ORIGINAL ARTICLE

Open Access

CrossMark

The impact of soft skills training on female youth employment: evidence from a randomized experiment in Jordan

Matthew Groh[1], Nandini Krishnan[1], David McKenzie[1,2,3,4*] and Tara Vishwanath[1]

* Correspondence:
dmckenzie@worldbank.org
[1]World Bank, Washington, DC, USA
[2]BREAD, Cambridge, MA, USA
Full list of author information is
available at the end of the article

Abstract

Employers around the world complain that youth lack the soft skills needed for success in the workplace. In response, a number of employment programs have begun to incorporate soft skills training, but to date there has been little evidence as to the effectiveness of such programs. This paper reports on a randomized experiment in Jordan in which female community college graduates were randomly assigned to a soft skills training program. Despite this program being twice as long in length as the average program in the region, and taught by a well-regarded provider, we find soft skills training does not have any significant employment impact in three rounds of follow-up surveys. We elicit expectations of academics and development professionals and reveal that these findings are novel and unexpected.

JEL codes: O12, O15, J08, J16

Keywords: Soft skills, Youth unemployment, Randomized experiment, Expectation elicitation

1 Introduction

Firms around the world 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, how to work in teams, how to act professionally, and even how to properly represent themselves in job interviews in the first place. For example, in a review of 28 studies using employer surveys in different countries, Cunningham and Villaseñor (2014) find that socio-emotional skills are the first priority of 76.5 % of the studies that rank employer skill preferences. Likewise, several surveys of US employers identify a lack of soft skills as the area where young job seekers have the largest deficiency (White, 2013).

As a result, efforts to improve these skills have become an increasing part of youth employment programs in many developing countries, especially in Latin America (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

Springer

Groh *et al. IZA Journal of Labor & Development* (2016) 5:9

Page 2 of 23

early evidence from Ibarraran et al.'s (2012) finding, this increased the formality of work but did not lead to any increase in overall employment in the short run. More recently, a series of pilot programs for adolescent girls have combined life skills training with technical or vocational training in Afghanistan, Haiti, Liberia (Adoho et al. 2014), Nepal (Ahmed et al., 2015), and Rwanda.

Yet despite the growing popularity of such programs, there is little empirical evidence on the effectiveness of teaching soft skills. There is a variety of evidence that shows associations between non-cognitive or soft skills and employment and other life outcomes (e.g., Bowles et al., 2001; Heckman et al., 2006; Heckman and Kautz, 2012; Groh et al., 2015), but it is less clear as to whether such skills can be effectively taught, especially in the types of 2- or 3-week courses used in the typical employment program. Moreover, studies in developing countries have typically combined life or soft skills training with vocational or technical training, so that the effect of the soft skills training alone cannot be measured.

Our paper provides the first experimental evidence on the effectiveness of soft skills training for youth in a developing country context.[1] The Jordan New Opportunities for Women (Jordan NOW) pilot randomly allocated 1347 female community college graduating in 2010 into a training group that received an invitation to attend a 45-h soft skills training course, and a control group which was not.[2] Follow-up surveys then measure employment impacts 6, 14, and 27 months later, allowing us to examine the program impacts at different post-training time horizons. Despite the course being taught by a well-regarded provider and offering twice as many hours of training as the average employability program does in the region, we find that soft skills training did not have any significant impact on any of a range of different employment measures for these young women over any of these time horizons. Examining non-employment outcomes, we find that in the short term, the training does appear to have made the participants more optimistic about the future, and have reduced restrictions on their mobility.

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). However, as Watts (2011) discusses, many findings in social science seem obvious once you know the answer, but the opposite conclusions could also often be easily justified ex post. 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 expectation elicitation exercise. This study and the results from a companion experiment on wage subsidies (Groh et al., forthcoming) were presented in seminars to academics and development professionals and described online on a popular development economics blog; point expectations of the treatment impacts were then elicited before the results were revealed. The results reveal considerable heterogeneity in the expected outcome of these two employment programs, with a general bias towards expecting soft skills training to be more effective than it was in practice, and for there to be a complementarity with wage subsidies which does not exist in practice. In addition to 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 results; Section 5 the expectation elicitation exercise; and Section 6 concludes.

2 Context, sample, and the intervention

Jordan's labor market is similar to others throughout the Middle East in that there is a bulging youth population facing high unemployment rates, particularly among relatively educated youth, with female youth employment rates particularly low. Youth aged 15–29 are the largest demographic group in the country, comprising 30 % of the total population. In 2011, the unemployment rate for 20–24-year-old females was 47.6 % compared to 23.1 % for 20–24-year-old males (Jordanian Department of Statistics, 2011).

2.1 The sample

We work with graduates of Jordan's community colleges. These community colleges offer 2-year programs, similar to community colleges in the USA, and like the USA tend to attract individuals with more limited financial means and/or lower academic scores than youth attending universities. At the request of the Government of Jordan, the World Bank developed the Jordan NOW pilot program to attempt to assist women graduating from these community colleges to find employment.

The pilot worked with the eight public community colleges with the largest female enrolment numbers, together comprising over 85 % of all female public community college enrolment. These 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 ease of exposition, since Amman is the capital of Jordan and two of the four colleges in Central Jordan are located within the city of Amman while the other two are located within a 45-min drive of Amman, we will herein refer to Central Jordan as inside Amman and Northern and Southern Jordan as outside Amman.

Baseline surveys were conducted in July 2010 for all second-year students in these colleges sitting for their final exams, giving data on 1755 female students. In August 2010, the baseline data was 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. The majority of students had specialized 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; or educational programs (10 %), which covers those aiming to be teachers.

2.2 The soft skills training intervention

Interventions that aim to teach soft 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.

Graduates receiving this intervention were invited to free intensive training on interpersonal skills in areas identified by Jordanian employers as essential yet typically missing in recent graduates. The training course was 45 h over a 9-day period (5 h 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

Groh *et al. IZA Journal of Labor & Development* (2016) 5:9

Page 4 of 23

Center (BDC), 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 training programs. Training 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 where BDC has satellite training facilities.

