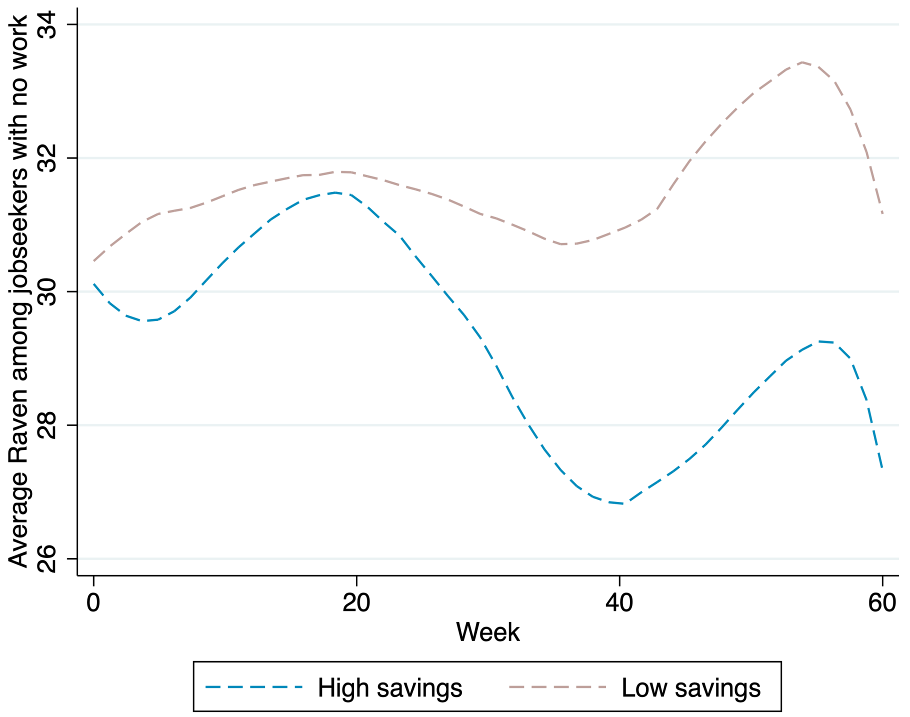
Online Appendix

The Selection of Talent: Experimental and Structural Evidence from Ethiopia

Girum Abebe, A. Stefano Caria, Esteban Ortiz-Ospina

A Figures and tables

Figure A1: The selection mechanism

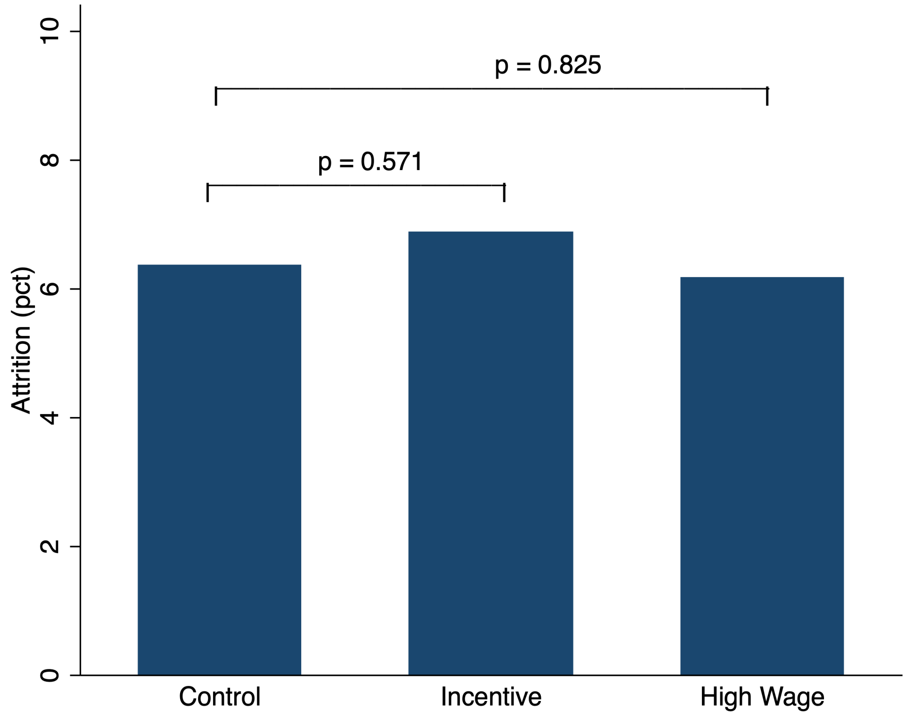


Notes: This figure reports data from the high-frequency survey of Abebe et al. (2020). We plot the average Raven test score among jobless individuals who are searching for work at a given point in time. Changes in this variable are due to movements in and out of job search over time. The figure is produced using Stata's `lpoly` command.

Figure A2: Timeline

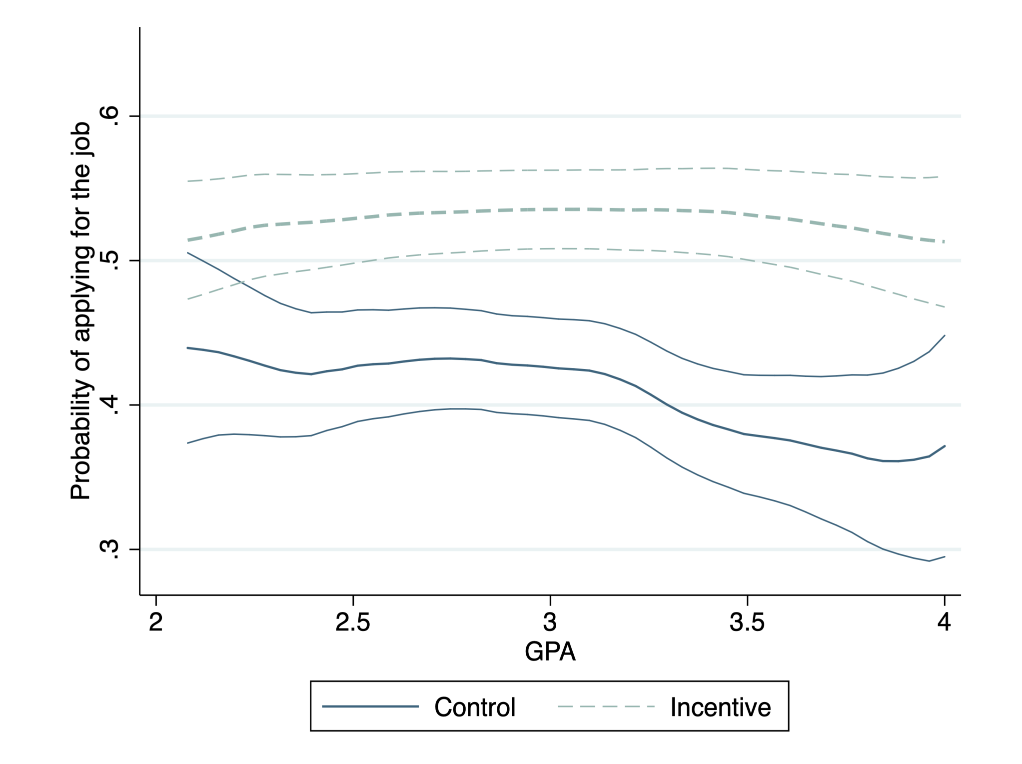
Notes: This figure shows the timeline of a typical hiring round. The position is advertised at the start of the week (day 1 in the timeline). Jobseekers can call to inquire about the position until close-of-business on Friday of that week (day 5). Those jobseekers who call to inquire about the position are then invited to make an in-person application on a randomly assigned day on the following week (days 7 to 12) or on the Monday of the third week (day 15). We thus have 6 application days, 2 of which are assigned to each experimental group. Finally, all jobseekers who call to inquire about the position are called again 30 days after the initial phone call. The jobseekers who are invited for an interview are told about this in a separate phone call shortly before the second phone call. Interviews are held shortly thereafter and the position starts right away. We collect data on jobseekers at each stage of the experiment. During the first phone call, we collect data on their socio-demographic characteristics, labour market experience and GPA. If jobseekers apply for the position, they have to complete several tests of ability (Raven and Stroop), and answer psychometric questions and questions about their economic preferences. Finally, in the second phone call, jobseekers are asked about job search in the last 30 days.

Figure A3: Attrition



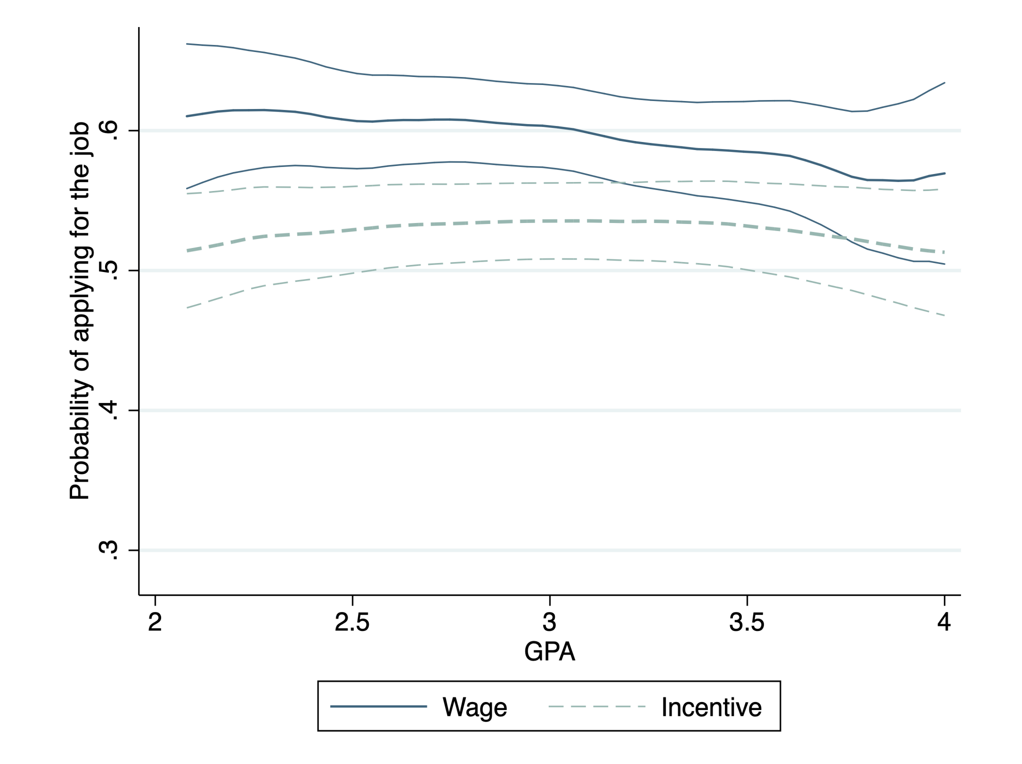
Notes: This figure shows descriptive data on attrition in the three experimental groups. An individual is considered attrited when our team is unable to contact them for the 30-days follow-up phone survey. The bars indicate total attrition in each group. Further, we report the p-values for a test of the null hypothesis of no differential attrition between a given treatment group and the control group. Sample used: baseline sample.

