent say they have given up because of marriage, pregnancy, or looking after other family members, and 12 percent because they are waiting on a response from a government agency they have applied to.

The main occupations for graduates employed outside Amman at the endline are teacher (23%), nurse (19%), pharmacist (16%), and clerk (11%). The labor market outside Amman in most of these occupations is likely to be relatively thin, with a limited number of openings for new graduates each year. It does therefore seem reasonable that graduates may have been competing for some of the same jobs, and that, in addition to the temporary additional hires firms made using the vouchers, they chose voucher or training graduates over the control group for positions they were planning on hiring anyway. As a result, the treatment effects seen in the endline outside of Amman are likely to reflect largely displacement, rather than added employment.

7. Conclusions

Wage subsidies and soft skills training are two popular types of policies that governments are turning to around the world as part of their efforts to deal with high youth unemployment. Our experimental analysis shows they do not appear to have had large impacts on generating sustained employment for young, relatively educated women in Jordan. Short-term wage subsidies generated large and significant increases in employment while the subsidies were in effect, but most of these jobs disappeared when the subsidies expired. High minimum wages may be one reason, with firms saying that graduates were not productive enough to be affordable without subsidies. Since our intervention ended, the minimum wage has been raised even higher, suggesting young women will continue to struggle to find paid employment.

Using an audience expectations elicitation exercise we show that there is considerable heterogeneity in beliefs about the likely effectiveness of such programs, and substantial underestimation of the short-term impacts of the voucher program. This both serves as an illustration of an approach that could be used in other studies to reveal the extent to which impact evaluations

confirm or contrast with existing priors, as well as serving to show that the results generated here are different from what many people would expect.

The wage subsidy intervention did succeed in getting graduates to have work experience they otherwise would not have had, while the soft skills training intervention resulted in improvements in positive thinking and mental health. It is possible that this experience and/or positive outlook may help graduates over a longer period of time, but the fact that we do not see employment impacts 16 months after graduation shows the tremendous challenge in getting this population into work. Interventions to address supply-side constraints that prevent firms from creating more jobs, especially jobs for young women, may instead be needed to address the problem of persistent low employment for women throughout most of the Middle East.

References

Almeida, Rita, Juliana Arbelaez, Maddalena Honorati,, Arvo Kuddo, Tanja Lohmann, Mirey Ovadiya, Lucian Pop, Maria Laura Sanchez Puerta and Michael Weber (2012) “Improving Access to Jobs and Earnings Opportunities: The Role of Activation and Graduation Policies in Developing Countries”, *World Bank Social Protection and Labor Discussion Paper no. 1204*.

Bell, Brian, Richard Blundell and John Van Reenen (1999) “Getting the Unemployed Back to Work: The role of targeted wage subsidies”, *International Tax and Public Finance* 6(3): 339-60.

Betcherman, Gordon, Karina Olivas, and Amit Dar (2004) “Impacts of Active Labor Market Programs: New Evidence from Evaluations with Particular Attention to Developing and Transition Countries”, *World Bank Social Protection Discussion Paper no. 402*.

Bowles Samuel, Herbert Gintis, and Melissa Osborne (2001) “Incentive-enhancing Preferences: Personality, Behavior and Earnings” *American Economic Review* 91(2):155–158

Bruhn, Miriam and David McKenzie (2009) “In pursuit of balance: randomization in practice in development field experiments”, *American Economic Journal: Applied Economics* 1 (4), 200–232.

Burtless, Gary (1985) “Are Targeted Wage Subsidies Harmful? Evidence from a Wage Voucher Experiment” *Industrial and Labor Relations Review*, 39(1): 105–15.

Cantril, Hadley (1965) *The pattern of human concerns*. New Brunswick, NJ: Rutgers University Press.

Card, David and Dean Hyslop (2005) “Estimating the effects of a time-limited earnings subsidy for welfare-leavers”, *Econometrica* 73(6): 1723-70.

Crépon, Bruno, Esther Duflo, Marc Gurgand, Roland Rathelot and Philippe Zamora (2011) “Do labor market policies have displacement effect? Evidence from a cluster randomized experiment”, Mimeo.