BDC designed the 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 CV and interviewing skills. Sessions were based on active participation and co-operative learning rather than lectures, with games, visual learning experiences, group exercises, and active demonstrations used to teach and illustrate concepts. Appendix 3 provides some more detail of the training schedule and content.

The World Bank and International Youth Foundation (2013) studied 57 organizations providing life skills or soft skills training in the Middle East. The IYF set benchmarks for such programs, among which is that training lasts a minimum of 30 h, finding that the average program in the region lasted only 20 h, with programs focused on employability having fewer hours and those focused more on civic education having more hours. The program we offered therefore offers more than twice as many hours of soft skills training as the average employability program does in the region.[3]

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 a 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 end of high school (Tawjihi) examination score 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 they are usually permitted to travel to the market alone (a measure of empowerment). These variables were chosen as potential predictors of future employment, for which stratification over would improve power (Bruhn and McKenzie, 2009), as well as variables for which there may be potential heterogeneity in treatment impact. The soft skills training intervention discussed in the paper was cross-randomized with a second intervention which provided wage subsidies (discussed in a companion paper, Groh et al., forthcoming). Appendix 1 shows that there was no significant interaction effect between the two treatments; hence, we can just compare all those who were allocated to soft skills training to all those who were not. Within each stratum, 44.4 % of the students were allocated to receive the soft skills training for a total treatment size of 599 and control group of 748.[4]

Groh *et al. IZA Journal of Labor & Development* (2016) 5:9

Page 5 of 23

Students were informed that this was a pilot program, with insufficient funding to enable program provision to everyone, and that a lottery was being used as a way to ensure everyone had an equal chance of being selected.

3.1 Baseline information by treatment status

Table 1 provides summary statistics on the experimental sample by treatment status. As one would expect given computerized randomization, the treatment and control groups look similar on observable characteristics, and a test of joint orthogonality cannot reject the null hypothesis that these characteristics are jointly unrelated to treatment status.

At baseline, the graduates express high levels of desire to work, with 93 % saying they plan to work after they graduate, and 82 % say they think it is very likely or somewhat likely that they will have a job within 6 months of graduating. We define individuals who do not plan to work, do not think they will find a job, or only expect to work part-time as having less desire to work, with 41 % in both groups falling into this category.

The average age of the youth in our study is 21, with only 14 % married at baseline. Only 16 % have any previous work experience, and only 7 % have a mother who currently works. At the time of the baseline survey, which took place only weeks before final examination results were available, only 8 % of students had already found a full-time job for work after graduation. As a result, the majority of our sample consists of young women who have expressed a desire to work, have just graduated, and have not yet found work or worked before.

Table 1 Comparison of means of baseline characteristics by treatment group

	Soft skills	Control	*p* value
Stratifying variables			
In Amman, Salt, or Zarwa	0.44	0.44	
Tawjihi score above median	0.55	0.55	
Low desire to work full-time	0.41	0.41	
Is allowed to travel to the market alone	0.51	0.51	
Other baseline variables			
Age	21.1	21.3	0.214
Married	0.14	0.14	0.850
Mother currently works	0.07	0.07	0.865
Father currently works	0.59	0.56	0.238
Has previously worked	0.17	0.15	0.430
Has a job set up for after graduation	0.09	0.06	0.076
Has taken specialized English training	0.26	0.31	0.068
Household owns car	0.64	0.63	0.771
Household owns computer	0.74	0.71	0.176
Household has internet	0.22	0.27	0.032
Prefers government work to private sector	0.80	0.81	0.582
Sample size	599	748	
Joint orthogonality test			0.190

Note: p values are from regressions which control for randomization strata. As a result, they are not available for the variables which were used for stratification

Groh *et al. IZA Journal of Labor & Development* (2016) 5:9

Page 6 of 23

3.2 Data collection

Data collection consisted of three follow-up surveys, conducted in April 2011, December 2011, and January 2013. This timing allows us to test both whether the soft skills training helps graduates find their first job faster, as well as to address the possibility that in an environment where it takes graduates some time to find a job, it may take more than 1 year to see impacts. All rounds collected data on employment outcomes, and the second follow-up survey also collected information on several well-being measures, including mental health, subjective well-being, and empowerment. The response rates were 92, 96, and 92 % for the first, second, and third follow-ups, respectively, including proxy reports from relatives for 3 and 9 % of the subjects in the second and third rounds. These attrition rates are low in absolute terms and are balanced by treatment status (see Appendix 4), and so we treat attrition as at random in our estimation.

In addition to the survey data, we obtained administrative data from the Social Security Corporation of Jordan. In March 2012, the Social Security Corporation sent us employment data for all graduates who provided their social security numbers in the follow-up surveys, which is 95 % of the sample.

3.3 Take-up of the soft skills training

The soft skills training course was completed by 373 of the 599 graduates (62 %) assigned to treatment. This take-up rate is similar to that of other training courses in developing countries. For example, McKenzie and Woodruff (2014) find an average take-up rate of 65 % for business training programs, with this average including a number of studies which made training a required part of obtaining a loan. Only five students attended a part of the course but did not finish. Participants were asked to give their assessment on a five-point scale of each of seven topics in the course at the end of the course: their rating were extremely positive, with 98 % of participants rating each component as "excellent," the top of the scale, and the remainder typically rating it as good (Business Development Center BDC 2010). Eighty-one percent of women who participated in the training say that they would have paid 150 JD to take the course had they not had the opportunity to take it for free. In qualitative comments, many participants said the course had given them more confidence and positivity, as well as noting their appreciation for learning practical topics not taught in college.

Table 2 examines the correlates of training take-up. Column 1 considers just the stratification variables, while column 2 adds a number of additional controls. We see training attendance is lower for those who have less desire to work full-time, and for individuals who are married. It is also lower for those with higher Tawjihi scores, perhaps reflecting that these higher-skilled graduates believe they have less need for these additional skills. Attendance is higher for those with brothers (who can perhaps oversee them traveling to training), for those with email, and for individuals studying administrative and business subjects.