Figure A4: Applications and GPA score: Control and application incentive groups



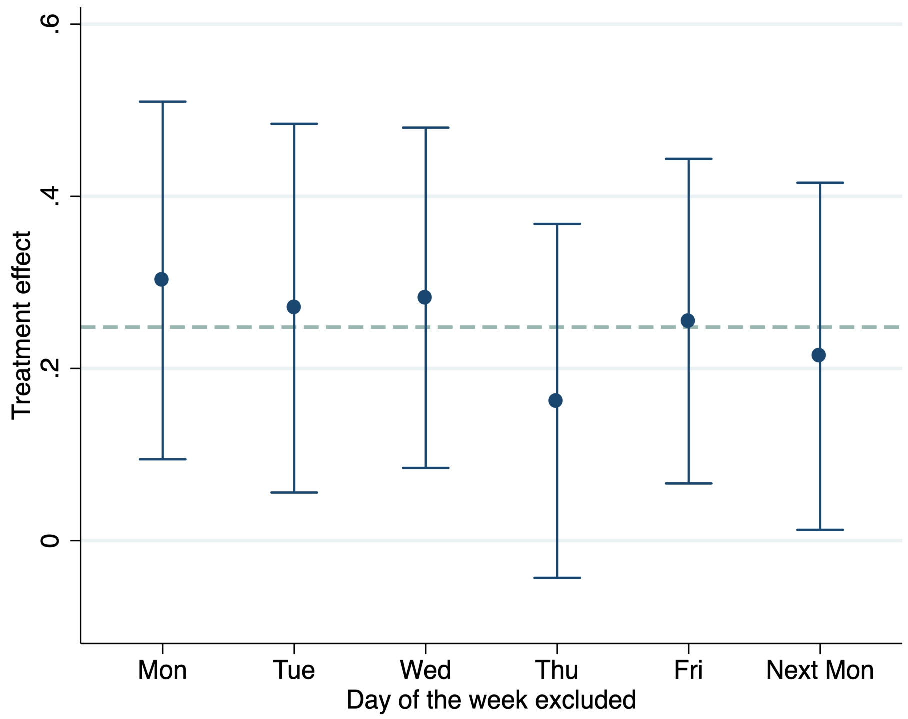
Notes: This figure shows the relationship between a jobseeker's GPA and the probability of applying for the experiment's job. We separately plot this relationship for individuals in the control and application incentive groups. We report 95 percent confidence intervals. The figure is produced using Stata's `lpolyci` command. Sample used: baseline sample (control and incentive groups).

Figure A5: Applications and GPA score: Incentive and high wage groups



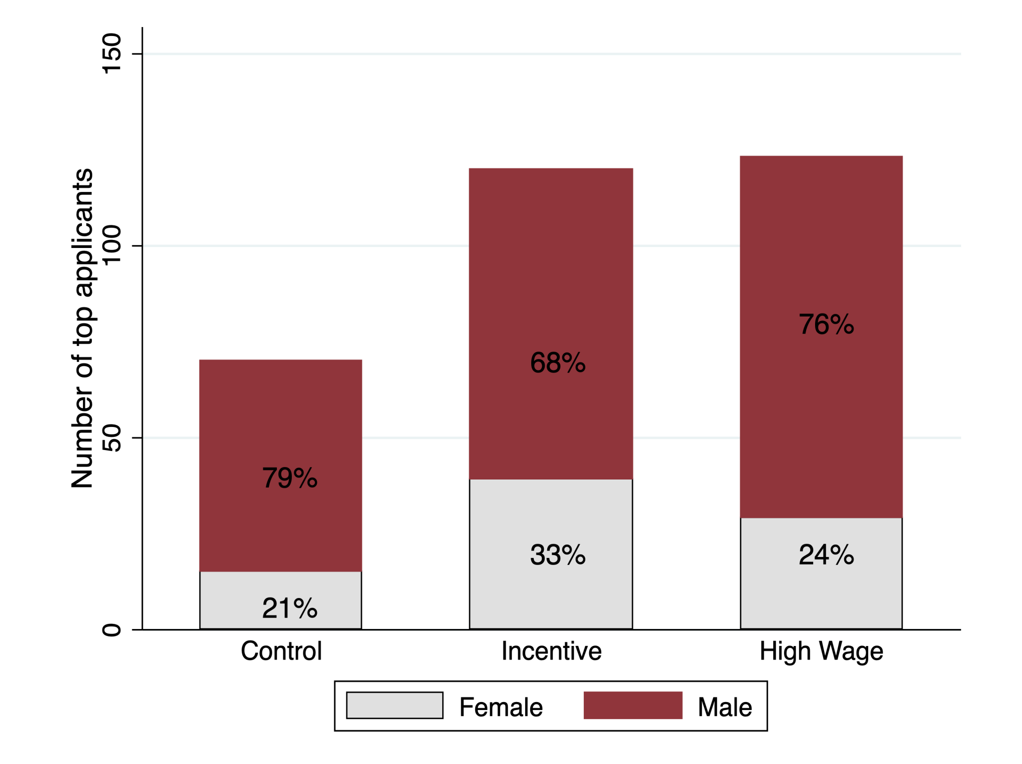
Notes: This figure shows the relationship between a jobseeker's GPA and the probability of applying for the experiment's job. We separately plot this relationship for individuals in the application incentive and high wage groups. We report 95 percent confidence intervals. The figure is produced using Stata's `lpolyci` command. Sample used: baseline sample (control and high wage groups).

Figure A6: Robustness to exclusion of selected days of the week



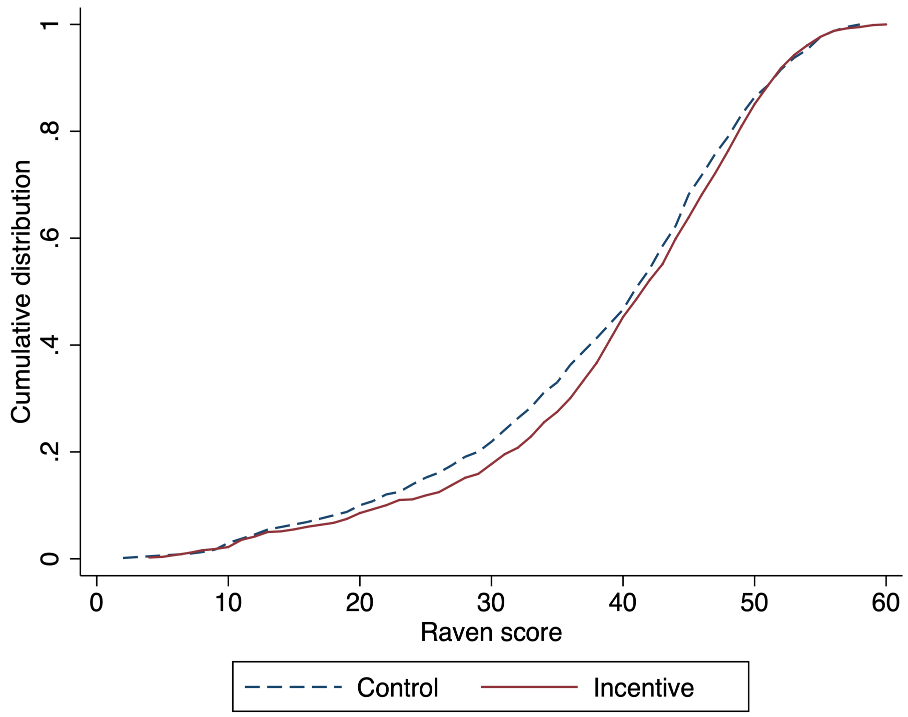
Notes: This figure shows the OLS estimates of the effect of the application incentive on average applicant ability for different samples. Each sample is obtained by dropping all individuals who are invited to take the test on a specific day of the week. The dashed horizontal line indicates the treatment effect for the full sample. Overall sample used: all applicants.

Figure A7: The proportion of female top applicants



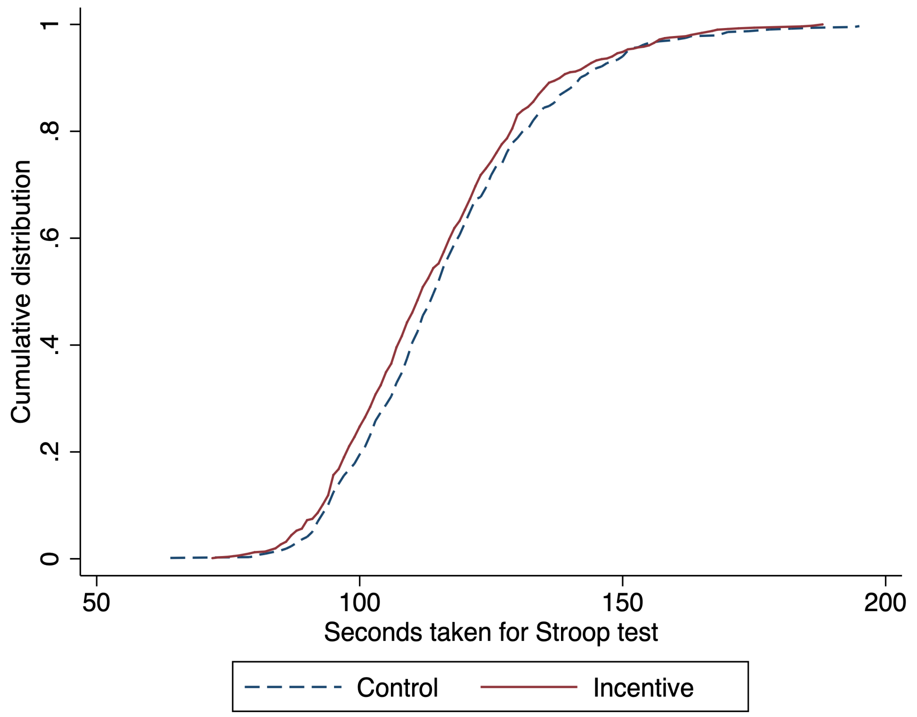
Notes: This figure plots the number of top applicants in each experimental group. Within each bar, we also report the proportion of top applicants in that experimental group that are female and the proportion of top applicants that are male. A 'top applicant' is defined as somebody whose cognitive ability is above the 90th percentile of the control group distribution of cognitive ability. Sample used: all top applicants.

