Online Supplement for: Jobseekers' Beliefs about Comparative Advantage and (Mis)Directed Search

Andrea Kiss ⓡ Robert Garlick ⓡ Kate Orkin ⓡ Lukas Hensel

August 10, 2024

Contents

A Summary Statistics for Skill Assessment Results

This appendix reports additional summary statistics for skill assessment results in the big and tight experiment, providing more details behind the discussion in Section 2 of the paper.

Table A.1: Joint Distribution of Skill Quintiles - Tight Experiment

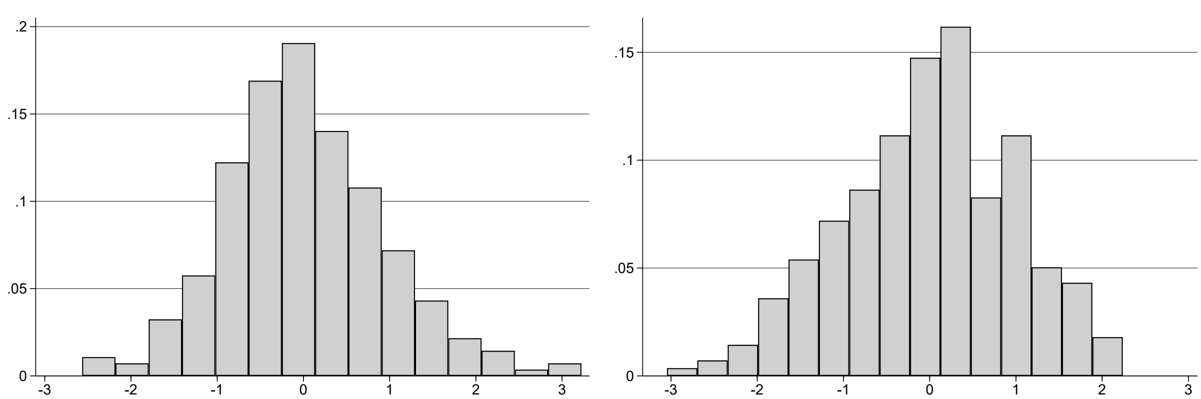
Numeracy quintile	Communication quintile				
	Bottom	Lower middle	Middle	Upper middle	Top
Bottom	0.00%	7.19%	6.47%	4.68%	4.68%
Lower middle	11.51%	0.00%	7.19%	7.91%	9.35%
Middle	3.96%	2.52%	0.00%	4.68%	2.88%
Upper middle	2.52%	4.32%	3.96%	0.00%	7.19%
Top	0.72%	2.16%	0.72%	5.40%	0.00%

Notes: **Table A.1 shows that there considerable variation in quintiles across numeracy and communication skills in the tight experiment.** The sample is restricted to the 278 jobseekers with a unique CA in skills. This forces entries on the main diagonal to be zero.

Figure A.1: Distribution of Standardized Assessment Scores - Tight Experiment

Panel A: Numeracy

Panel B: Communication



Notes: **Figure A.1 shows the distribution of standardized assessment scores in the tight experiment.** Panel A displays the fraction of standardized numeracy scores in bins. Panel B displays the same for communication scores.

Table A.2: Correlation between Skill Quintiles - Tight Experiment

	Numeracy (1)	Communication (2)	Concept formation (3)
Panel A: Restricted sample			
Numeracy	1.000	-0.004	0.284
Communication		1.000	0.216
Concept formation			1.000
Panel B: Full sample			
Numeracy	1.000	0.306	0.375
Communication		1.000	0.326
Concept formation			1.000

Notes: **Table A.2 shows that correlation coefficients between the skills in the tight experiment are mostly moderate-sized and positive.** Panel A shows results for the sample of 278 jobseekers with a unique CA between numeracy and communication. Therefore, the correlation between those two skills is zero in this subsample. Panel B shows results for the full sample of 373 jobseekers.

Table A.3: Correlation Between Skill Terciles - Big Experiment

	Concept formation (1)	Communication (2)	Numeracy (3)	Grit (4)	Planning (5)	Focus (6)
Concept formation	1.000	0.298	0.435	0.098	0.214	0.189
Communication		1.000	0.331	0.095	0.213	0.173
Numeracy			1.000	0.139	0.266	0.159
Grit				1.000	0.090	0.047
Planning					1.000	0.173
Focus						1.000

Notes: **Table A.3 shows that correlation coefficients between the skills in the big experiment are mostly moderate-sized and positive.**

B Firms' Valuation and Observation of Skills in Hiring

B.1 Firms' Valuation of Skills

This section of the supplement provides further information about firms' valuation of the skills we measure and how much firms can observe these skills at the time of hiring, providing more details behind the discussion in Section 2 of the paper.

First, the communication, concept formation, and numeracy assessments have been used to screen jobseekers by our partner, the Harambee Youth Employment Accelerator. By 2016, Harambee had been contracted by firms in South Africa to screen roughly 160,000 prospective workers using these assessments. This suggests a revealed willingness of firms to pay to learn the results of these assessments. However, this does not mean that assessment results are the only information firms use in their hiring decisions. We do not assume that firms use the information at their disposal optimally, and thus, we do not claim that these tests are the best predictors of jobseekers' productivity.

Second, we used an incentivized choice experiment to show that firms vary in their valuation of communication and numeracy skills and value both highly relative to some forms of education. For this data collection, we recruited 67 firms soon after the big experiment by going door-to-door in areas of Johannesburg where most of the jobseekers in the big experiment lived. 81% of firms are in the retail or hospitality sectors, where many jobseekers in both experiments applied for jobs. Recruited firms have a mean size of 15 workers, half of whom are in entry-level roles, and planned to hire an average of 4 new entry-level workers in the next year.

Importantly for our current argument, we measured the preferences of these firms over the six skills used in the big experiment, relative to each other and relative to additional education. Each firm was asked to rank multiple jobseeker profiles with different levels of skills and with or without a one-year post-secondary diploma, all with completed secondary school. To incentivize the choices, we used firms' rankings to match them with jobseekers with specific skill profiles from Harambee's database, in a similar spirit to Kessler et al. (2019).

Table B.1 shows the average ranking of numeracy, communication and education over the 67 firms. There are six different possible rankings of these three elements, each shown in a row. The shares of firms in these bins are shown in column 4. Column 6 collapses these shares based on the most important skill. In this column, we see that 57% of firms prefer a candidate with top-tercile numeracy skills, 34% prefer a candidate with top-tercile communication skills and only 9% firms prefer a candidate with a relatively better educational achievement, i.e. a candidate with a diploma but with only middle-tercile communication and numeracy skills.