Unfortunately, neither administrative data from the course itself nor our surveys provide any direct measure of soft skills. Soft skills are difficult to measure in a survey setting, and there are concerns about self-reported Likert scale measures for questions like "I can work with someone who has different opinions from mine". In follow-up work (Groh et al., 2015), we worked with an organizational psychologist to develop

Table 2 Correlates of soft skills training take-up

Marginal effects from probit estimation		
Amman, Salt, or Zarqa	0.067 (0.108)	−0.057 (0.114)
Tawjihi score above median	−0.232** (0.107)	−0.220** (0.109)
No desire to work full-time	−0.265** (0.107)	−0.270** (0.110)
Is allowed to travel to market alone	−0.083 (0.105)	−0.064 (0.107)
Also given wage subsidy	0.158 (0.105)	0.137 (0.107)
Married at baseline		−0.514*** (0.155)
Wealth index		0.008 (0.032)
Number of brothers		0.056** (0.028)
Number of sisters		−0.019 (0.024)
Has email		0.340** (0.165)
In admin/finance program		0.319*** (0.111)
Sample size	599	599

Note: robust standard errors in parentheses
*, **, and *** denote significance at the 10, 5, and 1 % levels, respectively

several measures of soft skills in a different Jordanian youth population and show these soft skills are predictive of future employment outcomes. However, this approach required individuals to come in person to a testing center, where they were tested in both individual and group settings, with the tests taking an hour to complete. As a result, this approach would be costly and likely subject to high attrition in a follow-up survey setting. The consequence is that we can only report the above information that participants expressed a high degree of satisfaction with the course content but are not able to distinguish how much they learned during training.

4 Results

To evaluate the impact of assignment to the soft skills training treatment, we estimate the following equation for graduate i via OLS over the three follow-up periods $t = 1, 2, 3$:

$$
\begin{aligned}
\text{Outcome}_{i,t} = {}& \alpha + \beta_1 \text{Soft}_i + \beta_2 \text{Soft}_i \cdot \text{Time2}_t + \beta_3 \text{Soft}_i \cdot \text{Time3}_t \\
& + \theta_2 \text{Time2}_t + \theta_3 \text{Time3}_t + \sum_{s=1}^{16} \delta_s d_{i,s} + \varepsilon_{i,t}
\end{aligned}
\tag{1}
$$

where Soft_i is a dummy variable taking the value 1 if graduate i was assigned to the soft skills training treatment, Time2 and Time3 are dummy variables for the second and third follow-up surveys, respectively, and the $d_{i,s}$ are the randomization strata dummies (Bruhn and McKenzie, 2009). The coefficient β_1 then measures the intention-to-treat (ITT) effect of soft skills training in the first follow-up period, a time period 6 to 7 months post training. β_2 and β_3 then enable us to examine whether this effect changes over longer time horizons. The total effect of the treatment in follow-ups 2 and 3 are then given by $\beta_1 + \beta_2$ and $\beta_1 + \beta_3$, respectively, and we test whether these effects are 0. The standard errors are clustered at the individual level. We focus on ITT impacts which give the average effect of being offered training. We choose not to estimate the treatment effect on the treated because of the possibility that the offer of training may affect job search behavior even if the treatments are not taken up. Nonetheless, *late*

estimates for the training effect can be easily approximated by dividing the ITT estimates by the take-up rate (0.62).

4.1 Average impacts on labor force participation and employment

Table 3 reports the results of estimating Eq. (1) for different employment outcomes, defined in detail in Appendix 2. Column 1 examines labor force participation, defined as either working or actively looking for work. At the time of the first follow-up survey, labor force participation was relatively high at 78 % in the control group. Soft skills training has no significant impact on this rate. Labor force participation then falls by over 20 percentage points in the two subsequent surveys with the training having no impact on reducing this drop. Note that the labor force participation variable captures job search activity, which is high to begin with, and falls over time. Soft skills training does not result in more job search.

Columns 2 through 4 examine employment. Column 2 looks at whether individuals are currently employed or have worked for cash in the last month. 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 the soft skills training has no significant impact on any of these measures, neither in the short run nor in either of the two longer-term follow-up surveys. The point estimates are close to 0, especially for formal work, with a 95 % confidence interval for the ITT effect of training on employment 6 months later of (−0.043, +0.067) and for formal employment registered with social security of (−0.029, +0.037). As Hoenig and Heisey (2001) argue, these confidence intervals, rather than post-experiment power calculations, are the appropriate guide for interpreting statistically insignificant results and show the range of possible effect sizes not refuted by our data enables us to rule out large impacts of the treatment.

The remaining columns of Table 3 consider dimensions of employment at the intensive margin hours worked and income earned. Columns 6 and 7 are unconditional estimates, which code the outcome as 0 for those not working, while column 8 looks descriptively at income conditional on work. We find no significant impact of the soft skills training treatment on either hours worked or income earned.

The point estimate for the short-term increase in income is 2.3 JD per month, or approximately US$3. The training cost is $250 per individual invited, so these gains (which were not statistically significant) would have to last for 82 months to recoup the cost of training. It is important to note that these results do not necessarily imply that soft skills are not important determinants of income. Since we do not observe soft skills directly, we cannot say whether the lack of effect is due to the training not being very effective at increasing soft skills or to there being little return to these skills. In either case, the results here imply that the training does not appear cost effective.