- Das, Jishnu, Quy-Toan Do, Jed Friedman and David McKenzie (2008) "Mental health patterns and consequences: Results from survey data in five developing countries", *World Bank Economic Review* 23(1): 31-55.
- De Mel, Suresh, David McKenzie and Christopher Woodruff (2010) "Wage Subsidies for Microenterprises", *American Economic Review Papers and Proceedings*, 100(2): 614-18.
- Dubin, Jeffrey, and Douglas Rivers (1993) "Experimental Estimates of the Impact of Wage Subsidies." *Journal of Econometrics*, 56(1/2): 219-42.
- Entra 21 (2009) *Final Report on the entra21 Program Phase I 2001-2007*. International Youth Foundation.
- Galasso, Emanuela, Martin Ravallion and Augustin Salvia (2004) "Assisting the Transition from Workfare to Work: A Randomized Experiment", *Industrial and Labor Relations Review* 57(5): 128-42.
- Heckman, James, Jora Stixrud, and Sergio Urzua (2006) "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior", *Journal of Labor Economics* 24(3):411-482.
- Heckman, James and Tim Kautz (2012) "Hard evidence on soft skills", IZA Discussion Paper no. 6580.
- Ibarraran, Pablo, Laura Ripani, Bibiana Tapoada, Juan Miguel Villa and Brigida Garcia (2012) "Life Skills, Employability and Training for Disadvantaged Youth: Evidence from a Randomized Evaluation Design", IZA Discussion Paper no. 6617.
- International Youth Foundation (IYF) (2010) "Building on hope: Findings from a rapid community appraisal in Jordan", Youth: World Jordan report, the International Youth Foundation, Baltimore, MD.
- Jordanian Department of Statistics "Table 2.4: Jordanian Population Age 15+ Years by Activity Status, Sex and Broad Age Groups (Percentage Distribution) – 2011", [http://www.dos.gov.jo/owa-user/owa/emp_unemp_y.show_tables1_y?lang=E&year=2011&t_no=18](http://www.dos.gov.jo/owa/user/owa/emp_unemp_y.show_tables1_y?lang=E&year=2011&t_no=18) [accessed 24 April, 2011].
- Kaldor, Nicholas (1936) "Wage subsidies as a remedy for unemployment", *Journal of Political Economy* 44(6): 721-42.
- Kahneman, Daniel and Angus Deaton (2010) "High income improves evaluation of life, but not emotional well-being", *PNAS* 107(38): 16489-16493
- Kanaan, Taher and May Hanania (2009) "The Disconnect between Education, Job Growth, and Employment in Jordan", *Generation in Waiting*, Brookings Institution.
- Katz, Lawrence (1998) "Wage subsidies for the disadvantaged", in Richard Freeman and Peter Gottschalk (eds.) *Generating Jobs: How to Increase Demand for Less-skilled workers*. Russell Sage Foundation: New York, NY.
- Layard, P.R.G. and S.J. Nickell (1980) "The case for subsidizing extra jobs", *Economic Journal* 90(357): 51-73.

McKenzie, David and Berk Özler (2011) "The impact of economics blogs", *World Bank Policy Research Working Paper no. 5783*.

Peebles, Dana, Nada Darwazeh, Hala Ghosheh, and Amal Sabbagh (2007) "Factors affecting women's participation in the private sector in Jordan", <http://www.almanar.jo/AlManaren/Portals/0/PDF2/Mayssa%20Gender%20report.pdf> [accessed 24 April, 2011].

Scheiber, Noam (2007) "Freaks and Geeks; How Freakonomics is ruining the dismal science", *The New Republic*, April 2. <http://www.tnr.com/article/freaks-and-geeks-how-freakonomics-ruining-the-dismal-science> [accessed July 9, 2012].

Urban Institute (1999). *Snapshots of America's Families: Appendix*, <http://www.urban.org/url.cfm?ID=900875>.

Veit, C.T. & Ware, J.E. (1983). "The structure of psychological distress and well-being in general populations", *Journal of Consulting and Clinical Psychology*, 51, 730-742

World Bank (2010) "Active labor market programs for youth: A framework to guide youth employment interventions". World Bank Employment Policy Primer No. 16.

World Bank (2011) *World Development Report 2012: Gender Equality and Development*. World Bank: Washington, D.C.

World Bank (2012) *World Development Report 2013: Jobs*. World Bank: Washington, D.C.

Yamazaki, S., Fukuhara, S., Green, J., (2005) "Usefulness of five-item and three-item Mental Health Inventories to screen for depressive symptoms in the general population of Japan" *Health and Quality of Life Outcomes*, 3/1/48.

Table 1: Most Common Courses of Study for Experimental Sample

Program Code Level	%	Specialization Level	%
The Administrative and Financial Program	43	Nursing	13
Program of Medical Assistance	24	Accounting	12
Educational Program	10	Electronic administration	12
Performing Arts Program	7	Management information systems	10
Social Action Program	6	Other -- Educational Programs	9
Information Management and Libraries Program	6	Pharmacy	5
Engineering Program	2	Interior Design & Graphic	5
Science Program of Sharia and Islamic Civilization	2	Special education	5
Hotels Program	1	Information technology	4
Agriculture Program	0	Accounting Information Systems	4