Figure A8: Impacts of incentives on the distribution of applicant Raven test score



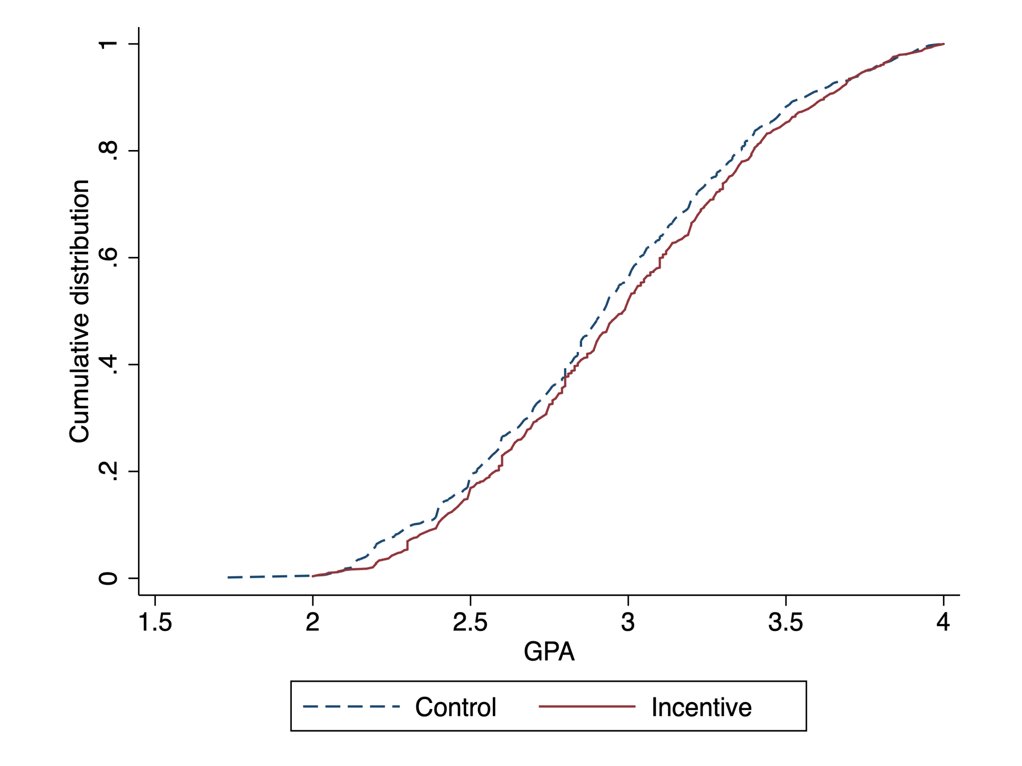
Notes: This figure plots the distribution of Raven test scores among control and application incentive applicants in the experiment. Sample used: all applicants (control and incentive groups).

Figure A9: Impacts of incentives on the distribution of applicant Stroop test performance



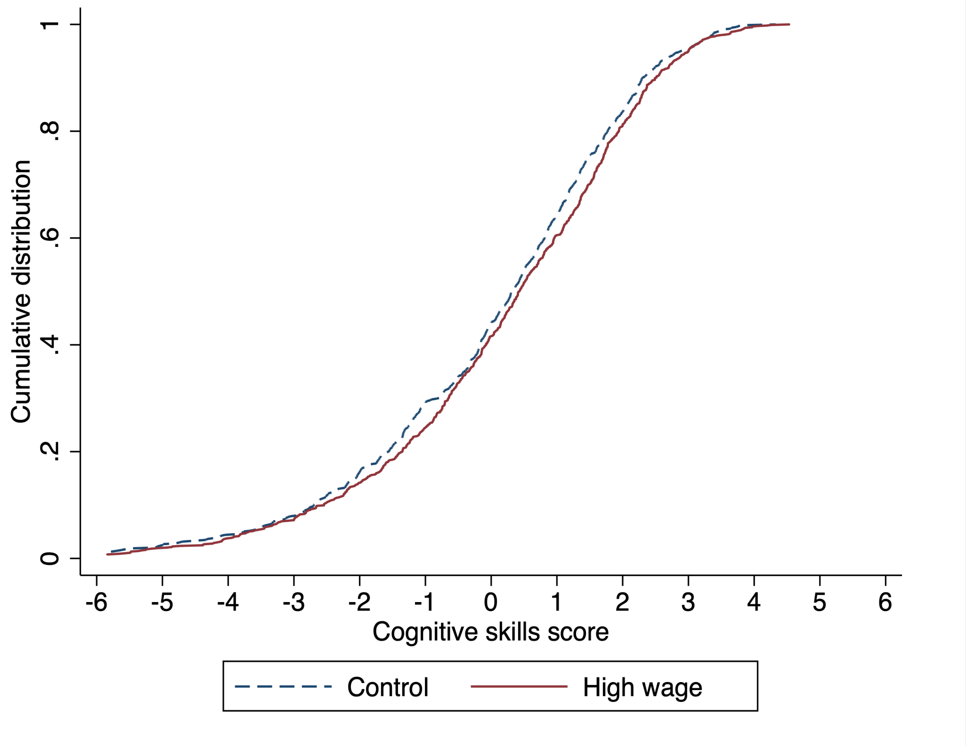
Notes: This figure plots the distribution of time taken to complete the Stroop test among control and application incentive applicants in the experiment. A smaller value indicates better performance. Sample used: all applicants (control and incentive groups).

Figure A10: Impacts of incentives on the distribution of applicant GPA score



Notes: This figure plots the distribution of GPA scores among control and application incentive applicants in the experiment. Sample used: all applicants (control and incentive groups).

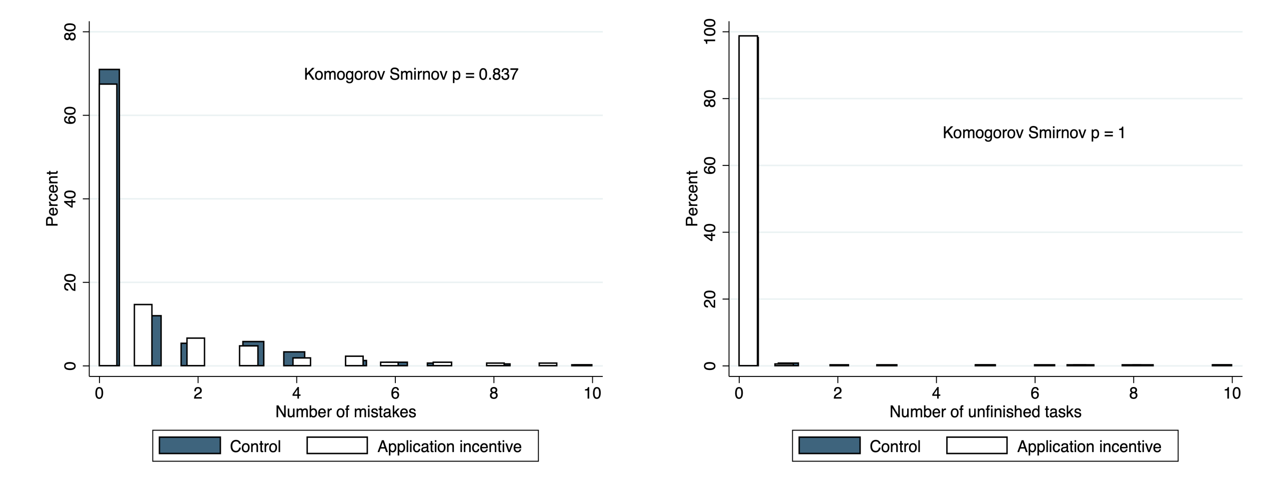
Figure A11: Impacts of high wage offer on the distribution of applicant cognitive ability



Notes: This figure plots the distribution of cognitive ability among control and application incentive applicants in the experiment. Sample used: all applicants (control and high wage groups).

Figure A12: Distribution of test effort

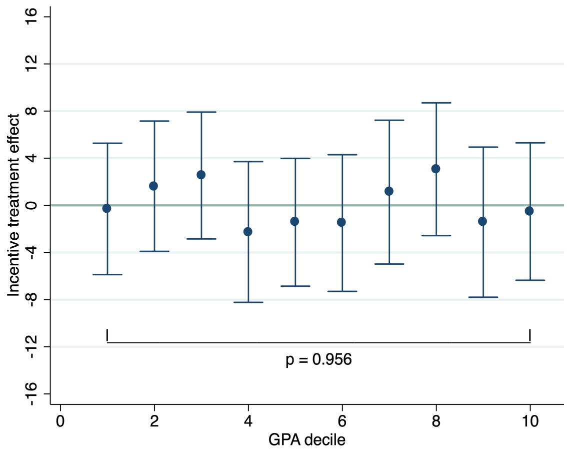
(a) No. mistakes (b) No. unfinished tasks



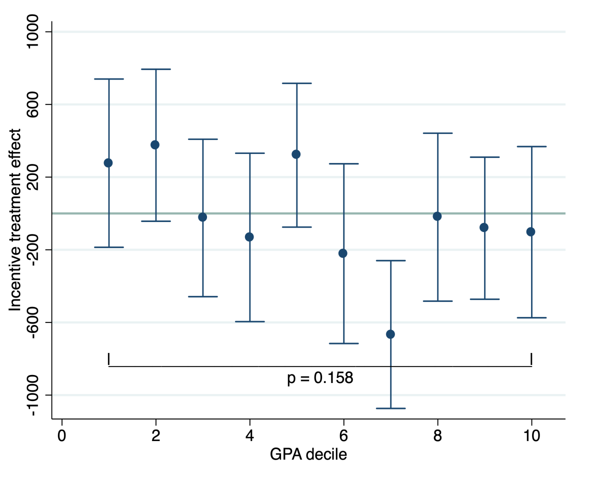
Notes: These figures show histograms of the two measures of test effort for applicants from the control and application incentive groups. The figures also report a p-value for a Kolmogorov Smirnov test of the equality of the two distributions. Sample used: all applicants (control and incentive groups).