4.2 Heterogeneity in employment impact

The lack of overall effectiveness of the soft skills training raises the question of whether training may be more effective if targeted at specific subgroups. To do this, we pool together follow-up rounds 2 and 3 to increase power and ease in interpretation of coefficients and estimate the following equation:

Table 3 Treatment impacts on employment

	Labor force participation	Employed	Formal contract	Employed and registered for social security (survey)	Employed and registered for social security (admin data)	Hours worked last week	Work income (not conditional on working)	Work income (conditional on working)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Assigned to soft skills training	0.026 (0.023)	0.012 (0.028)	0.005 (0.023)	0.004 (0.017)	−0.007 (0.015)	0.748 (1.062)	2.262 (4.715)	−0.019 (4.198)
Assigned to soft skills* follow-up 2 (December 2011)	−0.001 (0.031)	−0.015 (0.031)	0.006 (0.026)	−0.001 (0.020)	−0.003 (0.015)	−0.279 (1.249)	−2.971 (5.253)	−2.908 (7.109)
Assigned to soft skills* follow-up 3 (January 2013)	−0.028 (0.033)	0.021 (0.034)	0.017 (0.028)	0.013 (0.025)	n.a.	1.160(1.324)	8.177 (6.635)	6.435 (9.720)
Follow-up 2 (December 2011)	−0.256*** (0.021)	−0.155*** (0.021)	−0.068*** (0.018)	0.040*** (0.014)	0.066*** (0.010)	−3.839*** (0.827)	−17.296*** (3.637)	17.652*** (4.992)
Follow-up 3 (January 2013)	−0.214*** (0.023)	−0.124*** (0.023)	−0.027 (0.019)	0.070*** (0.017)	n.a.	−3.996*** (0.889)	−9.125** (4.430)	54.463*** (6.912)
Control mean in first follow-up	0.783	0.345	0.157	0.087	0.074	12.263	52.185	155.199
Sample size	3759	3759	3720	3603	2564	3721	3721	1049
p value: no effect at follow-up 2	0.382	0.909	0.531	0.883	0.577	0.663	0.876	0.679
p value: no effect at follow-up 3	0.937	0.195	0.312	0.439	0.623	0.078	0.078	0.498
p value: equality of follow-up 2 and 3 effects	0.419	0.160	0.639	0.480	0.813	0.199	0.034	0.320

Note: Huber-White standard errors in parentheses, clustered at the student level. All regressions also control for stratification dummies

n.a. denotes Administrative Data on Social Security Registration not available for the third follow-up period

*, **, and *** indicate significance at the 10, 5, and 1 % levels, respectively

Groh et al. IZA Journal of Labor & Development (2016) 5:9

Page 9 of 23

Groh *et al. IZA Journal of Labor & Development* (2016) 5:9

Page 10 of 23

$$\text{Employed}_{i,t} = \alpha + \beta_1 \text{Soft}_i + \beta_2 \text{Soft}_i \cdot \text{Time2or3}_t + \gamma_1 \text{Soft}_i \cdot X_i$$
$$+ \gamma_2 \text{Soft}_i \cdot X_i \cdot \text{Time2or3}_t + \theta_2 \text{Time2}_t + \theta_3 \text{Time3}_t + \delta_2 \text{Time2}_t \cdot X_i$$
$$+ \delta_3 \text{Time3}_t \cdot X_i + \sum_{s=1}^{16} \delta_s d_{i,s} + \mu X_i + \varepsilon_{i,t}$$

$$(2)$$

where X is the baseline characteristic of interest for heterogeneity testing and Time2or3 indicates the interaction with being in the second or third follow-up round. We examine heterogeneity with respect to the four variables used to stratify the randomization (and which can therefore be considered as pre-specified dimensions of heterogeneity to examine), as well as with respect to baseline marital status since this was a strong predictor of training take-up and of subsequent employment status.

Table 4 reports the results of estimating Eq. (2) to examine how the impact of soft skills training on employment varies with individual characteristics. The foot of the table then tests whether there is a significant impact of the training in the pooled round 2 and round 3 surveys for the case when the interaction variable is 0, and for when it is 1. We see that in no case can we reject the null hypothesis of no impact of treatment. That is, it is not the case that the soft skills training works particularly well for those in Amman, or those with higher ability, or those who are not married to begin with, etc. This is despite these variables having predictive power for employment outcomes. For example, we see that in the first follow-up only 22 % of those married are employed, compared to 36.5 % of those who were unmarried at baseline, but that we cannot reject that training has the same effect for both groups and that this effect is 0.

4.3 Impact on well-being, empowerment, and attitudes

As noted above, the training was rated extremely highly by those who attended it, with the participants saying that it had given them more confidence and positivity. This raises the possibility that it may have had impacts in other dimensions of life, even if it did not affect employment. We investigate this in Table 5.

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 5.2 on this scale, and being assigned to soft skills training lowers this by 0.23 rungs, significant at the 10 % level. Column 2 examines which step on the Cantril ladder graduates believe they will be on in 5 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 8.0 for the control group three rungs higher than they assess their current position to be. The training intervention leads to additional optimism: graduates assigned to this group think they will be 0.23 steps higher still, significant at the 5 % level, and equivalent to a 0.1 standard deviation increase. So, training participants are less satisfied with their current status but more optimistic about the future. Such a result could arise if the training causes them to think about a better possible future life for themselves than they had previously imagined, making their current status seem less desirable.