Table 2. Comparison of Means of Baseline Characteristics by Treatment Group

	Voucher Only	Training Only	Voucher & Training	Control Group
<i>Stratifying Variables</i>				
In Amman, Salt, or Zarqa	0.43	0.44	0.43	0.44
Tawjihi score above median	0.55	0.55	0.55	0.55
Low desire to work full time	0.41	0.41	0.41	0.41
Is allowed to travel to the market alone	0.51	0.51	0.51	0.51
<i>Other Baseline Variables</i>				
Age	21.2	21.1	21.1	21.3
Married	0.14	0.16	0.12	0.13
Mother Currently Works	0.07	0.06	0.08	0.06
Father Currently Works	0.59	0.61	0.57	0.53
Has Previously Worked	0.15	0.18	0.16	0.16
Has a Job Set Up for After Graduation	0.05	0.08	0.10	0.08
Has Taken Specialized English Training	0.31	0.26	0.26	0.30
Household Owns Car	0.62	0.66	0.62	0.64
Household Owns Computer	0.72	0.75	0.74	0.70
Household Has Internet	0.28	0.18	0.26	0.26
Prefers Government Work to Private Sector	0.82	0.81	0.79	0.81
Sample Size	299	300	299	449

Note: The only statistically significant difference across groups is internet access which is significant at the 10% level.

Table 3. Take-Up Regressions, and Correlates of Employment in the Control Group

	Training Take-up		Voucher Take-up		Midline Employment in Control group		Endline Employment in Control Group	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Stratifying Variables</i>								
Amman, Salt, or Zarqa	0.025 (0.040)	-0.022 (0.042)	-0.174*** (0.041)	-0.164*** (0.043)	0.171*** (0.040)	0.160*** (0.041)	0.270*** (0.042)	0.242*** (0.043)
Tawjihi score above median	-0.086** (0.040)	-0.079** (0.039)	0.001 (0.041)	-0.004 (0.041)	0.041 (0.037)	0.017 (0.037)	0.003 (0.039)	-0.026 (0.039)
No desire to work full time	-0.099** (0.041)	-0.097** (0.041)	-0.047 (0.042)	-0.034 (0.041)	-0.073** (0.037)	-0.054 (0.036)	-0.082** (0.038)	-0.069* (0.038)
Is allowed to travel to the market alone	-0.031 (0.040)	-0.022 (0.039)	0.069* (0.041)	0.071* (0.041)	0.012 (0.037)	0.019 (0.037)	-0.005 (0.039)	-0.010 (0.039)
Dual Treatment Group	0.059 (0.039)	0.051 (0.039)	0.007 (0.040)	0.006 (0.040)				
<i>Other Variables according to Baseline Status</i>								
Married at Baseline		-0.195*** (0.060)		-0.093 (0.061)		-0.100*** (0.036)		-0.127*** (0.036)
Wealth Index		0.003 (0.011)		0.013 (0.012)		-0.013 (0.011)		0.011 (0.011)
Number of brothers		0.020** (0.010)		0.015 (0.012)		-0.007 (0.010)		-0.008 (0.009)
Number of sisters		-0.007 (0.009)		-0.008 (0.009)		-0.006 (0.008)		0.005 (0.008)
Has E-mail		0.115** (0.054)		-0.016 (0.058)		0.072 (0.071)		0.102 (0.074)
In Admin/Finance Program		0.115*** (0.040)		-0.111*** (0.042)		-0.116*** (0.036)		-0.104*** (0.038)
Sample Size	599	599	598	598	398	398	419	419

Note: Huber-White standard errors in parentheses. *, **, and *** indicate significance at the 10, 5 and 1% levels respectively.

Table 4: Impacts on Different Dimensions of Employment

	Labor Force Participation	Employed	Employed and registered for Social Security (Survey)	Employed and registered for Social Security (Admin data)	Ever Employed	Months Employed Since Graduation	Hours Worked Last Week	Work Income (not conditional on working)	Work Income (conditional on working)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<i>Panel A: Midline Results (April 2011)</i>									
Assigned to Voucher	0.028 (0.031)	0.395*** (0.035)	0.045** (0.022)	0.005 (0.020)	0.357*** (0.036)	1.538*** (0.178)	13.416*** (1.301)	64.498*** (5.783)	23.730*** (8.175)
Assigned to Training	-0.002 (0.033)	0.031 (0.031)	0.034 (0.022)	-0.002 (0.019)	0.059* (0.034)	0.238 (0.174)	1.100 (1.172)	5.709 (4.828)	5.599 (10.350)
Assigned to Both	0.055 (0.046)	-0.022 (0.052)	-0.055 (0.034)	-0.010 (0.029)	-0.035 (0.054)	-0.159 (0.273)	-0.166 (1.963)	-4.245 (8.598)	-6.637 (11.142)
Sample Size	1,237	1,237	1,237	1,282	1,237	1,237	1,237	1,237	448
Control Mean	0.771	0.178	0.072	0.068	0.234	0.776	6.595	24.927	141.729
<i>Panel B: Endline Results (December 2011)</i>									
Assigned to Voucher	0.100*** (0.037)	0.028 (0.032)	0.016 (0.026)	-0.015 (0.026)	0.272*** (0.036)	2.456*** (0.393)	1.534 (1.426)	5.573 (6.061)	5.006 (10.096)
Assigned to Training	0.054 (0.037)	0.015 (0.032)	0.039 (0.027)	-0.005 (0.026)	0.036 (0.037)	0.115 (0.368)	1.534 (1.448)	0.790 (5.783)	-11.631 (10.602)
Assigned to Both	-0.044 (0.055)	-0.025 (0.048)	-0.066* (0.039)	-0.012 (0.038)	0.037 (0.054)	-0.104 (0.574)	-1.455 (2.201)	-1.291 (9.060)	13.850 (15.108)
Sample Size	1,287	1,287	1,249	1,282	1,250	1,249	1,249	1,249	312
Control Mean	0.484	0.229	0.148	0.126	0.388	2.626	9.579	39.534	168.958

Note: Huber-White standard errors in parentheses. *, **, and *** indicate significance at the 10, 5 and 1% levels respectively.