Figure A13: Belief updating and GPA

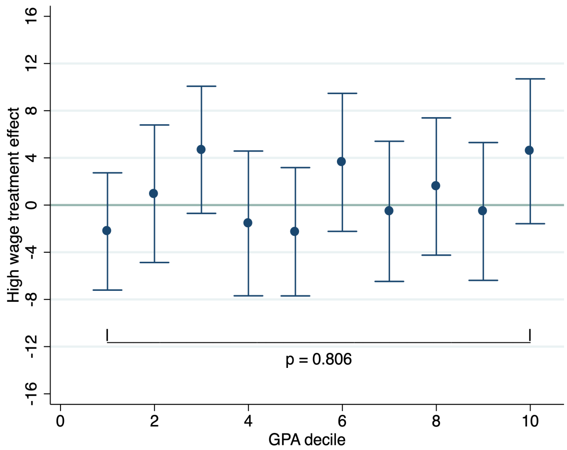
(a) Impact of incentive on E[getting the job]



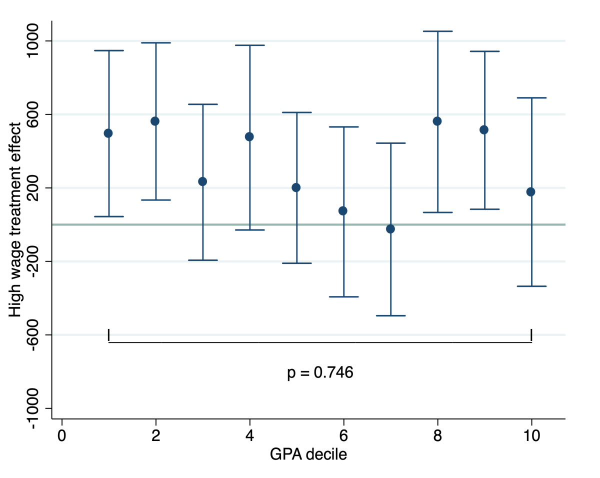
(b) Impact of incentive on E[wage]



(c) Impact of high wage on E[getting the job]



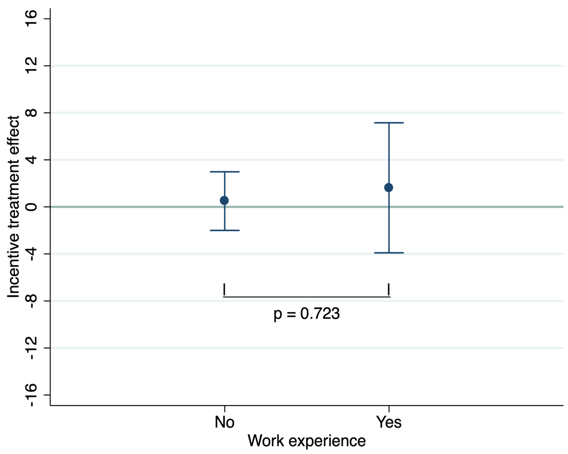
(d) Impact of high wage on E[wage]



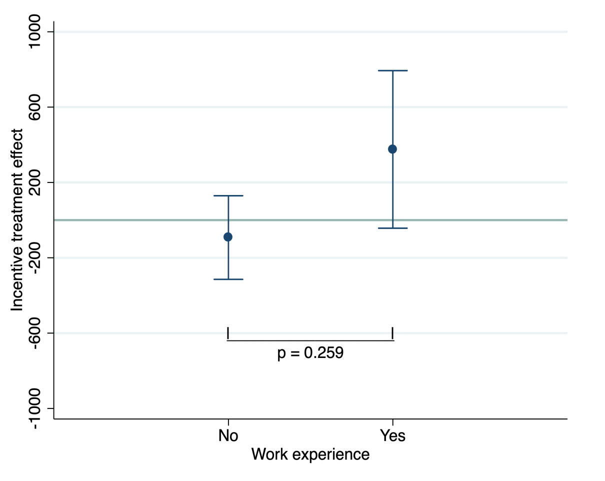
Notes: These figures show point estimates and 90% confidence intervals of the treatment effects on participants' beliefs about (i) the probability of getting the experiment's job, and (ii) the wage they would earn in their next job. Each figure also shows a p-value for a test of the null hypothesis that the coefficients are equal to each other. Sample used: baseline sample.

Figure A14: Belief updating and experience

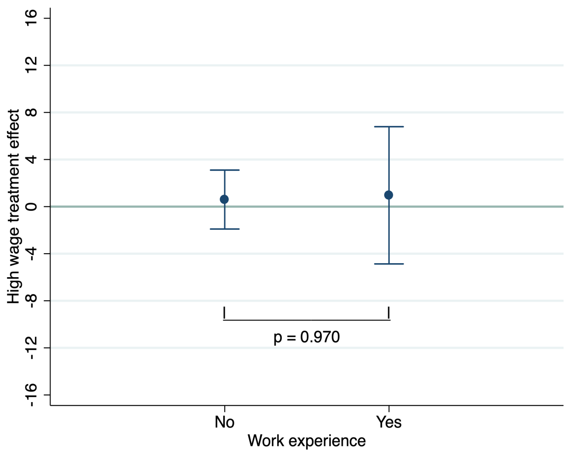
(a) Impact of incentive on E[getting the job]



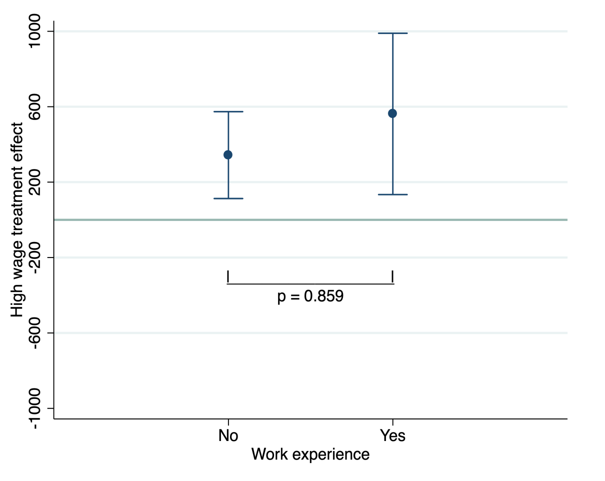
(b) Impact of incentive on E[wage]



(c) Impact of high wage on E[getting the job]



(d) Impact of high wage on E[wage]



Notes: These figures show point estimates and 90% confidence intervals of the treatment effects on participants' beliefs about (i) the probability of getting the experiment's job, and (ii) the wage they would earn in their next job. Each figure also shows a p-value for a test of the null hypothesis that the coefficients are equal to each other. Sample used: baseline sample.

Figure A15: Sample CV

Candidate 1

Age: 33

Education

Highest level of education: BA (BSc) degree

Field of study: Natural and Computational Sciences

Average grade/GPA: 2.57

Work Experience

Has work experience? Yes

Last employer: MAYLEKO LOAGE

Type of employer: Private business

Test scores

Cognitive ability score: 410

Non-cognitive ability score: 600

Table A1: Occupations included in the firm survey

Job code	Description	Examples
43-1	Supervisors of Support Workers	First-Line supervisors of office support workers
43-2	Communications Equipment Operators	Telephone operators
43-3	Financial Clerks	Bill and Account Collectors; Bookkeeping, accounting, and auditing clerks
43-4	Information and Record Clerks	Correspondence clerks, credit checkers, customer service representatives
43-9	Other Administrative Support Workers	Computer and data entry operators, claims processing clerks
13-1	Business Operations Specialists	Buyers and purchasing agents, cost estimators, claim checkers, logisticians
13-2	Legal Support Workers	Legal assistants, court workers

Table A2: Ability and labour market outcomes

Dep. var.	Ln(wage)				Employed			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Raven	0.090			0.090	0.046			0.046
	(0.035)			(0.035)	(0.018)			(0.018)
Conscientiousness	-0.032	0.019		0.018		0.007		0.004
	(0.025)	(0.030)		(0.031)		(0.018)		(0.019)
Neuroticism	-0.047		-0.005	-0.002			-0.009	-0.007
	(0.025)		(0.035)	(0.035)			(0.018)	(0.019)
Mean dep. var.	1938.37				0.56			
Obs.	424	424	424	424	780	780	780	780

Notes: OLS regression based on data from Abebe et al. (2020). The dependent variable in the first four columns is the natural logarithm of wages in the first endline survey. The dependent variable in the last four columns is a dummy equal to one if the respondent is employed at the time of the first endline survey. 'Raven' is the number of correct answers in a 60 item Raven test, which was administered shortly after the baseline interview. 'Conscientiousness' and 'neuroticism' are the conscientiousness and neuroticism scores obtained by administering a 10-items BFI inventory at baseline. All ability variables are standardised so that they have a mean of zero and a standard deviation of one. We report the mean of the dependent variables in the second-to-last row. For the wage variable, we report the mean wage in Ethiopian Birr units (as opposed to the mean of the natural logarithm).