Table B.1: Firms' Preference Ranking Over Communication Skills, Numeracy Skills, and Formal Education

	Top (1)	Middle (2)	Bottom (3)	Share (%) (4)	Most Important Skill (5)	Share (%) (6)
1	Num	Comm	Educ	52.24	Numeracy	56.72
2	Num	Educ	Comm	4.48		
3	Comm	Num	Educ	28.36	Communication	34.33
4	Comm	Educ	Num	5.97		
5	Educ	Num	Comm	1.49	Education	8.96
6	Educ	Comm	Num	7.46		

B.2 Observability of Skills and Firms' Valuation of Applicant Skill Match

As part of the tight experiment, we conducted a measurement exercise to show that firms partially observe assessed skills and value applicants whose skill profile matches their job requirements. During the job search workshop, we asked jobseekers to choose between applying for a real communication- or numeracy-heavy job at a firm that hires for a range of entry-level roles, including call center and data capture jobs. Jobseekers prepared a CV and a cover letter during the workshop, both designed for general use rather than tailored to these specific jobs. Two members of the firm's HR team evaluated every applicant for both jobs based on their CV and generic cover letter. Evaluators did not know which applicant applied for which job and were not shown applicants' skill assessment results. We received data on the evaluators' assessment of each applicant's communication skills, numeracy skills, suitability for each type of job, and whether they were recommended for an interview for each type of job.

This measurement exercise shows that the firm's HR team can get some information about jobseekers' skill levels from their application materials. Table B.2, panel A, column 5 shows that the HR team's assessments of skills are positively but not perfectly correlated with our measures of skill: they assigned a 0.22 standard deviation higher score to the skill dimension in which the jobseeker had a CA on our assessments ($p = 0.01$). The HR team's applicant ratings are also correlated with our measures of skill: they rate the applicant as 0.18 standard deviations more suitable for the job aligned with that applicant's CA (panel B, row 1, column 6). HR managers were also 9 percentage points more likely to recommend interviewing the candidate for the job aligned with that applicant's CA (panel B, row 2, column 4). This is a 21% increase relative to a 43% interview recommendation rate in the non-aligned job. These patterns show that skills are at least partly observable to the firm even when jobseekers could not tailor their resumes or cover letters to the specific role. Observability might be greater in natural job search where jobseekers can tailor their applications.

Table B.2: Employer Evaluation of Job Applicants Based on Skills

	Mean			Difference			Obs.
	Aligned	Non-aligned	SD	Δ	Δ/SD	$p(\Delta = 0)$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Skill levels							
Skill (1-5)	2.93	2.79	0.66	0.15	0.22	0.01	277
Panel B: Job-related evaluation							
Overall score (1-5)	3.00	2.84	0.86	0.16	0.18	0.05	277
Interview invitation (dummy)	0.43	0.34	0.49	0.09	0.18	0.07	277

Notes: **Table B.2 shows that the HR team of an employer can observe jobseekers' skills and evaluates applicants more highly if their assessed CA in skills matches the job's requirements.** One pair of the job choice task advertisements was from a firm that hires for a range of entry-level roles. We submitted the jobseekers' cover letter and resume to the firm based on which two members of the firm's HR team evaluated every applicant for both jobs. Evaluators rated the jobseekers' skill levels (Panel A) as well as their general suitability for the job (on a scale from 1 to 5) and whether they would invite the candidate for an interview (Panel B). We show the mean outcomes across evaluators for the skill/ job that is aligned with the jobseekers' CA in col. 1; and the outcomes for the misaligned skill / job in col. 2. The pooled standard deviation of the measures in col. 1 and 2 are in col. 3. Col. 4 shows the difference between cols. 1 and 2. Col. 5 shows this difference in terms of standard deviations. Col. 6 shows the p-value associated with a test of equality across cols. 1 and 2. Col. 7 shows the number of observations.

The evidence that jobseekers have different skills, that firms value these skills but differ in which skill they value more, and that firms can at least partly observe skills suggests that redirecting jobseekers' search towards jobs that match their CA in skills has the potential to improve their labor market outcomes.

C Beliefs About Wages and Job Offer Probabilities

The model predicts a three-stage reaction to treatment: beliefs about skill CA $\rightarrow$ search direction $\rightarrow$ search outcomes. We show evidence consistent with each of these stages in the paper. The model also suggests that treatment will shift expected returns to specific search activities. We evaluate this idea in this appendix.

We find that treatment effects on beliefs about search outcomes are broadly consistent with the model. But we view them as less important than the results on beliefs about skills, search direction, and search outcomes we include in the main paper. The questions about expected search outcomes rely on complex forecasts by jobseekers, as discussed in recent reviews (Delavande, 2023; Mueller & Spinnewijn, 2023). For example, we ask jobseekers about their expected search duration. This requires jobseekers to forecast both the expected number of offers and attributes of those offers (wages, hours, travel costs, working conditions, etc.), as the attributes determine whether they will accept offers. In contrast, our tight experiment is optimized to measure search direction in multiple different ways, including direct measures of behavior. Moreover, the survey questions about skill beliefs and actual labor market outcomes are simpler to answer.

We use only data from the tight experiment in this section. In the tight experiment, we collected beliefs about search outcomes on the same day as treatment. This means that treatment effects on beliefs reflect only information acquired during the workshops. In the big experiment, the endline survey occurred months after treatment. This means that any treatment effects on beliefs about search outcomes would reflect both the direct effect of treatment and any indirect effects arising from treatment-induced changes in search actions and their outcomes. The indirect causal channel is interesting but not a key part of the argument that we make in this paper.

Treatment effects on beliefs about returns to search direction: The model suggests that changes in search direction will be accompanied by changes in beliefs about the relative returns to searching for different types of jobs, i.e., beliefs about the returns to skill-directed job search. To evaluate this prediction, we conduct two exercises.

First, we surveyed jobseekers about their expected outcomes from applying to each type of job and estimate treatment effects on these measures. We asked for their expected number of job offers in the next 30 days, time to employment, and wage when employed. We asked these questions after treatment and just after asking their planned number of applications in the next 30 days, and we ask the expected offers and time to employment questions conditional on their planned number of applications. We asked all questions separately about all jobs, communication-heavy jobs, and numeracy-heavy jobs. We define the expected return to skill-directed job search as the expected outcome

Table C.1: Treatment Effects on Beliefs About Returns to Skill-Directed Job Search - Tight Experiment