Groh *et al. IZA Journal of Labor & Development* (2016) 5:9

Page 11 of 23

Table 4 Heterogeneity in treatment effects

Dependent variable: employed

	Interaction is between treatment and:				
	Amman	High aptitude	Low desire to work	Allowed to travel	Married
Assigned to soft skills	0.023 (0.038)	0.034 (0.042)	0.006 (0.037)	0.057 (0.041)	0.020 (0.031)
Assigned to soft skills*follow-up 2 or 3	0.017 (0.039)	−0.002 (0.044)	−0.005 (0.040)	−0.051 (0.042)	−0.007 (0.033)
Assigned to soft skills*interaction	−0.021 (0.057)	−0.040 (0.057)	0.017 (0.058)	−0.087 (0.057)	−0.060 (0.075)
Assigned to soft skills*follow-up 2 or 3*interaction	−0.035 (0.060)	0.008 (0.060)	0.018 (0.060)	0.105* (0.060)	0.069 (0.074)
Follow-up 2	−0.221*** (0.026)	−0.172*** (0.031)	−0.160*** (0.028)	−0.131*** (0.028)	−0.159*** (0.023)
Follow-up 3	−0.161*** (0.028)	−0.120*** (0.033)	−0.106*** (0.030)	−0.090*** (0.030)	−0.114*** (0.024)
Follow-up 2*interaction	0.134*** (0.042)	0.015 (0.042)	−0.009 (0.041)	−0.063 (0.041)	−0.029 (0.050)
Follow-up 3*interaction	0.103** (0.044)	0.007 (0.044)	−0.027 (0.044)	−0.053 (0.044)	−0.019 (0.056)
Control mean in first follow-up when interaction equals 0	0.310	0.319	0.378	0.307	0.365
Control mean in first follow-up when interaction equals 1	0.391	0.366	0.296	0.382	0.219
Sample size	3759	3759	3759	3759	3759
p value: no effect when interaction = 0 at follow-ups 2 and 3	0.122	0.276	0.983	0.840	0.574
p value: no effect for interaction = 1 at follow-ups 2 and 3	0.628	0.988	0.257	0.436	0.620
p value: equality of effect by interaction at follow-ups 2 and 3	0.189	0.445	0.406	0.670	0.864

Notes: Each column shows results from interacting treatment effect with the variable listed at the top of the column. All regressions also control for randomization strata dummies, while the last column also controls for marriage at baseline. Huber-White standard errors in parentheses, clustered at the student level
*, **, and *** indicate significance at the 10, 5, and 1 % levels, respectively

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

Table 5 Impacts on well-being, empowerment, attitudes, and marriage

	Life ladder	Life ladder future	MHI-5	Severely poor mental health	Mobility index	Empowerment index	Married at 2nd follow-up
Assigned to soft skills	−0.230*	0.233**	0.307	−0.033	0.381***	0.025	−0.019
	(0.133)	(0.105)	(0.209)	(0.022)	(0.096)	(0.051)	(0.027)
Control mean	5.207	8.043	19.193	0.204	4.966	4.822	0.323
Sample size	1249	1249	1249	1249	1249	1249	1249

Note: Robust standard errors in parentheses. All regressions include stratification dummies. Severely poor mental health denotes the MHI-5 index is below 17
*, **, and *** denote significance at the 10, 5, and 1 % levels, respectively

Groh *et al. IZA Journal of Labor & Development* (2016) 5:9

Page 12 of 23

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 has 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 small increase in the MHI-5 and reduction in the likelihood of major depression, but neither is statistically significant.

Columns 5 and 6 of Table 5 examine whether there were changes in two measures of empowerment: the ability to go to different places themselves, and attitudes towards women in society. We find a highly significant increase in the ability to travel to different places: the mean in the control group is being able to travel to 5.0 out of 6 places by herself, with the training increasing this by 0.38 places. In contrast, there is no significant change in attitudes towards women in society. Finally, we note that while at baseline (graduation from community college) only 13.7 % of the sample was married, this had increased to 31.6 % by the second follow-up 18 months later. Column 7 shows that training had no impact on the likelihood of marriage.

All together, there are seven different outcomes in Table 5, of which three are significant at the 10 % level. If we correct for multiple hypothesis testing using a Bonferroni or a step-down procedure, only the impact on mobility remains significant. Therefore, it appears that overall the impacts on other life outcomes are also relatively modest.

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. In a classic article, Lazarsfeld (1949) sets out six findings from a study of American soldiers (e.g., Southern soldiers were better able to stand the climate in the hot South Sea Islands than Northern soldiers), and then for each, a probable reader reaction (e.g., of course, Southerners are more accustomed to hot weather). He then reveals that each of these statements was false (e.g., Southerners had no greater ability than Northerners to adjust to a tropical climate). Watts (2011) refers to this as the "myth of common sense" and means that ex post, either a large positive impact, a zero impact, or a negative impact may all be easily explained away as obvious.

To understand whether this study merely confirms conventional wisdom or generates new and unexpected results, we undertook a novel audience expectation elicitation exercise. During the first four presentations of this research (two in academic departments, and two in international organizations),[5] 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 and of the cross-randomized wage subsidy experiment. We then distributed paper surveys throughout the audience and asked them to provide their point estimates of the first and second follow-up ITT impacts for each treatment group: soft skills training, wage subsidy only, and both combined. To complement the audience expectation elicitation exercises, we posted a description of the study on the World Bank's Development Impact blog[6] with a link to an online survey to capture reader expectations.[7] 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

Groh *et al. IZA Journal of Labor & Development* (2016) 5:9

Page 13 of 23

wider audience but offers less of a chance to explain the intervention, and has low and self-selected response rates.[8] In both cases, we explicitly asked the audience to give the ITT impact but gave them the take-up rate for the soft skills training so that the results would reflect beliefs about treatment impact and not treatment take-up.

Figure 1 provides histograms of the resulting expectations about the soft skills training impact, while Table 6 summarizes the resulting expectations and compares them to their expectations of the wage subsidy effects. We see first that there is tremendous heterogeneity in expectations about the likely effect of the soft skills training: the standard deviation is 10 percentage points for the first follow-up impact and 11 for the second. Second, on average, people are biased towards thinking soft skills training will be more effective than it was in practice. The mean expected impact of 9 to 10 percentage points is three to five times as large as the estimated ITTs. Only 63 % of the expectations lie within a 95 confidence interval of the treatment effect.

We also see that expectations were inaccurate for the wage subsidy program, with respondents dramatically underestimating the short-term effect of the subsidy and overestimating the longer-term effect. Only 4 % 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. Audiences believed the combination of the wage subsidy voucher and soft skills training was likely to have larger impact than either treatment alone, but this complementarity did not occur in practice. None of the respondents had all six of their expectations lie within the 95 % confidence intervals of the actual impacts. Furthermore, there was considerable heterogeneity among respondents in their relative rankings of the interventions: at the first (second) follow-up, 59 (47) % thought the wage subsidy would have a larger impact than the soft skills training, 11 (14) % that the impact would be the same, and 30 (39) % that the wage subsidy would have less impact than the soft skills training.