Table A3: Ability and labour market outcomes
with controls

Dep. var.	Ln(wage)				Employed			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Raven	0.084			0.084	0.048			0.047
	(0.034)			(0.034)	(0.018)			(0.018)
Conscientiousness	-0.032	0.015		0.014		0.007		0.004
	(0.025)	(0.030)		(0.032)		(0.018)		(0.019)
Neuroticism	-0.047		-0.001	0.000			-0.008	-0.007
	(0.025)		(0.034)	(0.035)			(0.018)	(0.019)
Mean dep. var.	1938.37				0.56			
Obs.	424	424	424	424	780	780	780	780

Notes: OLS regression based on data from Abebe et al. (2020). The dependent variable in the first four columns is the natural logarithm of wages in the first endline survey. The dependent variable in the last four columns is a dummy equal to one if the respondent is employed at the time of the first endline survey. 'Raven' is the number of correct answers in a 60 item Raven test, which was administered shortly after the baseline interview. 'Conscientiousness' and 'neuroticism' are the conscientiousness and neuroticism scores obtained by administering a 10-items BFI inventory at baseline. All ability variables are standardised so that they have a mean of zero and a standard deviation of one. We report the mean of the dependent variables in the second-to-last row. For the wage variable, we report the mean wage in Ethiopian Birr units (as opposed to the mean of the natural logarithm). All regressions include controls for age and age squared and a dummy for having worked in a permanent job in the past.

Table A4: The probability of finding a job for high and low-cost individuals

	Search-to-work transition	
	(1)	(2)
Raven (z-score)	0.022	0.016
	(0.019)	(0.018)
Low saving * Raven	-0.047	
	(0.025)	
High distance * Raven		-0.038
		(0.026)
Low saving	-0.032	
	(0.025)	
High distance		0.016
		(0.025)
Constant	0.235	0.207
	(0.020)	(0.016)
Obs.	2218	2218

Notes: OLS regression based on data from Abebe et al. (2020). The data is collapsed at the monthly level. We restrict the sample to unemployed people who are searching for work. The dependent variable is a dummy capturing whether the respondent is employed in at least one of the fortnights of the following month. The model enables to estimate the correlation between the probability of finding a job in the following month and (i) the z-score of the Raven test, (ii) a proxy for application costs, and (iii) the interaction between the two. Standard errors clustered at the individual level reported in parenthesis.

Table A5: Psychometrics Validity Checks: Cronbach α

	Raw	Ipsatised	Laajaj and Macours (2017)
Conscientiousness	0.59	0.70	0.51
Neuroticism	0.61	0.62	0.31
Grit	0.59	0.72	

Notes: In the first column, we report the value of Cronbach α for our three main measures of non-cognitive ability. In the second column, we show the Cronbach α for the ipsatised values of these variables. Variables are ipsatised by subtracting the individual acquiescence score — the mean of positive items and inverted items across scales — which is a measure of the tendency to agree with any statement (Laajaj and Macours, 2017). In third column, we report for reference the values of Cronbach α from a recent study in Kenya by Laajaj and Macours (2017). Sample used: all applicants.

Table A6: Indices of applicant quality

Index	Variable	Measure
Cognitive ability	Raven	No. of correct answers
	Stroop	Time in seconds
	Stroop	No. mistakes
Non-cognitive ability	Conscientiousness	BFI44 score
	Neuroticism	BFI44 score
	Grit	Score on grit scale
Experience	Routine tasks	No. months
	Managerial tasks	No. months
	Problem solving tasks	No. months

Table A7: Correlation between indices

	Cognitive ability	Non-cognitive ability	Experience
Cognitive ability	1		
Non-cognitive ability	0.205	1	
Experience	-0.002	0.064	1

Notes: Correlation coefficients. Sample used: all applicants.

Table A8: Summary statistics and balance for the sample of non-attriters

	Mean			StDev	N	Balance tests (p)		
	Control	Incentive	High wage	Control		Incentive = Wage = Control	Incentive = Control	Wage = Control
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	0.21	0.21	0.21	0.41	4386	0.90	0.71	0.68
Age	25.96	25.76	26.10	4.23	4383	0.11	0.20	0.41
Born in Addis Ababa	0.24	0.24	0.23	0.43	4386	0.59	0.81	0.45
First language is Amharic	0.68	0.70	0.67	0.47	4386	0.20	0.16	0.81
Heard about job in newspaper	0.55	0.58	0.55	0.50	4386	0.21	0.12	0.97
Engineering or hard science	0.51	0.49	0.49	0.50	4386	0.44	0.34	0.22
Economics	0.15	0.17	0.16	0.36	4386	0.62	0.34	0.54
Other social science	0.15	0.15	0.14	0.36	4386	0.45	0.68	0.41
Wage work experience (dummy)	0.52	0.51	0.52	0.50	4386	0.91	0.83	0.83
Wage work experience (months)	27.15	26.93	28.27	42.92	4386	0.71	0.89	0.51
Self-employed experience	0.33	0.34	0.35	0.47	4386	0.67	0.61	0.37
Currently unemployed	0.67	0.66	0.64	0.47	4386	0.15	0.33	0.05
Currently wage employed	0.24	0.25	0.27	0.43	4386	0.18	0.45	0.07
GPA	2.98	2.99	2.96	0.46	4168	0.20	0.48	0.28
Overall balance						0.24	0.53	0.38

Notes: Balance table. In columns 1-5, we report descriptive statistics and sample sizes for the sample of individuals that were interviewed in the second phone survey. In columns 6-8, we report the p-values of a battery of balance tests. In column 6, we test the null hypothesis that three experimental groups are balanced. In columns 7 and 8, we present pair-wise tests of a single treatment group against the control. In the last row, we report a joint test of orthogonality (following the recent literature, e.g. McKenzie (2017)). To perform the joint test of orthogonality we regress the treatment variable on all covariates and we then test the joint null hypothesis that all covariates have a zero coefficient. We do this by estimating a categorical logit model in column 6 (the treatment variable can take three values) and OLS models in columns 7 and 8 (here we drop one experimental group, so our treatment variable only takes two values). Sample used: second phone call sample.

Table A9: Additional balance tests

	Week							
	1	2	3	4	5	6	7	8
Incentive	-0.001	0.000	0.000	0.001	0.000	0.002	-0.001	-0.002
	(0.010)	(0.012)	(0.011)	(0.012)	(0.012)	(0.013)	(0.012)	(0.012)
High Wage	0.001	-0.001	-0.002	0.002	-0.001	0.001	-0.000	-0.000
	(0.010)	(0.012)	(0.011)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)

Notes: OLS regression testing balance of the day of the week when the application test took place. Robust standard errors reported in parenthesis. Sample used: baseline sample.

Table A10: Is the sample representative?

	Age	Female	Work experience	Unemployment duration
Experiment	25.13	0.23	0.34	4.99
	(3.85)	(0.42)	(0.47)	(7.02)
Labour Force Survey	26.72	0.36	0.29	9.16
	(8.38)	(0.49)	(0.46)	(7.52)

Notes: This table presents means and standard deviations of key variables in the experimental sample and in a representative sample of jobseekers from Addis Ababa who (i) use job boards or newspapers for job search and (ii) have the educational qualifications required to apply for the experiment's job (they hold a vocational diploma or a university degree). Unemployment duration is measured in months. Sample used: baseline sample and 2013 Labour Force Survey sample.

Table A11: Components of index

	Raven	Stroop time	Stroop mistakes
	(1)	(2)	(3)
Incentive	1.155	-2.601	-0.050
	(0.618)	(1.046)	(0.188)
	[0.092]	[0.039]	[0.791]
High wage	0.591	-0.982	-0.302
	(0.618)	(1.046)	(0.188)
	[0.337]	[0.337]	[0.297]
Control mean	38.593	117.304	3.854
Incentive = Wage (p)	0.307	0.078	0.098
Obs.	2397	2386	2388

Notes: OLS regression. The dependent variable is indicated in the column heading. 'Raven' is the number of correctly answered questions on the Raven test. 'Stroop time' is the number of seconds required to complete the Stroop test. 'Stroop mistakes' is the number of mistakes made in the Stroop task. The negative coefficients on 'Stroop time' and 'Stroop mistakes' indicate better performance. Sharpened q-values (Benjamini et al., 2006) reported in brackets. q-values control the false discovery rate for the multiple comparisons reported in the same row of the table. The second to last row reports the p-value of a test of the null hypothesis that the treatments have the same effect. Robust standard errors reported in parenthesis. Sample used: all applicants.

Table A12: Impacts on applicant GPA

	Mean	Percentile				
		90th	75th	50th	25th	10th
	(1)	(2)	(3)	(4)	(5)	(6)
Incentive	0.049	0.070	0.050	0.060	0.040	0.080
	(0.025)	(0.043)	(0.035)	(0.033)	(0.033)	(0.036)
High Wage	0.012	-0.010	-0.010	0.030	0.000	0.040
	(0.024)	(0.040)	(0.033)	(0.033)	(0.030)	(0.039)
Control value	2.94	3.56	3.27	2.93	2.60	2.32
Incentive = Wage (p)	0.088	0.022	0.045	0.360	0.181	0.187
Obs.	2285	2285	2285	2285	2285	2285

Notes: Estimates from OLS (Column 1) and quantile (Columns 2-6) regressions. The dependent variable is the applicant's GPA. The second to last row reports the p-value of a test of the null hypothesis that the treatments have the same effect. Robust standard errors reported in parenthesis. Sample used: all applicants.

Table A13: Proportion of applicants who score above a threshold

	Threshold (percentile in control group distribution)				
	90th	75th	50th	25th	10th
	(1)	(2)	(3)	(4)	(5)
Incentive	0.053	0.229	0.229	0.170	0.412
	(0.024)	(0.110)	(0.117)	(0.133)	(0.173)
High Wage	0.052	0.202	0.227	0.075	0.280
	(0.023)	(0.108)	(0.112)	(0.131)	(0.165)
Incentive = Wage (p)	0.966	0.795	0.983	0.448	0.371
Obs.	2386	2386	2386	2386	2386

Notes: OLS regression. The dependent variable is a dummy for whether the applicant's cognitive ability is above the threshold indicated in the column heading. Thresholds are defined with respect to specific percentiles of the control distribution of cognitive ability. The second to last row reports the p-value of a test of the null hypothesis that the treatments have the same effect. Robust standard errors reported in parenthesis. Sample used: all applicants.