	Index		ΔE [offers per app] (w)		- (Δ [sear. dur.]) (w)		ΔE [wage] (w)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	0.095	0.260*	0.043**	0.053**	-0.007	-0.079	0.073	23.981**
	(0.100)	(0.130)	(0.018)	(0.026)	(0.157)	(0.185)	(6.605)	(9.383)
Treatment × Aligned CA belief (bl)		-0.341*		-0.023		0.114		-45.738***
		(0.180)		(0.034)		(0.223)		(13.960)
Aligned CA belief (bl)		0.689***		0.107***		0.432**		35.845***
		(0.140)		(0.030)		(0.197)		(12.069)
Treatment effect:		-0.081		0.030		0.035		-21.757**
Aligned CA belief (bl)		(0.124)		(0.022)		(0.191)		(9.164)
Control mean	-0.000	-0.000	0.027	0.027	0.244	0.244	6.007	6.007
Observations	278	278	273	273	272	272	278	278

Notes: **Table C.1 shows that treatment improves jobseekers' their expected outcomes from searching for jobs that match their assessed CA relative to jobs that do not match their assessed CA.** All outcomes are defined as the expected outcome for jobs aligned with CA minus the expected outcome for jobs not aligned with CA, so positive values indicate higher expected returns from searching for aligned jobs. "CA" stands for comparative advantage in skills and "bl" stands for baseline. Outcomes are an inverse covariance-weighted average of outcomes in cols 3–8 (cols. 1–2) and differences, as defined above, in winsorized expected offers per application (cols. 3–4), winsorized expected search duration in months (cols. 5–6), and winsorized expected weekly wages in 2021 USD in purchasing power parity terms (cols. 7–8). Even-numbered columns show heterogeneity by whether individuals have aligned CA beliefs at baseline. Control variables are listed in footnote 18. Standard errors clustered at the treatment-day level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

for jobs aligned with CA minus the expected outcome for jobs not aligned with CA. For example, for a jobseeker with numeracy CA, the expected number of offers for numeracy-heavy jobs minus the expected number of offers for communication-heavy jobs. We also construct an inverse covariance-weighted average of the three measures.

Treatment effects on these survey measures of expected returns to skill-directed job search are mostly consistent with our model. Treated jobseekers expect 0.043 more job offers from skill-directed job search, equal to 165% of the control group mean (Table C.1, column 3, row 1, $p = 0.023$). This effect, like most in the paper, is driven by jobseekers with baseline misaligned CA beliefs (column 4). Treated jobseekers with baseline misaligned CA beliefs expect weekly earnings 24 USD higher from skill-directed job search, equal to roughly four times the control group mean (column 8, row 1, $p < 0.001$). Treatment has negligible effects on the expected length of time it will take to get a job from skill-directed job search (columns 5-6). We return to this result on page 11.

Treatment effects on an index combining these survey measures of expected returns to skill-directed job search are consistent with our model, although slightly imprecisely estimated. Treatment increases the expected return to skill-directed search by 0.095 standard deviations for the average jobseeker (column 1, row 1, $p = 0.35$). This result is driven by the same heterogeneity that we see elsewhere in the paper: jobseekers with misaligned baseline CA beliefs increase their expected returns by 0.26 standard deviations (column 2, row 1, $p = 0.054$) while jobseekers with aligned baseline CA beliefs do not increase their expected returns (column 2, row 4).

Reassuringly, we see a "sensible" relationship between baseline CA beliefs and expected returns to skill-directed job search in the control group. Jobseekers whose baseline CA belief matches their assessed CA have a 0.69 standard deviation higher expected return to skill-directed job search (column 2, row 3). The relationship is also positive for all three components of the index (columns 4, 6, 8).

Second, we survey jobseekers about their expected outcomes from applying to specific jobs during the job choice task. For each job in 5 of the 11 job pairs, we ask jobseekers about the probability of getting an offer if they applied, the expected starting wage, and the general desirability of the job. We estimate treatment effects on these measures with a prespecified regression of the belief measure on treatment, a dummy for job alignment with jobseeker CA, their interaction, job fixed effects, and prespecified controls. The relevant coefficient is on the interaction term. This captures the treatment effect on jobseekers' beliefs about the attributes of the job aligned with their skill CA, relative to the job not aligned with their CA.

Treatment effects on these measures of expectations are similar to effects on the first types of expectations, discussed above, but are less precisely estimated. Within each pair of jobs, treatment increases the expected offer probability and wage for the job aligned with the jobseeker's CA (Table C.2, columns 2–3). But the effects are small – roughly 2% of the control group mean – and are not statistically significant at conventional levels. So we do not view this as strong evidence supporting the model. These results might be less precise than the results using the survey measures of expected returns to skill-directed search discussed above because the questions asked during the job choice task only ask about five specific pairs of jobs, rather than allowing jobseekers to implicitly average over many skill-directed job application choices.

Relationship between search direction and expected return to skill-directed job search: Beliefs about job attributes predict jobseekers' choices in the job choice task, consistent with a role for belief-based job search. For 5 of the 11 pairs of jobs, we asked jobseekers the probability that they would get an offer if they applied to each job and their expected salary if offered a job. We regress the job choice on the job offer probability times the expected monthly wage using a logit regression model, following Wiswall & Zafar (2015). This is not an experimental analysis because we regress post-treatment choices on post-treatment beliefs. Column 1 of Table C.3 shows that a 100 USD increase in the expected weekly wage scaled by the offer probability is associated with a 9.2 pp increase in the probability of choosing that job ($p = 0.001$). This relationship is robust to

Table C.2: Treatment Effects on Beliefs About Jobs in the Job Choice Task - Tight Experiment

	Desirability (sd)	Expected earnings (w)	Job offer probability
	(1)	(2)	(3)
Treatment	-0.038	-3.240	-0.022
	(0.032)	(5.921)	(0.022)
Treatment× Aligned skill req.	0.008	3.873	0.011
	(0.029)	(4.006)	(0.013)
Aligned skill req.	0.030	-3.297	0.017
	(0.022)	(2.724)	(0.011)
Control mean	-0.000	193.588	0.544
Observations	2770	2770	2770

Notes: **Table C.2 shows that treatment very weakly increased jobseekers' expectations about jobs in the job choice task that were aligned with their assessed CA.** Beliefs were elicited for both jobs in each of 5 job pairs for each jobseeker. Columns indicate different outcome variables: desirability of jobs measured on a 0–10 Likert scale and standardized with respect to the control group (col. 1), expected weekly earnings winsorized at the 99^{th} percentile (col. 2), and expected job offer probability (col. 3). Analysis is at the job $\times$ jobseeker level and includes prespecified control variables listed in footnote 18, dummies for CA in numeracy and communication, fixed effects for the randomized order in which job pairs were shown, and job fixed effects. All monetary values are measured in 2021 USD in purchasing power parity terms. Standard errors clustered at the treatment-day level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

adding both jobseeker and job pair fixed effects (columns 2–4). The slope of the belief-choice relationship is somewhat different in the treatment and control groups, although the differences are not statistically significant. This suggests that treatment effects on these belief measures account for part but not all of the treatment effects on job choices.