All regressions also control for stratification dummies. Outcome of "Ever Employed" not available in midline survey.

Table 5. Heterogeneity of Employment Impact by Randomization Stratification Variables

Interaction	Amman	High Aptitude	Low Desire to work full-time	Allowed to Travel to Market Alone
Dependent Variable: Employed	(1)	(2)	(3)	(4)
<i>Panel A: Midline Results (April 2011)</i>				
Assigned to Voucher	0.504*** (0.043)	0.426*** (0.050)	0.380*** (0.046)	0.339*** (0.049)
Assigned to Training	0.019 (0.033)	0.062 (0.043)	0.022 (0.042)	0.049 (0.044)
Assigned to Both	0.021 (0.064)	-0.041 (0.076)	-0.013 (0.069)	0.024 (0.075)
Voucher*Interaction	-0.254*** (0.071)	-0.057 (0.069)	0.036 (0.070)	0.111 (0.069)
Training*Interaction	0.021 (0.064)	-0.056 (0.062)	0.023 (0.061)	-0.033 (0.062)
Both*Interaction	-0.086 (0.105)	0.035 (0.104)	-0.022 (0.105)	-0.094 (0.104)
Sample Size	1,237	1,237	1,237	1,237
Control Mean when Interaction=0	0.10	0.14	0.22	0.163
Control Mean when Interaction =1	0.28	0.21	0.12	0.19
<i>Panel B: Endline Results (December 2011)</i>				
Assigned to Voucher	0.085** (0.037)	0.013 (0.044)	-0.007 (0.043)	0.031 (0.045)
Assigned to Training	0.061* (0.035)	0.015 (0.045)	0.012 (0.044)	-0.026 (0.042)
Assigned to Both	-0.059 (0.057)	0.022 (0.068)	-0.019 (0.065)	-0.004 (0.066)
Voucher*Interaction	-0.130* (0.066)	0.028 (0.063)	0.088 (0.064)	-0.005 (0.064)
Training*Interaction	-0.104 (0.066)	0.001 (0.063)	0.009 (0.062)	0.080 (0.063)
Both*Interaction	0.078 (0.099)	-0.085 (0.096)	-0.016 (0.096)	-0.039 (0.096)
Sample Size	1,287	1,287	1,287	1,287
Control Mean when Interaction=0	0.11	0.21	0.28	0.22
Control Mean when Interaction =1	0.39	0.25	0.16	0.23

Note: Huber-White standard errors in parentheses. *, **, and *** indicate significance at the 10, 5 and 1% levels respectively. All regressions also control for stratification dummies.

Table 6: Impacts on Wellbeing, Empowerment, Attitudes, and Marriage at Endline

	Life Ladder (1)	Life Ladder Future (2)	MHI 5 (3)	Severely Poor MH (4)	Mobility Index (5)	Empowerment Index (6)	Married (7)
Assigned to Voucher	0.577*** (0.180)	-0.201 (0.159)	-0.177 (0.276)	0.017 (0.031)	-0.526*** (0.141)	-0.060 (0.068)	0.026 (0.036)
Assigned to Training	0.283 (0.185)	0.266** (0.131)	0.582** (0.281)	-0.048* (0.029)	-0.072 (0.120)	-0.018 (0.072)	-0.003 (0.036)
Assigned to Both	-1.005*** (0.265)	-0.071 (0.210)	-0.553 (0.415)	0.030 (0.045)	0.834*** (0.186)	0.084 (0.102)	-0.032 (0.053)
Sample Size	1,249	1,249	1,249	1,249	1,249	1,249	1,249
Control Mean	4.970	8.128	19.266	0.197	5.259	4.847	0.313

Note: Huber-White standard errors in parentheses. *, **, and *** indicate significance at the 10, 5 and 1% levels respectively. All regressions also control for stratification dummies.

Severely Poor MH indicates MHI-5 index is below 17.