Table A14: Average ability of top candidates

	Top 20		Top 10		Top 5	
	Cognitive	GPA	Cognitive	GPA	Cognitive	GPA
	(1)	(2)	(3)	(4)	(5)	(6)
Incentive	0.359	0.102	0.243	0.182	0.150	0.233
	(0.078)	(0.050)	(0.091)	(0.069)	(0.114)	(0.095)
High Wage	0.457	0.066	0.354	0.079	0.282	0.144
	(0.078)	(0.052)	(0.092)	(0.073)	(0.116)	(0.100)
Control group st. dev.	0.71	0.45	0.59	0.44	0.53	0.42
Incentive = Wage (p)	0.173	0.480	0.206	0.138	0.260	0.324
Obs.	480	466	240	233	120	116

Notes: Estimates from OLS regressions. The sample comprises the top 20, 10 and 5 applicants for job offered (top applicants are defined using the cognitive ability score, following the procedures used in the field experiment). The employer offers one job per fortnight for each treatment group. In a few cases, the employer combines two fortnights for the same treatment group and offers only one job for these two fortnights. For the present analysis, however, we consider the top applicants from each fortnight separately. The second to last row reports the p-value of an F-test of the null hypothesis that the treatments have the same effect. Robust standard errors reported in parenthesis. Sample used: all top applicants.

Table A15: Non-cognitive ability

	Mean	Percentile				
		90th	75th	50th	25th	10th
	(1)	(2)	(3)	(4)	(5)	(6)
Incentive	-0.095	-0.155	-0.161	-0.004	-0.141	-0.137
	(0.125)	(0.134)	(0.132)	(0.157)	(0.199)	(0.229)
	[0.566]	[0.369]	[0.336]	[1.000]	[0.717]	[1.000]
High Wage	0.170	-0.000	-0.039	0.257	0.162	0.260
	(0.118)	(0.127)	(0.129)	(0.147)	(0.203)	(0.222)
	[0.225]	[1.000]	[0.764]	[0.245]	[0.637]	[0.725]
Control value	0.0000	2.688	1.721	0.241	-1.227	-2.885
Incentive = Wage (p)	0.015	0.235	0.288	0.038	0.095	0.091
Obs.	2373	2373	2373	2373	2373	2373

Notes: Estimates from OLS (Column 1) and quantile (Columns 2-6) regressions. The dependent variable is the index of non-cognitive ability. The second-to-last row reports the p-value of a test of the null hypothesis that the treatments have the same effect. Robust standard errors are reported in parenthesis. Sharpened q-values (Benjamini et al., 2006) are reported in brackets. q-values control the false discovery rate for the multiple tests of the same hypothesis for different indices of ability. A Wilcoxon rank-sum test fails to reject the equality of the distribution of non-cognitive ability in the control and incentive groups (p=.411) and fails to reject the equality of the distribution of non-cognitive ability in the control and wage groups (p=.255). Sample used: all applicants.

Table A16: Experience

	Mean	Percentile				
		90th	75th	50th	25th	10th
	(1)	(2)	(3)	(4)	(5)	(6)
Incentive	-0.091	0.163	-0.044	0.000	0.000	0.000
	(0.158)	(0.850)	(0.117)	(0.008)	(0.004)	(0.002)
	[0.566]	[0.848]	[0.704]	[1.000]	[1.000]	[1.000]
High Wage	-0.063	0.302	-0.088	0.000	0.000	0.000
	(0.157)	(0.733)	(0.147)	(0.008)	(0.004)	(0.002)
	[0.687]	[1.000]	[0.764]	[1.000]	[1.000]	[1.000]
Control value	0.0000	2.217	0.064	-1.225	-1.225	-1.225
Incentive = Wage (p)	0.808	0.811	0.718	1.000	1.000	1.000
Obs.	2311	2311	2311	2311	2311	2311

Notes: Estimates from OLS (Column 1) and quantile (Columns 2-6) regressions. The dependent variable is the experience index. The second-to-last row reports the p-value of a test of the null hypothesis that the treatments have the same effect. Robust standard errors are reported in parenthesis. Sharpened q-values (Benjamini et al., 2006) are reported in brackets. q-values control the false discovery rate for the multiple tests of the same hypothesis for different indices of ability. A Wilcoxon rank-sum test fails to reject the equality of the distribution of experience in the control and incentive groups (p=.354) and fails to reject the equality of the distribution of experience in the control and wage groups (p=.718). We report results obtained using an alternative definition of the index in Table . used: all applicants.

Table A17: Cognitive ability, weighted index

	Mean	Percentile				
		90th	75th	50th	25th	10th
	(1)	(2)	(3)	(4)	(5)	(6)
Incentive	0.079	0.068	0.060	0.062	0.114	0.022
	(0.037)	(0.035)	(0.038)	(0.045)	(0.060)	(0.074)
High Wage	0.066	0.065	0.067	0.028	0.085	0.037
	(0.037)	(0.035)	(0.036)	(0.043)	(0.057)	(0.070)
Control value	-0.0000	0.769	0.499	0.113	-0.403	-0.864
Incentive = Wage (p)	0.699	0.917	0.819	0.396	0.575	0.847
Obs.	2386	2386	2386	2386	2386	2386

Notes: Estimates from OLS (Column 1) and quantile (Columns 2-6) regressions. Cognitive ability index obtained by weighting observations by the inverse of the covariance matrix. The second-to-last row reports the p-value of a test of the null hypothesis that the treatments have the same effect. Robust standard errors are reported in parenthesis. A Wilcoxon rank-sum test rejects the equality of the distribution of cognitive ability in the control and incentive groups (p=.057) and marginally rejects the equality of the distribution of cognitive ability in the control and wage groups (p=.096). Sample used: all applicants.

Table A18: Non cognitive ability, weighted index

	Mean	Percentile				
		90th	75th	50th	25th	10th
	(1)	(2)	(3)	(4)	(5)	(6)
Incentive	-0.036	-0.059	-0.069	0.015	-0.068	-0.051
	(0.042)	(0.050)	(0.043)	(0.052)	(0.068)	(0.091)
High Wage	0.057	0.003	-0.035	0.102	0.068	0.109
	(0.039)	(0.047)	(0.043)	(0.050)	(0.065)	(0.082)
Control value	0.0000	0.929	0.585	0.059	-0.423	-0.988
Incentive = Wage (p)	0.010	0.177	0.387	0.049	0.018	0.053
Obs.	2373	2373	2373	2373	2373	2373

Notes: Estimates from OLS (Column 1) and quantile (Columns 2-6) regressions. Non-cognitive ability index obtained by weighting observations by the inverse of the covariance matrix. The second-to-last row reports the p-value of a test of the null hypothesis that the treatments have the same effect. Robust standard errors are reported in parenthesis. A Wilcoxon rank-sum test fails to reject the equality of the distribution of non-cognitive ability in the control and incentive groups (p=.371) and fails to reject the equality of the distribution of non-cognitive ability in the control and wage groups (p=.239). Sample used: all applicants.

Table A19: Experience, weighted index

	Mean	Percentile				
		90th	75th	50th	25th	10th
	(1)	(2)	(3)	(4)	(5)	(6)
Incentive	-0.029	0.007	-0.014	0.000	0.000	0.000
	(0.053)	(0.142)	(0.058)	(0.002)	(0.001)	(0.000)
High Wage	-0.020	0.063	-0.022	0.000	0.000	0.000
	(0.052)	(0.127)	(0.057)	(0.002)	(0.001)	(0.000)
Control value	-0.0000	0.753	0.017	-0.409	-0.409	-0.409
Incentive = Wage (p)	0.820	0.685	0.830	1.000	1.000	1.000
Obs.	2311	2311	2311	2311	2311	2311

Notes: Estimates from OLS (Column 1) and quantile (Columns 2-6) regressions. Experience ability index obtained by weighting observations by the inverse of the covariance matrix. The second-to-last row reports the p-value of a test of the null hypothesis that the treatments have the same effect. Robust standard errors are reported in parenthesis (for the 50th, 25th and 10th we estimate coefficients of zero and are thus unable to calculate robust standard errors; for these percentiles we report raw standard errors instead). A Wilcoxon rank-sum test fails to reject the equality of the distribution of experience in the control and incentive groups (p=.354) and fails to reject the equality of the distribution of experience in the control and wage groups (p=.719). Sample used: all applicants.

Table A20: Non-cognitive ability, ipsatised

	Mean	Percentile				
		90th	75th	50th	25th	10th
	(1)	(2)	(3)	(4)	(5)	(6)
Incentive	-0.113	-0.200	-0.213	0.024	-0.075	-0.174
	(0.118)	(0.141)	(0.126)	(0.154)	(0.178)	(0.226)
High Wage	0.090	-0.082	-0.092	0.180	0.166	0.228
	(0.112)	(0.138)	(0.120)	(0.142)	(0.172)	(0.224)
Control value	-0.00	2.68	1.60	0.15	-1.28	-2.83
Incentive = Wage (p)	0.050	0.325	0.278	0.221	0.166	0.051
Obs.	2373	2373	2373	2373	2373	2373

Notes: Estimates from OLS (Column 1) and quantile (Columns 2-6) regressions. The dependent variable is an index of ipsatised non-cognitive ability. The index is based on ipsatised values of conscientiousness, neuroticism and grit. Variables are ipsatised by subtracting the individual acquiescence score — the mean of positive items and inverted items across scales — which is a measure of the tendency to agree with any statement (Laajaj and Macours, 2017). The second to last row reports the p-value of a test of the null hypothesis that the treatments have the same effect. Robust standard errors reported in parenthesis. A Wilcoxon rank-sum test fails to reject the equality of the distribution of ipsatised non-cognitive ability in the control and incentive groups (p=.354) and fails to reject the equality of the distribution of ipsatised non-cognitive ability in the control and wage groups (p=.466). Sample used: all applicants.