Treatment effects on beliefs about search outcomes: The preceding analysis focuses on jobseekers' beliefs about the returns to searching for jobs requiring specific skills. We also construct measures of jobseekers' expected outcomes from searching for any type of job. We estimate treatment effects on these beliefs using the specification in equation (4).

Treatment has a positive but imprecisely estimated effect on expected wages, driven by jobseekers with baseline misaligned CA beliefs. Columns 3–4 of Table C.4 show that treatment increases expected weekly wage by 9.6 USD for the average treated jobseeker and 22.1 USD for the average treated jobseeker with misaligned baseline CA belief. Both effects are relatively imprecisely estimated (standard errors 7.7 and 14.9 respectively), perhaps reflecting the difficult of forecasting wage offers for respondents with limited work experience. Effects on jobseekers' reservation wages, beliefs about the minimum and maximum wages they might earn if employed, and an index combining all of these wage beliefs follow the same qualitative pattern (columns 1–2 and 5–10).[1] These positive

[1] In standard job search models, the reservation wage is both a decision rule and a feature of the wage distribution. This is not inconsistent with our interpretation of reservation wages as another proxy for wage expectations.

effects on wage beliefs in the tight experiment are consistent with the positive effects on actual wages in the big experiment, although we do not compare the magnitudes to evaluate forecast accuracy because the estimates come from two different experiments.

Treatment has a positive effect on the expected probability of formal employment, again driven by jobseekers with baseline misaligned CA beliefs. Columns 9–12 of Table C.5 show that treatment increases the probability of employment in 1–3 months by 3–4 percentage points for the average treated jobseeker and 8–9 percentage points for the average treated jobseeker with misaligned baseline CA belief. Effects on other, less direct proxies for employment probability – callbacks and offers per application and search duration – are closer to zero (columns 3–8). These results might differ because the callback, offer, and search duration questions all explicitly condition on the jobseeker's planned number of applications, while the probability of employment questions are asked later in the survey and do not include this explicit conditioning.[2] The explicit conditioning might mean jobseekers put more mental weight on the role of search effort relative to search direction when answering these questions, but this is a speculative suggestion that future work could better evaluate. These results from the tight experiment are qualitatively consistent with the big experiment's positive effect on employment with a written contract. However, the magnitudes are substantially different, perhaps in part because the control group's expectations are much higher than realized outcomes.

[2]For example, we ask "How many job applications do you plan to submit in the next 30 days?" and then "If you submit X job applications in the next 30 days, how many months starting from today do you think it will take you to find a formal job, with an employment contract where you are paid a regular salary?

Table C.3: Association between Job Choices, Offer Probabilities, and Expected Wages - Tight Experiment

	Marginal effects on choice of numeracy job (logit estimate)							
	Control group				Control and treatment groups			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$E[Wage^{num}] - E[Wage^{com}]$ (w)	0.092***	0.097***	0.119***	0.127***	0.092**	0.099**	0.120***	0.124***
	(0.028)	(0.030)	(0.038)	(0.040)	(0.038)	(0.039)	(0.038)	(0.039)
Treatment $\times$ $E[Wage^{num}] - E[Wage^{com}]$ (w)					0.023	0.022	-0.076	-0.078
					(0.052)	(0.051)	(0.055)	(0.055)
Observations	695	695	600	600	1385	1385	1195	1195
Job pair fixed effects	No	Yes	No	Yes	No	Yes	No	Yes
Individual fixed effects	No	No	Yes	Yes	No	No	Yes	Yes

Notes: **Table C.3 shows that expected earnings predict job choice in the tight experiment.** All estimates are average marginal effects from logit regressions of indicators for choosing the numeracy-heavy job on the difference in expected returns between the communication-heavy and numeracy-heavy jobs. The difference in expected returns is the expected weekly wage (in 100s of USD in 2021 purchasing power parity terms) times the expected probability of a job offer for the numeracy job, minus the equivalent quantity for the communication job, winsorized at the 1% and 99% levels. Columns 1–4 use only the control group and columns 5–8 use both the treatment and control groups. Columns 3–4 and 7–8 include jobseeker fixed effects. Sample sizes drop from columns 1–2 and 5–6 to columns 3–4 and 7–8 because some jobseekers choose the numeracy-heavy job in all pairs, so their fixed effects are perfect predictors. Standard errors shown in parentheses are heteroskedasticity-robust in all columns, clustering by treatment date where appropriate, and using a 500-repetition bootstrap in columns with fixed effects, where analytical heteroskedasticity-robust adjustments are not feasible. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.4: Treatment Effects on Beliefs about Wages - Tight Experiment

	Index		Wage expectations (w)		Minimum expected wage (w)		Maximum expected wage (w)		Reservation wage (w)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treatment	0.099	0.319*	9.566	22.119	8.934	19.180*	0.650	3.115	1.318	3.805
	(0.104)	(0.179)	(7.740)	(14.852)	(6.090)	(10.607)	(11.984)	(24.259)	(3.628)	(5.948)
Treatment × Aligned CA belief (bl)		-0.461*		-26.041		-21.065		-6.425		-5.366
		(0.258)		(21.034)		(17.531)		(38.832)		(6.578)
Aligned CA belief (bl)		0.247		10.018		4.604		20.878		5.287
		(0.208)		(17.725)		(15.279)		(29.176)		(4.039)
Control mean	-0.000	-0.000	212.045	212.045	131.392	131.392	324.076	324.076	109.567	109.567
Observations	278	278	278	278	278	278	278	278	277	277

Notes: **Table C.4 shows that treatment effects on beliefs about wages in the tight experiment are positive but not consistently statistically significant.** "CA" stands for comparative advantage in skills and "bl" stands for baseline. Columns indicate different outcome variables: an inverse covariance-weighted average of the other outcomes (cols. 1–2), the expected wage (cols. 3–4), the minimum expected wage (cols. 5–6), the maximum expected wage (cols. 7–8), and the reservation wage (cols. 9–10). Wage expectations are answered to the questions "What is the (lowest / highest possible) monthly take-home salary you think you would earn?" in a permanent, formal job with six months tenure. Even-numbered columns show heterogeneity by whether jobseekers had aligned CA beliefs at baseline. Control variables are listed in footnote 18. All outcomes are winsorized at the 99^{th} percentile. All monetary values are measured in 2021 USD in purchasing power parity terms per week. Standard errors clustered at the treatment-day level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.5: Treatment Effects on Beliefs about Offers and Search Duration - Tight Experiment