While the sample size is relatively small to examine in detail the heterogeneity of expectations across groups, we can examine the extent to which expectations differ among the five different audiences (four seminars and the blog). There are three results of interest
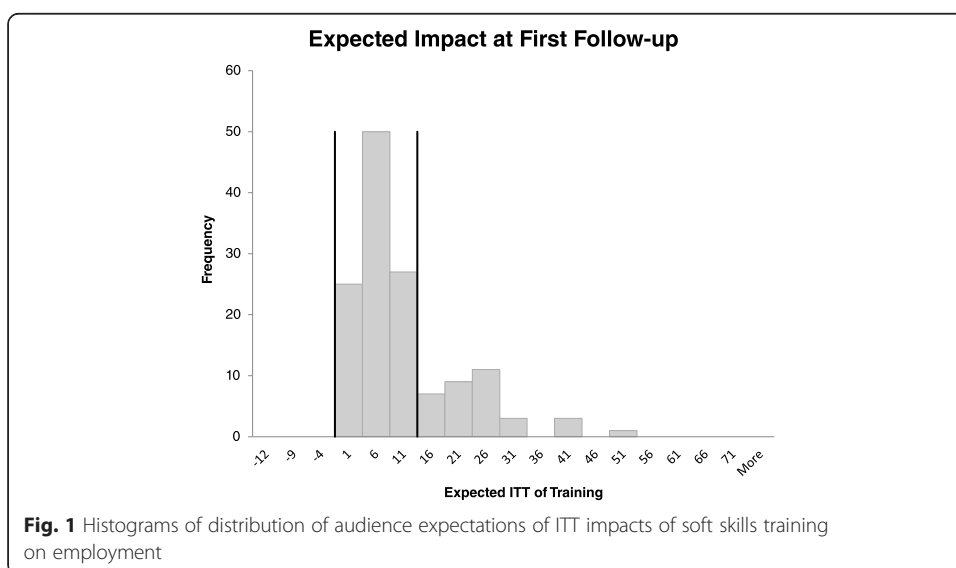

Fig. 1 Histograms of distribution of audience expectations of ITT impacts of soft skills training on employment

Groh *et al. IZA Journal of Labor & Development* (2016) 5:9

Page 14 of 23

Table 6 Expected impact of the interventions from seminar participants and online respondents

	Number	Actual treatment effect	$\mu_{expected}$	$\sigma_{expected}$	median$_{expected}$	% of expectations within 95 % CI
Expected impact at first follow-up						
Wage subsidies alone	136	40	12	12	8	4
Soft skills training alone	136	3	9	10	5	63
Wage subsidy and soft skills	135	40	17	14	11	9
Expected impact at second follow-up						
Wage subsidies alone	137	3	10	12	5	60
Soft skills training alone	139	2	10	11	5	63
Wage subsidy and soft skills	135	2	16	15	10	38

here. First, the results are not driven by any one audience. Second, overall, the most ac-curate responses were from blog readers, although these responses were not that different from participants in a World Bank seminar. Third, the least accurate responses were from the Paris School of Economics audience, who were most optimistic about the soft skills training having longer-term effects. We cannot reject that expectations of the soft skills training impact were the same among the other four groups ($p = 0.19$ short-run, $p = 0.52$ long-run). We speculate that this might reflect a European view that is more positive about active labor-market policies generating employment than is the belief in the USA, which could be interesting to test in the future such exercises.[9]

Finally, a simplified version of the expectations was also elicited from nine Jordanian policymakers in a presentation of the results made in Jordan, who were just asked ex-pectations in five bin ranges (<0, 0, 0 to 5 %, 5 % to 10 %, more than 10 %). The policy-makers' responses also show considerable heterogeneity in responses, and the majority of policymakers underestimated the short-term impact of the wage subsidy voucher.

Overall, these results therefore show that there is considerable uncertainty about what the likely impact of soft skills training is likely to be but that people are on aver-age over-optimistic about its impacts and over-optimistic about the possibility of com-plementarities between policies.

6 Conclusions

Soft skills training programs have become an increasingly common type of youth em-ployment policy in developing countries, but to date, there has been little evidence as to their impact. Our randomized experiment finds that even though our soft skills training program was provided by a well-regarded provider, was of twice the length as the regional average, and had very positive post-training participant assessments, it did not have any significant impact on employment outcomes of young women over either the short or the medium term. These findings are in contrast to the expected impact of such programs from academic audiences and policymakers. The expectation elicitation exercise conducted here 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.

While these results are disappointing, they are consistent with findings from recent overviews of vocational training programs (Blattman and Ralston, 2015) and business training programs (McKenzie and Woodruff, 2014) which have noted that these other forms of training programs have also often struggled to show significant employment impacts. This other literature suggests two possible reasons for the lack of effectiveness of soft skills training here. The first is that relatively short courses may only result in only small changes in the skills that they aim to teach. Measuring soft skills well is difficult, and we were not able to observe the extent to which the training actually improved soft skills for the participants. Devoting more attention to how to measure changes in soft skills therefore seems an important area for future work, with Groh et al. (2015) providing one example of how to do this on a different sample of Jordanian youth. It would also be useful to examine the effectiveness of such programs for young men, in addition to the young women studied here. The second possible reason for a lack of effect is that the diagnosis that a lack of skills is the key constraint to employment may be incorrect. The more binding constraint may be more general constraints that prevent productive firms from growing and employing more workers, in which case policy efforts may need to be directed more towards helping firms than helping workers.

Endnotes

[1] A recent working paper by Honorati (2015) provides a second example that helps confirm our results. She tests a 2-week program in Kenya and finds no significant impact on employment outcomes.

[2] The program also cross-randomized a wage subsidy component which is covered in a companion paper (Groh et al., forthcoming). Appendix I shows this program had no interaction effect with the soft skills training.

[3] Our course is also similar in length to the 40-h life skills training in Nepal (Ahmed et al., 2015) and the 2-week training in Kenya (Honorati, 2015).