Table 7a: Expected Impacts of the Interventions from seminar participants and online respondents

	N	Actual Treatment Effect	μ_{expected}	σ_{expected}	median _{expected}	% of expectations within 95% C.I.
Expected Impact at Midline						
Voucher only	136	40	12	12	8	4
Training only	136	3	9	10	5	63
Voucher and Training	135	40	17	14	11	9
Expected Impact at Endline						
Voucher only	137	3	10	12	5	60
Training only	139	2	10	11	5	63
Voucher and Training	135	2	16	15	10	38

Table 7b. Impact Expectations of Jordanian Policymakers (N=9)

	Midline Impact				Endline Impact			
	0% or less	0 to 5%	5 to 10%	10% or more	0% or less	0 to 5%	5 to 10%	10% or more
Voucher	22%	22%	11%	44%	22%	22%	33%	22%
Training	33%	22%	11%	33%	33%	11%	44%	11%

Table 8: Employment Transitions (Percentage Transitioning from one state to another between survey rounds)

	Voucher Group		Training Group		Both Treatments		Control Group	
	Endline unemp.	Endline employ.	Endline unemp.	Endline employ.	Endline unemp.	Endline employ.	Endline unemp.	Endline employ.
<i>Panel A: Entire Sample</i>								
Midline unemployed	37.5	5.2	66.1	13.3	37.1	4.7	70.4	12.2
Midline employed	36.8	20.5	8.1	12.6	37.8	20.5	6.6	10.9
Sample Size	288		278		271		395	
<i>Panel B: Amman, Salt and Zarqa</i>								
Midline unemployed	36.8	9.6	50.8	17.5	41.9	7.3	52.1	20.4
Midline employed	28.0	25.6	13.3	18.3	25.8	25.0	7.8	19.8
Sample Size	125		120		124		167	
<i>Panel C: Outside Greater Amman</i>								
Midline unemployed	38.0	1.8	78.2	9.9	33.1	2.6	83.8	6.1
Midline employed	43.6	16.6	4.0	8.0	47.4	16.9	5.7	4.4
Sample Size	163		151		154		228	

Table 9: Employment rates of 20-25 year old Female Community College Graduates

	2007	2008	2009	2010	Control group	Voucher group	Training only
Central Jordan	31	32	31	29	40	34	36
Northern and Southern Jordan	19	19	20	18	11	19	16
<i>Difference</i>	13	13	11	11	29	15	20

Source: 2007 to 2010 from Jordan Employment and Unemployment Surveys for 20-25 year old females with intermediate diplomas. Survey standard error is approximately 2 percentage points on sample means.
Voucher group includes both voucher only and voucher plus training group.