	Index		Callbacks / apps (w)		Offers / apps (w)		Month to job (w)		p(employed in 1 months)		p(employed in 3 months)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treatment	0.095	0.264	-0.001	0.039	-0.006	0.016	0.011	0.228	3.493	8.674**	4.235*	8.431**
	(0.122)	(0.194)	(0.023)	(0.036)	(0.021)	(0.031)	(0.247)	(0.392)	(2.238)	(3.748)	(2.315)	(4.042)
Treatment × Aligned CA belief (bl)		-0.360		-0.086		-0.049		-0.444		-10.821*		-8.678*
		(0.247)		(0.057)		(0.042)		(0.470)		(6.003)		(5.126)
Aligned CA belief (bl)		0.234		0.061		0.049		0.069		4.789		2.647
		(0.228)		(0.043)		(0.036)		(0.386)		(4.416)		(4.186)
Control mean	0.000	0.000	0.391	0.391	0.257	0.257	2.466	2.466	54.094	54.094	68.230	68.230
Observations	278	278	276	276	274	274	275	275	278	278	278	278

Notes: **Table C.5 shows that treatment effects on beliefs about job offers and search duration in the tight experiment are mostly positive but not consistently statistically significant.** "CA" stands for comparative advantage in skills and "bl" stands for baseline. Columns indicate different outcome variables: an inverse covariance-weighted average of the other outcomes (cols. 1–2), expected callbacks per application (cols. 3–4), expected offers per application (cols. 5–6), expected months to find a full-time job (cols. 7-8), and expected probability of being employed one month (cols. 9–10) and three months after baseline (cols. 11–12). Expected months to employment is multiplied by minus one before being included in the average so that higher values correspond to better search outcomes, as for the other outcomes. Winsorized variables (w) are winsorized at the 99^{th} percentile. Even-numbered columns show heterogeneity by whether jobseekers had aligned CA beliefs at baseline. Control variables are listed in footnote 18. Standard errors clustered at the treatment-day level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D Additional Mechanism Tables

This appendix shows results related to the "additional mechanisms" discussed in Section 6 and general equilibrium considerations discussed in Section 7.1 of the paper.

Table D.1 shows that treatment effects on self-esteem and education investments in the big experiment are very small and not statistically significant. Table D.2 shows that treatment effects on labor market outcomes are driven by jobseekers who did not attach their skill report with any job applications. This suggests that labor market effects are unlikely to be driven by firms learning about assessment results that are shared by jobseekers.

Table D.3 shows that the treatment did not induce congestion effects, i.e., it did not change the share of applications directed to numeracy-heavy relative to communication-heavy jobs. This suggests that interventions providing information about skill CA can help different types of jobseekers apply to different types of jobs, rather than concentrating applications toward one type of job.

Table D.4 shows treatment effects on the three different willingness-to-pay (WTP) measures in the tight experiment and Figure D.1 displays the distribution of willingness-to-pay values by treatment status.. Treatment lowers jobseekers' average WTP for a numeracy workbook but does not affect average WTP for a communication workbook. Treatment also does not affect jobseekers' WTP for information about jobs' skill demands. These results suggest a limited role for skill investment or information about jobs' skill demand in explaining the labor market effects.

Willingness-to-pay measurement details: As part of the job search workshop we elicited jobseekers' willingness-to-pay for three products: a document that revealed the expert-assessed skill requirements of jobs included in the job choice task (delivered after 30 days to avoid it influencing their search behavior), printed self-study materials for to improve numeracy skills, and printed self-study materials for to improve written communication skills. The self-study materials were taken from the "Mind the Gap!" series which targets grade 12 students and was developed by the Department of Basic Education in South Africa.

We elicited willingness-to-pay using multiple price lists with six items and announced that one randomly selected choice would be implemented in practice. The choice was always between receiving the product and receiving a changing monetary amount in airtime for their mobile phone (the "price"). For the skill document revealing the skill requirements of jobs the prices were: 0, 25, 50, 100, 150, and 200 Rand. For the self-study materials, we use the following prices: 0, 15, 30, 50, 75, and 100 Rand. We randomized whether jobseekers were asked about the highest or lowest price first. We forced a unique switching point by ending the elicitation for an item at the first question they chose the item (for those starting with the highest price) or the monetary amount (for those starting with the lowest price) and assumed that they would choose consistently on the omitted questions. We then randomly chose one question among the 18 questions, implemented their choice in practice, and announced this at the end of the session.

To ensure that jobseekers were familiar with the procedure, they completed a practice round where we elicit the jobseekers' willingness-to-pay for a bar of soap. If their choices were inconsistent, the enumerator explained the elicitation again and asked for the reason for the inconsistency to ensure that they understood the process.

To construct monetary willingness-to-pay measures, we assign jobseekers the average of the price they switched and the previous price. For example, if a jobseeker preferred the item at price zero but the money at price 25 Rand we assign them a willingness-to-pay of 12.5 Rand. For jobseekers who switched immediately, we assign -10 Rand (for those with negative willingness-to-pay) and 110 or 210 Rand (for those whose willingness-to-pay exceeds the maximum price). Given that there is substantially more mass on the upper end of the willingness-to-pay range, this conservative assumption is likely to lead us to underestimate the true willingness to pay.

Table D.1: Treatment Effects on Additional Mechanisms - Big Experiment

	Self-esteem				Education investment		
	SMS (z) (1)	SMS above med. (2)	Endline (z) (3)	Endline above med. (4)	Any (5)	Apprenticeship (6)	Formal (7)
Treatment	0.003 (0.008)	0.014 (0.014)	0.007 (0.021)	0.014 (0.015)	0.011 (0.011)	0.006 (0.005)	0.009 (0.011)
Control mean	-0.000	0.483	0.000	0.471	0.224	0.036	0.185
Observations	3334	3334	4206	4206	4205	4205	4205

Notes: **Table D.1 shows that treatment does not affect average self-esteem or education investment in the big experiment.** Cols. 1–2 show effects on self-esteem in the SMS survey 2–3 days after the workshops. Cols. 3–4 show effects on self-esteem in the endline survey 3.5 months after the workshops. The SMS and phone surveys use respectively one and five items from the Rosenberg (1965) scale, both answered on five-point Likert scales. Cols. 1 and 3 use standardized measures and columns 2 and 4 use dummies for above-median values. Cols. 5, 6, and 7 show effects on dummies for enrolling in respectively any education, an apprenticeship, and a formal degree/diploma. Control variables are listed in footnote 25. Standard errors clustered at the treatment-day level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D.2: Heterogeneous Treatment Effects on Labor Market Outcomes by Skill Report Attachment to Applications - Big Experiment

	Worked in last 7d	Earnings (w)	Cond. earnings (w)
	(1)	(2)	(3)
Treatment	0.014	8.572***	25.961**
	(0.016)	(3.020)	(9.850)
Treatment × Attached report w. application	-0.013	-5.108	-13.035
	(0.024)	(4.229)	(13.444)
Control mean	0.309	25.424	85.826
Observations	3991.000	3983.000	1215.000