[4] One individual in each of the treatment and control groups had his/her gender mis-identified in the data and was dropped from the study.

[5] 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.

[6] http://blogs.worldbank.org/impactevaluations/are-our-blog-readers-better-predictors -of-impact-results-than-seminar-audiences-evaluating-programs

[7] 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.

[8] 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, 2014).

[9]A breakdown of results by these five groups can be found on the Development Impact blog http://blogs.worldbank.org/impactevaluations/the-results-are-in-did-our-blog-readers-do-better-than-our-seminar-audiences-at-predicting-the-impac [accessed February 6, 2016].

Appendix 1

6.1 Cross-randomized intervention and no complementarities

The soft skills training intervention discussed in the paper was cross-randomized with a second intervention which provided wage subsidies (discussed in a companion paper, Groh et al., forthcoming). 22.2 % of the students were allocated to receive the wage subsidy voucher only, 22.2 % allocated to receive the soft skills training only, 22.2 % allocated to receive both, and 33.3 % allocated to the pure control group. This resulted in 300 in each treatment group and 449 in the pure control group, with two of the groups of 300 dropping to 299 after two individuals whose gender was misreported were removed. Appendix Table 7 shows that the soft skills training had no significant interaction effect with the voucher treatment. We therefore pool together the wage subsidy alone and control group (and refer to this as the control group in the main body of the paper) and pool together the soft skills training only and wage subsidy voucher plus soft skills group (and refer to this as the soft skills group in the main body of the paper). Since the pure control group was larger in size than the other three groups, we reweight the data by the inverse of the probability of being assigned to that group.

Table 7 Pooling of treatments and no complementarity

	Employed	Monthly income
Assigned to voucher only	0.395*** (0.035)	64.453*** (5.802)
Assigned to soft skills only	0.029 (0.030)	5.227 (4.779)
Assigned to both voucher and soft skills	0.402*** (0.036)	65.593*** (5.911)
Voucher only*follow-up 2	−0.360*** (0.039)	−56.869*** (6.799)
Voucher only*follow-up 3	−0.379*** (0.043)	−57.665*** (8.443)
Soft Skills only*follow-up 2	−0.013 (0.034)	−3.604 (5.643)
Soft skills only*follow-up 3	0.005 (0.039)	3.855 (7.625)
Both voucher and soft skills*follow-up 2	−0.387*** (0.039)	−60.827*** (6.626)
Both voucher and soft skills*follow-up 3	−0.353*** (0.043)	−46.807*** (8.793)
Follow-up 2	0.032 (0.021)	12.340*** (3.520)
Follow-up 3	0.073*** (0.025)	20.900*** (4.625)
Control mean in first follow-up	0.178	0.178
Sample size	3759	3721
p value for testing no effect of soft skills training	0.689	0.484
p value for equality of voucher only and both	0.542	0.306

Notes: Robust standard errors in parentheses, clustered at the student level. All regressions control for stratification dummies
*, **, and *** denote significance at 10, 5, and 1 % levels, respectively

Appendix 2

6.2 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 0 to 6 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"
- *High aptitude*—a dummy indicating an above median score on the Tawjihi 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 work after they 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 as follows:

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 0 to 6 with a higher score indicating more places a girl can travel by herself
- *Severely poor mental health*—a dummy indicating 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 3

6.3 Further details of soft skills training

Soft skills training by BDC was intended to provide job seekers with interpersonal and professional skills that employers look for when hiring new graduates. The aim of this training is to equip young graduates with particular cross-cutting professional skills, which are lacking in typical community college curricula. Such skills include some elements of resume preparation and interview taking; professional conduct; business communication; teamwork; and presentation skills.

All candidates have been trained on the following topics: team building, communication skills, customer service, positive thinking, business writing, presentation skills, interviewing skills, and CV writing. The training approach used is based on adult learning principles, which means that it is interactive, relevant and practical. Training methods used included Ice-breakers and energizers, power point presentations, discussion and sharing information, group case studies experiences, brain storming and team building activities, and role models and videos. A detailed training outline is as follows:

- Communication and presentation skills

General objective:

This program aims to introduce participants to modern concepts of the communication process as well as develop their skills in communicating and dealing with others in a positive manner.

Topics:

- The communication process: concept, importance, elements
- Stages of communication and obstacles
- Communication skills
- Listening skills
- Oral skills
- Writing reports

• Team work/team building

General objective:

This program aims to provide participants with knowledge for team building and the importance of teamwork in business performance in addition to the skills necessary to adopt the teamwork spirit which will reflect positively on their performance and attitudes in their professional lives.

Topics:

- Teams: concept, importance, composition
- Stages team creation/building
- Types and roles of team members
- Characteristics of the effective team
- Principles and characteristics of team work
- The evaluation and measurement of team performance

• Positive thinking

General objective:

This program aims to provide participants with the needed skills to switch negative to positive and to be able to utilize that in business situations.

Topics:

- How to identify negative thinking and behavior habits
- Methods, models and techniques for switching negative to positive
- Thinking systems for success—planning the positive future
- Visualization techniques to transform business situations

• CV writing and interviewing skills

General objective:

By the end of the training, university students/graduates will have written their Curriculum Vitae and will gain an insight on how to perform well in a job interview.

Topics for CV writing

1. Introduction
   (a) What is a "CV"?
   (b) Why do we write a "CV"?
   (c) Different types of "CV"
   (d) CV main rules
2. The cover letter
   (a) Cover letter hints
3. Before writing the CV
   (a) Visual layout
   (b) Fonts and sizes
   (c) Results orientated CV
   (d) How to test your CV
4. Writing the CV
   (a) How to write a CV
   (b) The order
   (c) Heading
   (d) Heading examples
   (e) Good heading example
   (f) Objective
   (g) Professional experience
   (h) Training/part-time jobs
   (i) Education
   (j) Certifications and awards
   (k) Successful achievements
   (l) Languages
   (m) Skills
5. Updating your CV
   (a) What updates?
   (b) Visual layout updates

Topics for interviewing skills:

- The ins and outs of interviewing

1. Interviewing—what is it?
2. Types of interviews
3. Pre-interview preparation
4. Arriving
5. Greeting and introduction
6. Body language: what signals are you sending?
7. Types of interview questions
8. General interview strategies
9. Dressing for success
   - Business writing

General objective:

The main objective of the session is to introduce the concept of business writing to the participants and the need for systematic writing. The participants will be introduced to different types of correspondences and writing methods.