Table A21: Job search outcomes in 30 days after first phone call

	Applications	Money (USD)	Time	Interviews	Offers	Has job
	(1)	(2)	(3)	(4)	(5)	(6)
Incentive	-0.009	0.043	28.295	-0.014	-0.017	-0.004
	(0.080)	(0.244)	(28.017)	(0.027)	(0.014)	(0.008)
High Wage	-0.072	0.037	-10.242	-0.038	-0.027	-0.019
	(0.073)	(0.240)	(26.027)	(0.025)	(0.013)	(0.007)
Control group mean	1.573	2.742	392.509	.309	.103	.048
Incentive = Wage (p)	0.401	0.982	0.152	0.360	0.437	0.049
Obs.	4328	4328	4328	4328	4328	4328

Notes: OLS regression. The dependent variable is indicated in the column heading. All dependent variables are measured in the second phone call (see Figure A2 for a timeline). Further, all dependent variables refer to jobs other than the experiment's job. These variables are collected through a short application roster where the respondent is asked a number of questions about each application they have made in the 30 days between the two phone calls. This includes information about the application process and its outcome. The variables 'applications', 'interviews' and 'offers' capture the total number of applications and interviews made, and offers received. 'Has job' is a dummy variable capturing whether the respondent is currently employed in one of the jobs they have applied for in the period between the two phone calls. 'Cost' and 'time' are, respectively, the total amount of money and time that the respondent reports to have spent on all job applications they have made in the 30 days period. The second to last row reports the p-value of a test of the null hypothesis that the treatments have the same effect. Robust standard errors reported in parenthesis. Sample used: second phone call sample.

Table A22: Additional job search outcomes in 30 days after first phone call (quality of applications)

	Number of applications				
	Total (incl. exp. job)	Occup. matched	Skills matched	Long run	Permanent
	(1)	(2)	(3)	(4)	(5)
Incentive	0.114	0.019	-0.028	-0.033	0.011
	(0.083)	(0.061)	(0.034)	(0.043)	(0.061)
High Wage	0.118	-0.033	-0.063	0.008	-0.029
	(0.077)	(0.061)	(0.034)	(0.044)	(0.060)
Control group mean	1.983	1.299	0.357	0.502	1.309
Incentive = Wage (p)	0.962	0.390	0.303	0.347	0.510
Obs.	4328	4328	4328	4328	4328

Notes: OLS regression. The dependent variable is indicated in the column heading. All dependent variables are measured in the second phone call (see Figure A2 for a timeline). 'Applications' is the total number of applications made. This variable includes the application to the experiment's job and thus differs from the variable 'Applications' reported in Table A21. On the other hand, the variables reported in columns (2)-(5) do not include the application to the experiment's job. For each application made to these other jobs, respondents are asked a number of questions about the position: the occupation of the job, whether they feel they have the right skills for the job, a rating from 0 to 10 indicating whether they see themselves doing that particular job in the long run (we create a dummy that splits this variable at the median), and whether the job has an open-ended contract. We use these responses to construct three variables: (2) 'Occup. matched' is the number of applications to positions that match the occupation the jobseeker would like to find, (3) 'Skills matched' is the number of applications to positions that match the skills of the jobseeker (i.e. the jobseeker does not feel overqualified for the position) (4) 'Long run' is the number of applications to positions that the jobseeker sees herself doing in the long run and (5) 'Permanent' is the number of applications to positions that offer an open-ended contract. The second to last row reports the p-value of a test of the null hypothesis that the treatments have the same effect. Robust standard errors reported in parenthesis. Sample used: second phone call sample.

Table A23: Additional job search outcomes in 30 days after first phone call (quality of interviews and jobs)

	Interviews for jobs that are...			Job is...		
	Matched	Long run	Permanent	Matched	Long run	Permanent
	(1)	(2)	(3)	(4)	(5)	(6)
Incentive	-0.015	-0.026	-0.004	-0.005	-0.007	-0.006
	(0.025)	(0.016)	(0.025)	(0.008)	(0.005)	(0.007)
High Wage	-0.039	-0.023	-0.027	-0.018	-0.013	-0.020
	(0.024)	(0.017)	(0.024)	(0.008)	(0.006)	(0.007)
Control group mean	0.273	0.129	0.261	0.045	0.024	0.043
Incentive = Wage (p)	0.305	0.853	0.338	0.070	0.243	0.049
Obs.	4328	4328	4328	4328	4328	4328

Notes: OLS regression. The dependent variable is indicated in the column heading. All dependent variables are measured in the second phone call (see Figure A2 for a timeline). The first set of dependent variables are defined as the number of interviews for jobs that: (1) match the occupation the jobseeker would like to find, (2) the jobseeker sees herself doing in the long run and (3) offer an open-ended contract. The second set of dependent variables are defined as a dummy variable for working in a job that (4) matches the occupation the jobseeker would like to find, (5) the respondent sees herself doing in the long run and (6) offers an open-ended contract. All variables exclude the application to the experiment's job. The second to last row reports the p-value of a test of the null hypothesis that the treatments have the same effect. Robust standard errors reported in parenthesis. Sample used: second phone call sample.

Table A24: Search method

	Board	Newspaper	Direct	Network
	(1)	(2)	(3)	(4)
Incentive	0.037	-0.032	-0.002	-0.054
	(0.306)	(0.162)	(0.064)	(0.199)
High Wage	-0.624	-0.152	-0.049	-0.064
	(0.294)	(0.162)	(0.061)	(0.235)
Control Mean	6.326	3.895	0.567	2.272
St. dev.	8.077	4.365	1.725	6.016
Incentive = Wage (p)	0.028	0.460	0.436	0.963
Obs.	4357	4366	4370	4349

Notes: OLS regressions. The dependent variable is reported in the column headings. 'Board' is the number of days in the last 30 days when the respondent visited the job vacancy board. 'Newspaper' is the number of times in the last 30 days when the respondent consulted the job insert in the newspaper. 'Direct' is the number of days in the last 30 days when the respondent visited employers to inquire about vacancies. 'Network' is the number of social contacts that the person has talked to about job opportunities in the last 30 days. The second to last row reports the p-value of a test of the null hypothesis that the treatments have the same effect. Robust standard errors reported in parenthesis. Sample used: all applicants.

Table A25: Heterogeneous impacts on applications

Heterogeneity by	Gender	Experience	Unemployed	Length unemployment	Age	Value job
	(1)	(2)	(3)	(4)	(5)	(6)
Impacts for	Female	Low experience	Unemployed	Long spell	Young	High value
Incentive	0.073	0.092	0.097	0.108	0.101	0.091
	(0.039)	(0.025)	(0.021)	(0.033)	(0.027)	(0.029)
High wage	0.190	0.217	0.214	0.242	0.196	0.206
	(0.038)	(0.024)	(0.021)	(0.032)	(0.027)	(0.026)
Control mean	0.429	0.492	0.473	0.464	0.464	0.504
Incentive = Wage (p)	0.002	0.000	0.000	0.000	0.000	0.000
Impacts for	Male	High experience	Employed	Short spell	Old	Low value
Incentive	0.125	0.136	0.158	0.081	0.121	0.137
	(0.020)	(0.025)	(0.030)	(0.030)	(0.023)	(0.026)
High wage	0.185	0.156	0.147	0.205	0.179	0.171
	(0.020)	(0.025)	(0.030)	(0.029)	(0.023)	(0.025)
Control mean	0.407	0.329	0.270	0.482	0.372	0.318
Incentive = Wage (p)	0.003	0.436	0.704	0.000	0.015	0.179
No het. incentive (p)	0.235	0.209	0.099	0.537	0.567	0.186
No het. wage (p)	0.891	0.080	0.065	0.383	0.631	0.302
Obs.	4689	4686	4689	3020	4686	4686

Notes: OLS regressions. The dependent variable is a dummy capturing whether the jobseeker applied for the experiment's job. Column headings indicate the dimension of heterogeneity studied. For each treatment, the last panel reports the p-value of a test of the null hypothesis that the effect of treatment is not heterogenous across the dimension under study. Robust standard errors reported in parenthesis. The standard errors reported in column (6) are bootstrapped to reflect the uncertainty in the estimation of the present value of the job. Sample used: baseline sample.

Table A26: Heterogeneous impacts on other job search

Heterogeneity by	Gender	Experience	Unemployed	Length unemployment	Age	Value job
	(1)	(2)	(3)	(4)	(5)	(6)
Impacts for	Female	Low experience	Unemployed	Long spell	Young	High value
Incentive	-0.014	0.065	0.021	0.068	-0.010	0.013
	(0.142)	(0.120)	(0.105)	(0.166)	(0.120)	(0.142)
High wage	-0.128	-0.065	-0.058	0.021	-0.174	-0.133
	(0.137)	(0.097)	(0.092)	(0.129)	(0.117)	(0.113)
Control mean	1.407	1.811	1.779	1.472	1.780	1.821
Incentive = Wage (p)	0.383	0.259	0.405	0.762	0.110	0.206
Impacts for	Male	High experience	Employed	Short spell	Old	Low value
Incentive	0.005	-0.073	-0.027	0.015	0.002	-0.012
	(0.094)	(0.102)	(0.100)	(0.142)	(0.107)	(0.095)
High wage	-0.083	-0.129	-0.107	-0.060	-0.025	-0.040
	(0.084)	(0.105)	(0.101)	(0.136)	(0.089)	(0.097)
Control mean	1.617	1.322	1.093	2.046	1.410	1.312
Incentive = Wage (p)	0.300	0.500	0.398	0.547	0.792	0.742
No het. incentive (p)	0.908	0.382	0.738	0.805	0.940	0.873
No het. wage (p)	0.781	0.653	0.717	0.667	0.315	0.515
Obs.	4328	4325	4328	2804	4325	4325

Notes: OLS regressions. The dependent variable is the number of applications to jobs other than the experiment's job. Column headings indicate the dimension of heterogeneity studied. For each treatment, the last panel reports the p-value of a test of the null hypothesis that the effect of treatment is not heterogenous across the dimension under study. Robust standard errors reported in parenthesis. The standard errors reported in column (6) are bootstrapped to reflect the uncertainty in the estimation of the present value of the job. Sample used: second phone call sample.