Notes: **Table D.2 shows that treatment effects on labor market outcomes are driven by jobseekers that do not use their skill reports in any job applications in the big experiment.** These results should be interpreted with caution because the right-hand side of the regression includes the interaction between treatment and a post-treatment outcome, "Attached report w. application." The interaction term is included in the regression but "Attached report w. application" is omitted because no control individual received a report. Outcomes are the same as in Table 5: a dummy indicating any work for pay (col. 1), unconditional earnings (col. 2), and earnings conditional on working (col. 3), all in the seven days before the endline survey. Winsorized variables (w) are winsorized at the 99^{th} percentile. Control variables are listed in footnote 25. All monetary values are measured in 2021 USD purchasing power parity terms. Standard errors clustered at the treatment-day level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D.3: Treatment Effects on Concentration of Job Search - Tight Experiment

	Degree of job search concentration		
	Individual level measure		Job-pair level measure
	(1)	(2)	(3)
Treatment	-0.007	0.009	-0.029**
	(0.017)	(0.019)	(0.013)
Treatment × Aligned CA belief (bl)		-0.032	
		(0.028)	
Aligned CA belief (bl)		0.050**	
		(0.024)	
Treatment effect:		-0.023	
Aligned CA belief (bl)		(0.024)	
Control mean	0.181	0.181	0.077
Observations	278	278	22

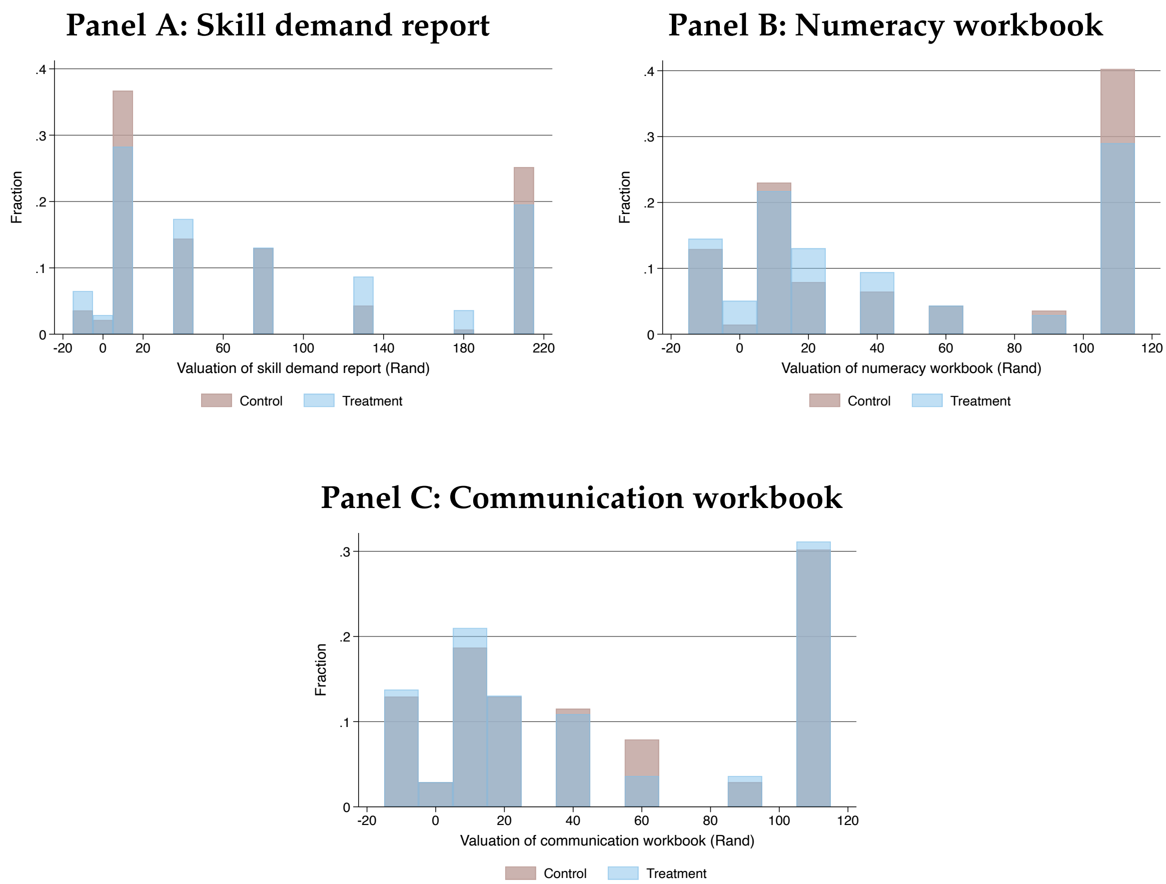
Table D.3 shows that treatment weakly decreases the concentration of jobseekers' applications in the job choice task. Cols. 1 and 2 show effects on the concentration of job choices in the job choice task using one observation per jobseeker. This concentration measure is the absolute deviation of the fraction of chosen numeracy jobs from 0.5, averaged across job pairs at the jobseeker level. Col. 3 shows the effect on the concentration of job choices in the job choice task using one observation per job pair. This concentration measure is constructed in two steps. First, we calculate the fraction of jobseekers choosing the numeracy job in each job-pair × treatment group combination. Second, we calculate the absolute deviation of this measure from 0.5. For both concentration measures, higher numbers indicate more concentrated job choices. We do not estimate heterogeneous treatment effects by baseline aligned CA belief for the job-pair level analysis in col. 3 because the outcome of interest is averaged across jobseekers. Controls used in cols. 1 and 2 are listed in footnote 18. Standard errors are clustered at the treatment-day level in cols. 1 and 2 and are heteroskedasticity-robust in col. 3. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D.4: Treatment Effects on Willingness-to-pay - Tight Experiment

	Info on skill requirements						Numeracy materials						Communication materials					
	Pooled		Num. CA		Comm. CA		Pooled		Num. CA		Comm. CA		Pooled		Num. CA		Comm. CA	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Treatment	-0.719	-6.851	-12.823	-14.035	1.708	-8.179	-13.410***	-15.309**	-14.527*	-20.583**	-10.046**	-3.217	-0.822	-0.247	5.252	-2.571	-3.821	3.862
	(5.766)	(11.255)	(13.609)	(15.888)	(8.396)	(17.726)	(3.539)	(6.775)	(8.049)	(8.944)	(3.898)	(11.201)	(3.790)	(6.747)	(9.785)	(12.162)	(5.882)	(11.719)
Treatment $\times$ Aligned CA belief (bl)		12.287		8.804		14.077		5.019		36.732**		-9.825		-0.847		45.821**		-11.987
		(18.773)		(38.367)		(22.913)		(11.915)		(15.806)		(16.400)		(10.474)		(21.394)		(14.833)
Aligned CA belief (bl)		0.174		-12.737		7.989		-14.744		-38.444***		-4.144		-3.753		-43.991***		7.831
		(16.202)		(28.007)		(16.687)		(8.948)		(11.399)		(11.074)		(9.163)		(11.637)		(10.853)
Treatment effect: Aligned CA belief (bl)		5.436		-5.231		5.898		-10.290		16.149		-13.042*		-1.094		43.251**		-8.125
		(10.847)		(33.252)		(11.314)		(7.268)		(14.781)		(7.247)		(6.130)		(16.512)		(7.682)
Control mean	78.867	78.867	79.337	79.337	78.611	78.611	54.964	54.964	55.408	55.408	54.722	54.722	48.327	48.327	49.184	49.184	47.861	47.861
Observations	277	277	105	105	172	172	277	277	105	105	172	172	277	277	105	105	172	172