Figure 1: Employment rates over time by location and treatment status

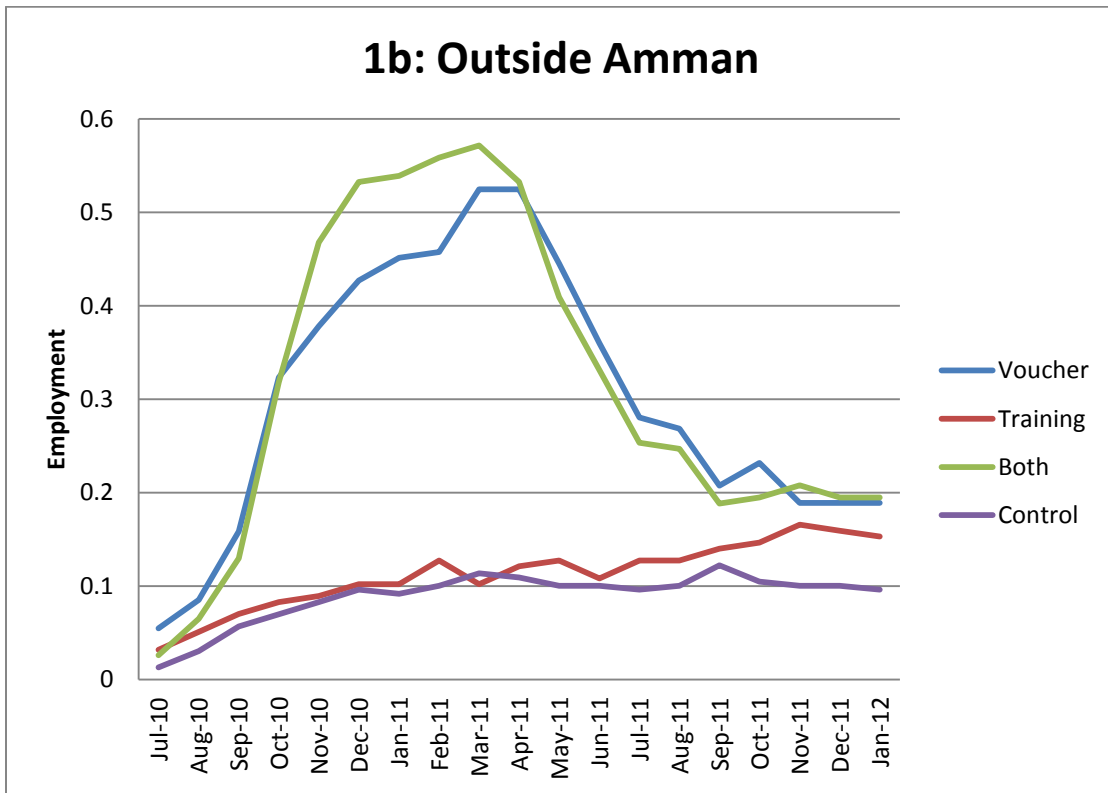
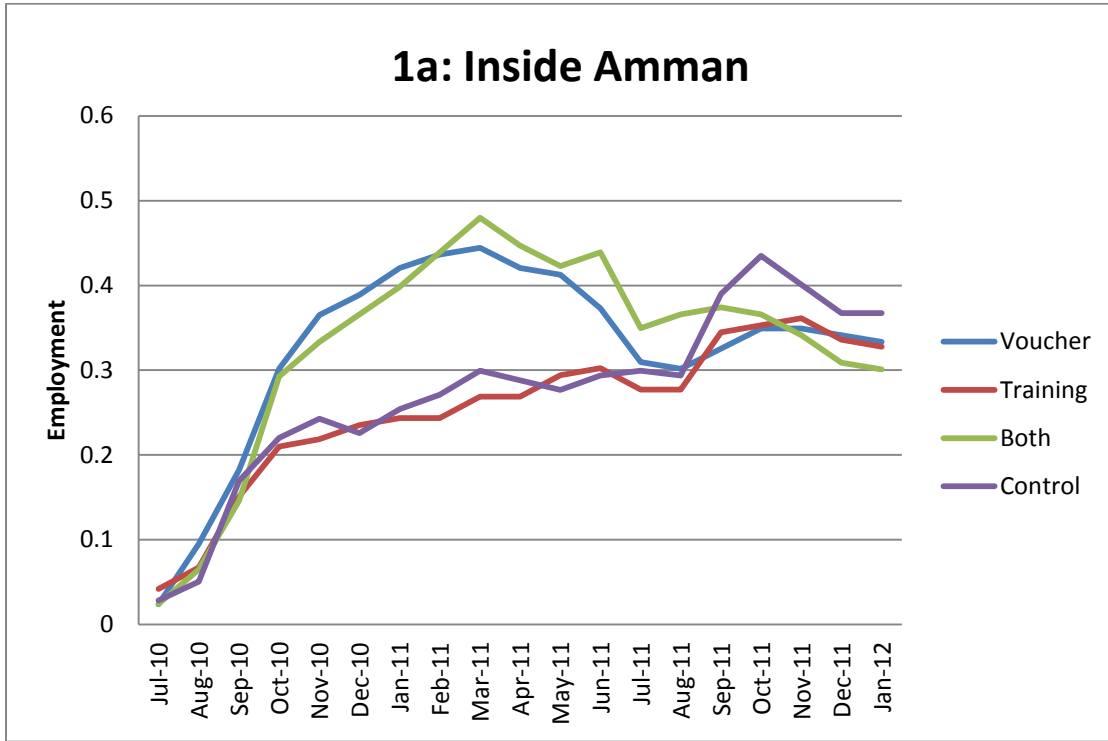
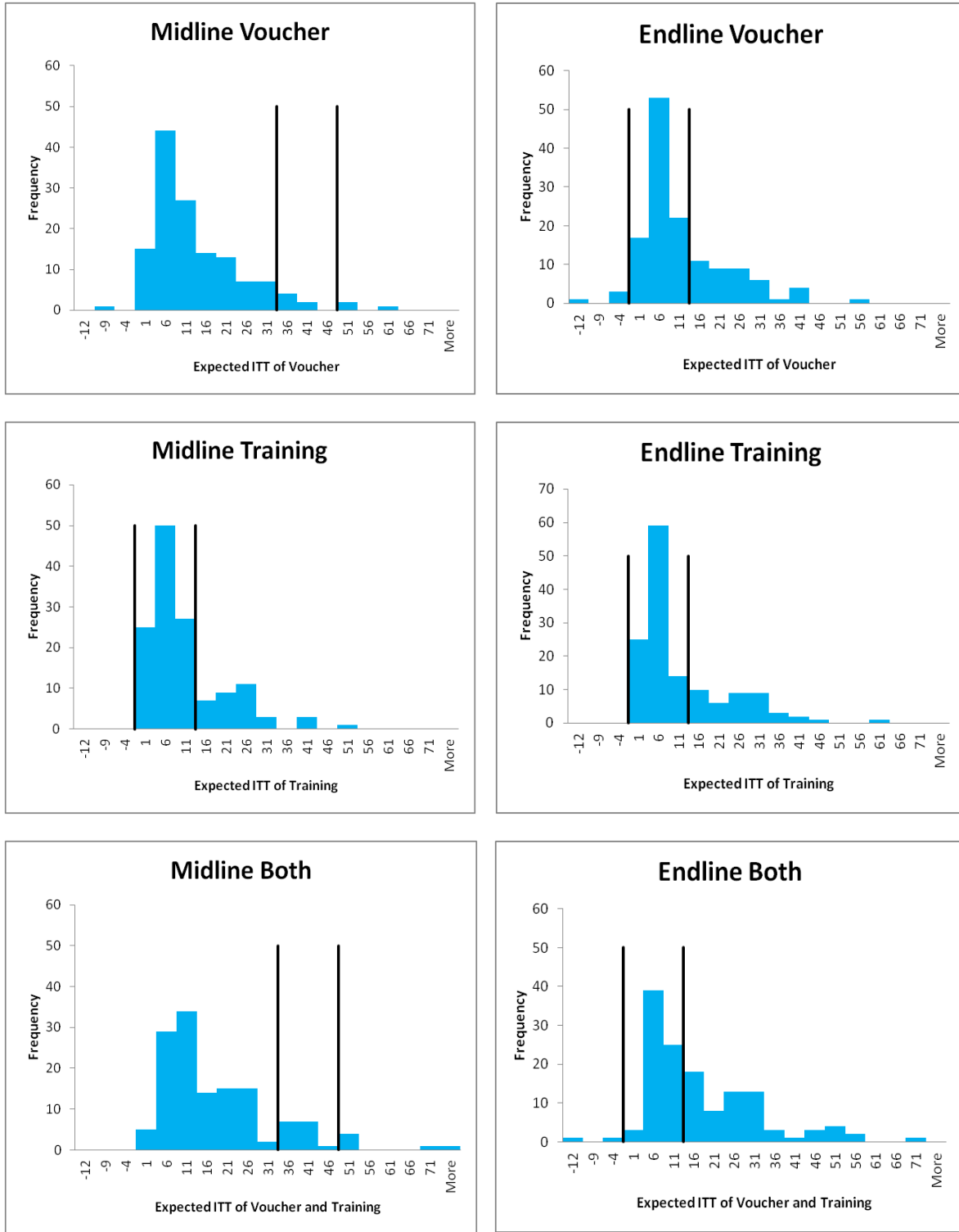