Topics:

- Objects of business writing
- Fundamentals of business writing
- Types of correspondences
- Means of delivery
- Correspondences layout
- Template
- Date and reference
- Attention
- Subject
- Body
- Signature
- Common phrases
- Fax
- E-mail
- Memo
- Commercial offers

• Excellence in service

General objective:

This program aims to provide participants with the concept of customer service in the aim of raising the level and quality of services, as well as develop the participants' skills in dealing with customers, clients, and the public.

Topics:

- A general introduction of the "service"
- The importance of the service in labor market
- The client/customer and its importance
- What is a high quality service?
- How to gain and maintain a clients' satisfaction
- The mechanism of communicating with a unsatisfied customer
- Types of customers
- Handling complaints
- The art of dealing with the public (clients and customers)
- The importance of listening skills
- Skills needed for call centers

Appendix 4
6.4 Survey response rates by treatment status
Appendix Table 8 shows that the survey response rates are balanced across the treatment and control groups.

Table 8 Response rates by treatment status

	Soft skills	Control	*p* value
First follow-up	0.918	0.918	0.987
Second follow-up	0.955	0.956	0.933
Third follow-up	0.913	0.920	0.663
Social security	0.962	0.944	0.131

Author details
[1]World Bank, Washington, DC, USA. [2]BREAD, Cambridge, MA, USA. [3]CEPR, London, UK. [4]IZA, Bonn, Germany.

References
Adoho F, Chakravarty S, Korkoyah Jr. D, Lundberg M, Tasneem A (2014) The impact of an adolescent girls program: the EPAG Project in Liberia. World Bank Policy Research Working Paper No. 6832.
Ahmed A, Chakravarty S, Lundberg M, Nikolov P (2015) The role of training programs for youth employment in Nepal: Impact Evaluation Report on the Employment Fund, AGI Results Brief., http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2015/02/09/000477144_20150209083940/Rendered/PDF/941570BRI00PUB00AGI00results0series.pdf
Blattman C, Ralston L (2015) Generating employment in poor and fragile states: evidence from labor market and entrepreneurship programs. Columbia University, Mimeo
Bowles S, Gintis H, Osborne M (2001) Incentive-enhancing preferences: personality, behavior and earnings. Am Econ Rev 91(2):155–158
Bruhn M, McKenzie D (2009) In pursuit of balance: randomization in practice in development field experiments. Am Econ J: Appl Econ 1(4):200–232
Business Development Center (BDC) (2010) "Jordan Now Training Program: Training report 25th September – 2nd November, 2010", Report delivered to the World Bank.
Cantril H (1965) The pattern of human concerns. Rutgers University Press, New Brunswick, NJ
Cunningham W, Villaseñor P (2014) Employer voices, employer demands, and the implications for public skills development policy, World Bank Policy Research Working Paper no. 6853.
Das J, Do Q-T, Friedman J, McKenzie D (2008) Mental health patterns and consequences: results from survey data in five developing countries. World Bank Econ Rev 23(1):31–55
Entra 21 (2009) Final Report on the entra21 Program Phase I 2001–2007. International Youth Foundation.
Groh M, McKenzie D, Vishwanath T (2015) Reducing information asymmetries in the youth labor market of Jordan with psychometrics and skill based tests. World Bank Econ Rev Papers Proc 29(suppl 1):S106–S117
Groh M, Krishnan N, McKenzie D, Vishwanath T (forthcoming) Do wage subsidies provide a stepping stone to employment for recent college graduates? Evidence from a randomized experiment in Jordan, Review of Economics and Statistics. https://drive.google.com/file/d/0B9C9RwWKZrUNOUZvWnFmOFY1Wk0/view?usp=sharing.
Heckman J, Kautz T (2012) Hard evidence on soft skills, IZA Discussion Paper no. 6580
Heckman J, Stixrud J, Urzua S (2006) The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. J Labor Econ 24(3):411–482
Hoenig J, Heisey D (2001) The abuse of power: the pervasive fallacy of power calculations for data analysis. Am Stat 55(1):1–6
Honorati M (2015) The impact of private sector internship and training on urban youth in Kenya, World Bank Policy Research Working Paper no. 7404.
Ibarraran P, Ripani L, Tapoada B, Villa JM, Garcia B (2012) Life skills, employability and training for disadvantaged youth: evidence from a randomized evaluation design, IZA Discussion Paper no. 6617.
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&year1=2011&t_no=18 [accessed 24 April, 2011].

Kahneman D, Deaton A (2010) High income improves evaluation of life, but not emotional well-being. PNAS 107(38): 16489–16493

Lazarsfeld P (1949) The american soldier—an expository review. Public Opin Q 13(3):377–404

McKenzie D, Özler B (2014) Quantifying some of the impacts of economics blogs. Econ Dev Cult Chang 62(3):567–597

McKenzie D, Woodruff C (2014) What are we learning from business training and entrepreneurship evaluations around the developing world? World Bank Res Obs 29(1):48–82

Scheiber N (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. Accessed March 2, 2016.

Veit C, Ware J (1983) The structure of psychological distress and well-being in general populations. J Consult Clin Psychol 51:730–742

Watts D (2011) Everything is obvious (*once you know the answer): how common sense fails Us. Crown Business, New York

White M (2013) The real reason new college grads can't get hired, Time, November 10, http://business.time.com/2013/11/10/the-real-reason-new-college-grads-cant-get-hired/

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 and International Youth Foundation (2013) Preparing youth for success: an analysis of life skills training in the MENA region. 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.