Notes: **Table D.4 shows treatment effects on willingness-to-pay (WTP) for different products relevant to job search in the tight experiment.** "CA" stands for comparative advantage in skills and "bl" stands for baseline. Cols 1–6 show effects on the WTP for a document with the expert-assessed skill requirements for the 11 job pairs in the job choice task. Cols. 7–12 show effects on WTP for a numeracy training resource. Cols. 13–18 show effects on WTP for a communication training resource. Cols. 1–2, 7–8, and 13–14 show results for the full sample. Cols. 3–4, 9–10, and 15–16 show results for jobseekers with a CA in numeracy. Cols. 5–6, 11–12, and 17–18 show results for jobseekers with a CA in communication. All currency values are in 2022 South African Rands, which was the currency used in the actual WTP measurement exercise and is consistent with the supplemental WTP appendix posted at `https://bit.ly/3E34oYH`. Multiplying these values by 0.140 converts them into the same currency units as the rest of the paper, 2021 USD in purchasing power parity terms. Control variables are listed in footnote 18. Standard errors clustered at the treatment-day level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure D.1: Distribution of Willingness-to-pay Measures - Tight Experiment



Notes: **Figure D.1 shows the distribution of elicited willingness-to-pay for three goods.** Panel A displays the distribution of willingness-to-pay for a report revealing the skill demand of jobs in the jobs choice task. Panel B displays the distribution of willingness-to-pay for a numeracy workbook. Panel C displays the distribution of willingness-to-pay for a communication workbook. All values are in 2022 South African Rands to illustrate the decisions as made by jobseekers. All other currency values in the paper are reported in 2021 USD purchasing power parity at the exchange rate 1 Rand = 0.140 USD PPP.

E Gender

This appendix reports descriptive statistics and treatment effects separately for women and men. Tables E.1 and E.2 display gender differences in baseline beliefs about skills in respectively the tight and big experiments. Gender differences in beliefs are small and are generally not statistically significant after adjusting for gender differences in demographics, education, and assessment results. Tables E.3 and E.4 display treatment effects on skill beliefs by gender in respectively the tight and big experiments. The treatment effects on skill beliefs do not differ by gender. The lack of gender differences in treatment effects on beliefs about skills makes gender differences in treatment effects on downstream outcomes (e.g. search direction, earnings) unlikely, so we do not report gender-disaggregated treatment effects on these outcomes.

Table E.1: Gender Differences in Beliefs about Skills - Tight Experiment

	Female	Male	Δ	p($\Delta = 0$)	$\Delta(adjusted)$	p($\Delta(adjusted) = 0$)	N
Aligned CA belief	0.50	0.46	-0.04	0.53	-0.01	0.83	278
Fraction aligned beliefs	0.23	0.21	-0.02	0.46	-0.05	0.15	278
Fraction overcofident beliefs	0.62	0.59	-0.03	0.58	0.04	0.11	278
Fraction underercofident beliefs	0.15	0.20	0.05	0.12	0.01	0.79	278

Notes: **Table E.1 shows that gender differences in baseline beliefs about skills in the tight experiment are small.** Adjusted differences control for prespecified covariates described in footnote 18, except the baseline value of the outcome variable. "CA" stands for comparative advantage in skills. P-values are estimated from regressions that use heteroskedasticity-robust standard errors.

Table E.2: Gender Differences in Beliefs about Skills - Big Experiment

	Female	Male	Δ	p($\Delta = 0$)	$\Delta(adjusted)$	p($\Delta(adjusted) = 0$)	N
Aligned CA belief	0.19	0.23	0.03	0.01	0.03	0.02	4312
Fraction aligned beliefs	0.35	0.42	0.07	0.00	0.01	0.19	4378
Fraction overconfident beliefs	0.53	0.45	-0.08	0.00	0.00	0.73	4378
Fraction underconfident beliefs	0.11	0.13	0.01	0.06	-0.01	0.03	4378

Notes: **Table E.2 shows that gender differences in baseline beliefs about skills in the big experiment are small**. Adjusted differences control for prespecified covariates described in footnote 25, except the baseline value of the outcome variable. "CA" stands for comparative advantage in skills. P-values are estimated from regressions that use heteroskedasticity-robust standard errors.

Table E.3: Heterogeneous Treatment Effects on Beliefs about Skills by Gender - Tight Experiment

	Aligned CA belief		Fraction aligned beliefs	
	(1)	(2)	(3)	(4)
Treatment	0.181*	0.178**	0.087	0.063
	(0.091)	(0.079)	(0.079)	(0.051)
Treatment × Female	-0.074	-0.062	0.059	0.022
	(0.107)	(0.103)	(0.095)	(0.062)
Female	0.082	0.019	-0.038	-0.033
	(0.066)	(0.063)	(0.060)	(0.042)
Treatment effect:	0.107	0.116**	0.146***	0.085**
Female	(0.065)	(0.047)	(0.049)	(0.032)
Control mean	0.475	0.475	0.183	0.183
Observations	278	278	278	278
Controls	No	Yes	No	Yes

Notes: **Table E.3 shows that treatment effects on skill beliefs do not substantially differ by gender in the tight experiment.** "CA" stands for comparative advantage in skills. Columns 1 and 2 show effects on a dummy equal to one if jobseekers' beliefs about their skill CA are aligned with their assessment results. Columns 3 and 4 show treatment effects on the fraction of skill beliefs that align with measured skill quintiles. Control variables are listed in footnote 18. Standard errors clustered at the treatment-day level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table E.4: Heterogeneous Treatment Effects on Beliefs about Skills by Gender - Big Experiment