Figure 2: Histograms of Distribution of Audience Expectations of ITT Impacts on Employment



Vertical lines show 95 percent confidence interval for estimated treatment effect. Histograms show distribution of audience expectations.

Appendix 1: Robustness of Results to Attrition

Table A1 shows attrition rates by treatment status and round. Attrition levels are low, but do vary slightly with treatment status, being lowest for the job voucher group. To assess the sensitivity of our employment results to survey attrition, Table A2 provides conservative upper and lower bounds on our point estimates. To construct the lower bound, we assume that all missing control group and training only group individuals were employed, whereas all missing voucher and voucher and training individuals were unemployed. The upper bound reverses this. Our main findings of a large and strongly significant impact of the voucher on employment at midline, and a small and insignificant effect at endline continues to hold at either bound.

Table A1. Attrition by Treatment Status

	Midline	Endline	Social Security
Voucher	0.03	0.01	0.04
Training	0.09	0.05	0.04
Both	0.07	0.04	0.03
Control	0.11	0.07	0.07
Total	0.08	0.04	0.05

Table A2: Upper and Lower Bounds on Voucher Effect due to attrition

	Midline Employment			Endline Employment		
	Point Estimate	Lower Bound	Upper Bound	Point Estimate	Lower Bound	Upper Bound
	(1)	(2)	(3)	(4)	(5)	(6)
Assigned to Voucher	0.395*** (0.0347)	0.284*** (0.0354)	0.431*** (0.0331)	0.0285 (0.0319)	-0.0261 (0.0323)	0.0508 (0.0313)
Assigned to Training	0.0313 (0.0309)	0.0117 (0.0330)	0.0321 (0.0282)	0.0153 (0.0315)	-0.000514 (0.0327)	0.0197 (0.0302)
Assigned to Both	-0.0221 (0.0519)	-0.0247 (0.0526)	-0.00830 (0.0491)	-0.0250 (0.0478)	-0.0161 (0.0480)	-0.00294 (0.0473)
Sample Size	1,237	1,347	1,347	1,287	1,347	1,347

Note: Huber-White standard errors in parentheses. *, **, and *** indicate significance at the 10, 5 and 1% levels respectively. All regressions also control for stratification dummies.