	Aligned CA belief		Fraction aligned beliefs	
	(1)	(2)	(3)	(4)
Treatment	0.156***	0.155***	0.150***	0.153***
	(0.021)	(0.020)	(0.019)	(0.015)
Treatment × Female	-0.032	-0.026	-0.021	-0.016
	(0.027)	(0.026)	(0.020)	(0.017)
Female	-0.029	-0.011	-0.051***	0.003
	(0.018)	(0.016)	(0.015)	(0.012)
Treatment effect:	0.124***	0.130***	0.128***	0.136***
Female	(0.013)	(0.015)	(0.013)	(0.010)
Control mean	0.196	0.196	0.388	0.388
Observations	4191	4118	4205	4195
Controls	No	Yes	No	Yes

Notes: **Table E.4 shows that treatment effects on skill beliefs do not substantially differ by gender in the big experiment.** "CA" stands for comparative advantage in skills. Columns 1 and 2 show effects on a dummy equal to one if jobseekers' beliefs about their skill CA are aligned with their assessment results. Columns 3 and 4 show treatment effects on the fraction of skill beliefs that align with measured skill quintiles. Control variables are listed in footnote 25. Standard errors clustered at the treatment-day level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

F Preregistration Appendix

The tight experiment is preregistered at AEARCTR-0010000. The research question, experimental design, estimating equations, covariates, inference methods, outcomes, and heterogeneity analyses by aligned baseline CA beliefs and baseline confidence about levels of skills are all prespecified. We depart from the pre-analysis plan for the tight experiment in four ways:

1. We restrict the sample to jobseekers who have a clear CA because the skill-directed search measures can only be sensibly defined for these jobseekers. When we use this restricted sample, we omit one prespecified control variable, the dummy variable about whether the jobseeker has a clear CA, because this control variable has no variation in the restricted sample. We show the main results of the paper for the full sample as a robustness check in Table C.1. The interpretation of results remains unchanged.

2. We added further skill-directed search outcomes: clicks on jobs sent in SMS messages and job search on the SAYouth.mobi platform. We obtained these measures from our partner organization after we had lodged the pre-analysis plan. We correct for the inclusion of these additional measures by creating a summary index of all skill-directed search outcomes. We this additional outcome and the index in Table 4.

3. To align the search direction and search effort measures, we add the SMS click rate and the number of observed application clicks to the main search effort table and, again, construct a summary index. We show results for the prespecified platform search index in Table D.2.

4. Following recent methodological critiques (Chen & Roth, 2023; Mullahy & Norton, 2022) we use winsorization instead of the prespecified inverse hyperbolic sine transformation for outcomes with potentially long right tails (e.g. earnings, number of applications). Using the inverse hyperbolic sine transformation produces qualitatively similar results, which we show for earnings in Table C.4.

The tight experiment was set up to test the research question that arose from the exploratory analysis of data from the big experiment. All our analysis of the big experiment uses the same covariates and inference methods as Carranza et al. (2022), following their preanalysis plan. The new outcome variables and the heterogeneity analyses by aligned baseline CA beliefs that we add in this paper were not prespecified and were not reported in Carranza et al. (2022).

G Benefit-Cost Comparison

This appendix reports the variable costs of the assessment operation and compares these to the earnings gains experienced by treated jobseekers. All calculations use data from the big experiment in 2021 USD in purchasing power parity terms.

We calculate that the average variable cost of the assessment operation is 45.74 USD using data from Harambee and J-PAL Africa's accounting records. This consists of 12.17 for rent and utilities for the assessment center; 10.14 for depreciation of the computers used in assessment; 0.35 for assessment and software licenses; 8.11 for airtime and data, mainly for internet access for the assessment computers and for contacting jobseekers; 3.82 for Harambee salaries for assessment staff, psychologists delivering briefings about results, and administrative support; 1.40 for J-PAL Africa research staff salaries to help run the assessments and results briefings; and 9.74 for jobseekers' attendance payments.

The average treatment effect on weekly earnings was 6.52 USD at the time of the follow-up survey, which occurred an average of 14.5 weeks after treatment (Table 5, column 1). Using this to forecast the average lifetime earnings gain for treated jobseekers requires very strong modeling assumptions. Instead, we note that this earnings gain needs to hold for only 7 weeks for the average earnings gain to exceed the average variable cost ($6.52 \times 7 > 45.74$). This would hold if, for example, the earnings gain held for at least half the period between workshop and endline. This is plausible, particularly because treated participants have a higher employment rate than untreated participants for part of this period, raising the possibility of even higher earnings from a higher employment rate (Table 5, columns 10-12). If the earnings gain held in every week between the workshop and endline, treated jobseekers would accumulate 94.5 USD higher earnings. This implies that the average benefit / average variable cost would be **94.50 / 45.74 = 1.82**.

We view this benefit-cost comparison as suggestive rather than conclusive because, like all such calculations, it requires some simplifying assumptions. We use a conservative approach to estimating lifetime benefits, we omit average fixed costs because these are very dependent on the scale of the assessment service, and we do not consider general equilibrium effects or the benefits that might be accrued from alternative uses of money. However, our measure of average variable costs is relatively broad, as it includes semi-fixed costs such as facility and equipment rental, most staff costs, and assessment and software licenses. The only costs that we exclude are those of actually creating the organization, senior management time, and general functions such as accounting.

There is scope to run similar interventions far more cheaply through online platforms that incorporate assessments and personalized, automated feedback on results. This approach would reduce or eliminate most of the variable costs of in-person assessment: facility rental (27% of average variable cost), equipment rental (22%), and transport (21%).

References

Carranza, E., R. Garlick, K. Orkin & N. Rankin (2022) "Job Search and Hiring with Limited Information about Workseekers' Skills," *American Economic Review*, 112 (11), 3547–83.

Chen, J. & J. Roth (2023) "Logs With Zeros? Some Problems and Solutions," Working Paper.

Delavande, A. (2023) "Expectations in Development Economics," in Bachmann, R., G. Topa & W. van der Klaauw eds. *Handbook of Economic Expectations*, Chap. 9, 261–291: Academic Press.

Kessler, J. B., C. Low & C. D. Sullivan (2019) "Incentivized Resume Rating: Eliciting Employer Preferences without Deception," *American Economic Review*, 109 (11), 3713–44.

Mueller, A. I. & J. Spinnewijn (2023) "Expectations Data, Labor Market, and Job Search," in Bachmann, R., G. Topa & W. van der Klaauw eds. *Handbook of Economic Expectations*, Chap. 22, 677–713: Academic Press.

Mullahy, J. & E. C. Norton (2022) "Why Transform Y? A Critical Assessment of Dependent-Variable Transformations in Regression Models for Skewed and Sometimes-Zero Outcomes," NBER Working Paper 30735.

Rosenberg, M. (1965) *Society and the Adolescent Self-Image*: Princeton University Press.

Wiswall, M. & B. Zafar (2015) "Determinants of College Major Choice: Identification Using an Information Experiment," *Review of Economic Studies*, 82 (2), 791–824.