Appendix II: Variable Definition

- *Amman* – a dummy indicating the respondent attended community college in Amman, Salt, or Zarqa, which are all located in Central Jordan
- *Employed* – a dummy indicating the respondent worked for pay in the last month or is currently employed full or part-time.
- *Employed and Registered for Social Security (Administrative)* – a dummy indicating the respondent is formally recorded in the Social Security Corporation’s administrative database
- *Employed and Registered for Social Security (Survey)* – a dummy indicating the respondent reports being formally registered with the Social Security Corporation of Jordan
- *Empowerment Index* – a proxy for female empowerment on a scale from zero to six with a higher score indicating greater empowerment. The index is computed as the number of pro-female responses: 1. Disagree, “A thirty year old woman who has a good job but is not yet married is to be pitied” 2. Agree, “Women should occupy leadership positions in society” 3. Agree, “Women should be allowed to work outside of home” 4. Disagree, “Educating boys is more important than educating girls” 5. Agree, “Boys should do as much domestic work as girls” 6. Disagree, “A girl must obey her brother’s opinion even if he’s younger than her”
- *Ever Employed (Endline)* – a dummy indicating the respondent reports having at least one paid job since she graduated from community college
- *Ever Employed (Midline)* – a dummy indicating the respondent reports being currently employed or having worked for pay within the last three months
- *High Aptitude* – a dummy indicating an above median score on the Tawjihhi test for our sample
- *Hours Worked Last Week* – the number of hours that the respondent reports working in the previous week
- *Labor Force Participation* – a dummy indicating the respondent is *Employed* or has looked for a job in the last month
- *Life Ladder* – a measure of how the respondent feels about her life at the moment measured by the eleven point Cantril self-anchoring striving scale with higher scores indicating positivity
- *Life Ladder Future* – how the respondent feels about how her life will be in five years measured by the eleven point Cantril self-anchoring striving scale with higher scores indicating positivity
- *Low Desire to Work Full Time* – a dummy indicating either the respondent doesn’t plan to graduate, is pessimistic about finding a job after graduation, or expects to work only part-time if she found a job

- *MHI 5* – a measure of psychological well-being and the presence of psychological distress in the past month measured by the Mental Health Inventory (MHI-5) of Veit and Ware (1983) with higher scores indicating better mental health. The five questions are:

1. *During the past month, how much of the time were you a happy person?*

(1=All of the time, 2=Most of the Time, 3=Some of the time, 4=A little of the time, 5=None of the time)

2. *How much of the time, during the past month, have you felt calm and peaceful?*

(1=All of the time, 2=Most of the Time, 3=Some of the time, 4=A little of the time, 5=None of the time)

3. *How much of the time, during the past month, have you been a very nervous person?*

(1=All of the time, 2=Most of the Time, 3=Some of the time, 4=A little of the time, 5=None of the time)

4. *How much of the time, during the past month, have you felt down-hearted and blue?*

(1=All of the time, 2=Most of the Time, 3=Some of the time, 4=A little of the time, 5=None of the time)

5. *How much of the time, during the past month, did you feel so down in the dumps that nothing could cheer you up? (1=Always, 2=Very often, 3=Sometimes, 4=Almost Never, 5=Never)*

Questions 1 and 2 are reverse-scored, so that answer one receives score 5, answer two score 4, and so on. Questions 3 through 5 are scored as they appear. This gives a maximum MHI-5 score of 25, and a minimum of 5, with higher scores representing better mental health.

- *Mobility Index* – a proxy for mobility on a scale from zero to six with a higher score indicating more places a girl can travel by herself
- *Months Employed Since Graduation (Midline)* –months that the respondent was employed in her current job
- *Months Employed Since Graduation (Endline)* –months that the respondent was employed in her current or most recent job plus the months employed in the her first job
- *Severely Poor Mental Health* – a dummy indicating the the respondent reports an *MHI5* score of less than 17, which indicates severe depression according to several studies (e.g. Urban Institute, 1999; Yamazaki et al., 2005)
- *Work Income* – the monthly salary reported by the respondent
- *Wealth Index* – a proxy for wealth with a higher index indicating greater wealth created by a principal component analysis on household assets

Appendix III: Correlates of Treatment Take-up

The first two columns of Table 3 examine the correlates of training attendance among those invited to training. Column 1 first examines how the four stratifying variables relate to take-up: we see no significant relationship with geographic location or with ability to travel to the market unaccompanied, suggesting that the choice of trusted locations in many governorates was successful in reducing

geographical and mobility constraints. Attendance is 9.9 percentage points less likely for those who do not believe they are likely to be working full-time after graduation, and 8.6 percentage points less likely for those with tawjihi scores above the median, perhaps reflecting that more academically skilled individuals felt they had less need for such training. There was no difference in training take-up among those who also were assigned the job voucher (which was known to graduates at the time of deciding whether to attend training).

Column 2 of Table 3 then adds several other characteristics of the graduate and their household. We see participation in training is 18.4 percentage points less likely for those graduates who were already married at baseline, even conditional on their expectations of working, empowerment and household wealth levels. Attendance is significantly higher for those taking administrative or financial courses, perhaps because they expect to be in more of a position to need business writing skills than graduates going into nursing or teaching, and higher for those who have email at home.

Columns 3 and 4 of Table 3 examine the correlates of voucher use. We see significantly lower use of the voucher in greater Amman than outside of it, even conditioning on the other stratification variables, and a weakly significant positive effect of having fewer mobility restrictions. Voucher use is not significantly different for those who also were assigned to the training treatment. Being married at baseline has a negative, but not statistically significant impact on using the voucher, while having more female siblings has a significant negative effect of similar magnitude to the relationship to training take-up. In contrast to the impact on training take-up, graduates of administrative or financial courses are less likely to use the voucher than graduates of other specializations.