Table A27: Heterogeneous impacts on cognitive ability

Heterogeneity by	Gender	Experience	Unemployed	Length unemployment	Age	Value job
	(1)	(2)	(3)	(4)	(5)	(6)
Impacts for	Female	Low experience	Unemployed	Long spell	Young	High value
Incentive	1.153	0.428	0.377	0.548	0.411	0.407
	(0.270)	(0.155)	(0.129)	(0.218)	(0.170)	(0.155)
High wage	0.447	0.354	0.272	0.377	0.244	0.297
	(0.272)	(0.153)	(0.127)	(0.216)	(0.171)	(0.153)
Control mean	-0.317	-0.080	0.011	-0.340	0.030	0.006
Incentive = Wage (p)	0.000	0.563	0.344	0.309	0.242	0.389
Impacts for	Male	High experience	Employed	Short spell	Old	Low value
Incentive	0.008	-0.006	-0.116	0.217	0.096	0.033
	(0.121)	(0.155)	(0.215)	(0.165)	(0.148)	(0.161)
High wage	0.123	-0.044	-0.064	0.215	0.152	0.051
	(0.119)	(0.152)	(0.214)	(0.156)	(0.142)	(0.162)
Control mean	0.089	0.124	-0.044	0.281	-0.028	-0.010
Incentive = Wage (p)	0.296	0.798	0.794	0.990	0.672	0.903
No het. incentive (p)	0.000	0.048	0.049	0.226	0.162	0.091
No het. wage (p)	0.275	0.065	0.176	0.545	0.679	0.259
Obs.	2386	2385	2386	1738	2384	2384

Notes: OLS regressions. The dependent variable is the cognitive ability index. Column headings indicate the dimension of heterogeneity studied. For each treatment, the last panel reports the p-value of a test of the null hypothesis that the effect of treatment is not heterogenous across the dimension under study. Robust standard errors reported in parenthesis. The standard errors reported in column (6) are bootstrapped to reflect the uncertainty in the estimation of the present value of the job. Sample used: all applicants.

Table A28: Heterogeneous impacts on Raven score

Heterogeneity by	Gender	Experience	Unemployed	Length unemployment	Age	Value job
	(1)	(2)	(3)	(4)	(5)	(6)
Impacts for	Female	Low experience	Unemployed	Long spell	Young	High value
Incentive	4.791	2.291	1.894	1.666	2.758	2.104
	(1.308)	(0.820)	(0.700)	(1.079)	(0.900)	(0.826)
High wage	0.966	1.311	0.932	0.141	1.387	0.639
	(1.351)	(0.799)	(0.681)	(1.036)	(0.908)	(0.824)
Control mean	37.179	38.765	38.912	38.281	38.676	39.602
Incentive = Wage (p)	0.001	0.164	0.119	0.107	0.085	0.036
Impacts for	Male	High experience	Employed	Short spell	Old	Low value
Incentive	0.188	-0.276	-0.621	1.778	-0.315	0.154
	(0.697)	(0.920)	(1.271)	(0.960)	(0.842)	(0.961)
High wage	0.490	-0.422	-0.335	1.377	-0.097	0.741
	(0.678)	(0.926)	(1.313)	(0.928)	(0.814)	(0.975)
Control mean	38.990	38.327	37.307	39.582	38.515	36.971
Incentive = Wage (p)	0.633	0.868	0.809	0.639	0.775	0.499
No het. incentive (p)	0.002	0.037	0.083	0.938	0.013	0.117
No het. wage (p)	0.753	0.157	0.392	0.374	0.224	0.934
Obs.	2397	2396	2397	1743	2395	2395

Notes: OLS regressions. The dependent variable is the Raven test score. Column headings indicate the dimension of heterogeneity studied. For each treatment, the last panel reports the p-value of a test of the null hypothesis that the effect of treatment is not heterogenous across the dimension under study. Robust standard errors reported in parenthesis. The standard errors reported in column (6) are bootstrapped to reflect the uncertainty in the estimation of the present value of the job. Sample used: all applicants.

Table A29: Heterogeneous impacts on GPA

Heterogeneity by	Gender	Experience	Unemployed	Length unemployment	Age	Value job
	(1)	(2)	(3)	(4)	(5)	(6)
Impacts for	Female	Low experience	Unemployed	Long spell	Young	High value
Incentive	0.129	0.078	0.067	0.083	0.064	0.075
	(0.053)	(0.029)	(0.027)	(0.041)	(0.034)	(0.034)
High wage	0.081	0.067	0.036	0.061	0.021	0.049
	(0.050)	(0.029)	(0.027)	(0.040)	(0.034)	(0.033)
Control mean	2.797	2.987	2.973	2.897	3.027	3.003
Incentive = Wage (p)	0.360	0.672	0.224	0.546	0.163	0.349
Impacts for	Male	High experience	Employed	Short spell	Old	Low value
Incentive	0.024	0.009	-0.010	0.047	0.035	0.016
	(0.027)	(0.043)	(0.055)	(0.038)	(0.035)	(0.044)
High wage	-0.008	-0.073	-0.075	0.018	0.004	-0.040
	(0.026)	(0.041)	(0.054)	(0.036)	(0.034)	(0.040)
Control mean	3.013	2.949	2.959	3.019	2.925	2.933
Incentive = Wage (p)	0.179	0.028	0.161	0.406	0.315	0.124
No het. incentive (p)	0.079	0.183	0.212	0.514	0.556	0.254
No het. wage (p)	0.118	0.005	0.063	0.426	0.729	0.074
Obs.	2285	2284	2285	1670	2283	2283

Notes: OLS regressions. The dependent variable is the applicant's GPA. Column headings indicate the dimension of heterogeneity studied. For each treatment, the last panel reports the p-value of a test of the null hypothesis that the effect of treatment is not heterogenous across the dimension under study. Robust standard errors reported in parenthesis. The standard errors reported in column (6) are bootstrapped to reflect the uncertainty in the estimation of the present value of the job. Sample used: all applicants.

Table A30: Heterogeneous impacts on top applicants (75th percentile)

Heterogeneity by	Gender	Experience	Unemployed	Length unemployment	Age	Value job
	(1)	(2)	(3)	(4)	(5)	(6)
Impacts for	Female	Low experience	Unemployed	Long spell	Young	High value
Incentive	0.240	0.081	0.073	0.077	0.094	0.075
	(0.052)	(0.031)	(0.027)	(0.040)	(0.035)	(0.039)
High wage	0.097	0.061	0.057	0.060	0.063	0.047
	(0.047)	(0.030)	(0.026)	(0.038)	(0.034)	(0.036)
Control top applicants	0.209	0.259	0.259	0.216	0.270	0.276
Incentive = Wage (p)	0.004	0.494	0.535	0.653	0.352	0.345
Impacts for	Male	High experience	Employed	Short spell	Old	Low value
Incentive	0.003	0.018	0.005	0.065	0.014	0.029
	(0.026)	(0.035)	(0.046)	(0.039)	(0.031)	(0.038)
High wage	0.039	0.040	0.041	0.057	0.044	0.065
	(0.026)	(0.035)	(0.047)	(0.037)	(0.031)	(0.038)
Control top applicants	0.260	0.233	0.206	0.295	0.229	0.206
Incentive = Wage (p)	0.143	0.503	0.378	0.828	0.311	0.256
No het. incentive (p)	0.000	0.188	0.205	0.834	0.091	0.320
No het. wage (p)	0.280	0.662	0.771	0.956	0.671	0.694
Obs.	2386	2385	2386	1738	2384	2384

Notes: OLS regressions. The dependent variable is a dummy capturing whether the jobseeker scored above the 75th percentile of the control distribution of the cognitive ability index. Column headings indicate the dimension of heterogeneity studied. For each treatment, the last panel reports the p-value of a test of the null hypothesis that the effect of treatment is not heterogenous across the dimension under study. Robust standard errors reported in parenthesis. The standard errors reported in column (6) are bootstrapped to reflect the uncertainty in the estimation of the present value of the job. Sample used: all applicants.

Table A31: Heterogeneous impacts on quality at the top of the distribution

Heterogeneity by	Gender		Experience		Unemployed		Length unemployment		Age		Value job	
	p90	p75	p90	p75	p90	p75	p90	p75	p90	p75	p90	p75
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Impacts for	Female		Low experience		Unemployed		Long spell		Young		High value	
Incentive	0.582	.848	0.244	.176	0.227	.185	0.259	.29	0.131	.131	0.191	.118
	(0.241)	(0.230)	(0.136)	(0.155)	(0.122)	(0.134)	(0.188)	(0.204)	(0.161)	(0.168)	(0.147)	(0.159)
High wage	0.247	.257	0.165	.136	0.136	.139	0.146	.257	0.164	.103	0.117	.035
	(0.246)	(0.225)	(0.129)	(0.153)	(0.120)	(0.129)	(0.205)	(0.178)	(0.168)	(0.175)	(0.144)	(0.149)
Control quality	-0.317	2.267	-0.080	2.404	0.011	2.421	-0.340	2.223	0.030	2.535	0.006	2.442
Incentive = Wage (p)	0.074	0.002	0.534	0.756	0.416	0.682	0.519	0.870	0.821	0.835	0.563	0.495
Impacts for	Male		High experience		Employed		Short spell		Old		Low value	
Incentive	0.081	-.069	0.252	.124	0.283	.033	0.104	.085	0.266	.086	0.337	.124
	(0.126)	(0.131)	(0.189)	(0.168)	(0.217)	(0.258)	(0.208)	(0.179)	(0.160)	(0.150)	(0.152)	(0.180)
High wage	0.175	.126	0.098	.262	0.167	.189	0.045	.118	0.143	.244	0.269	.333
	(0.121)	(0.130)	(0.185)	(0.161)	(0.222)	(0.252)	(0.192)	(0.173)	(0.152)	(0.139)	(0.175)	(0.169)
Control quality	0.089	2.339	0.124	2.211	-0.044	2.059	0.281	2.654	-0.028	2.211	-0.010	2.1
Incentive = Wage (p)	0.420	0.075	0.320	0.358	0.549	0.513	0.738	0.822	0.408	0.260	0.658	0.170
No het. incentive (p)	0.066	0.001	0.973	0.821	0.821	0.602	0.580	0.449	0.553	0.842	0.530	0.977
No het. wage (p)	0.793	0.614	0.768	0.572	0.903	0.860	0.719	0.576	0.927	0.528	0.496	0.197
Obs.	2386	2386	2385	2385	2386	2386	1738	1738	2384	2384	2384	2384

Notes: Quantile regressions. The dependent variable is the cognitive ability score at a given percentile of the control distribution (90th or 75th). Column headings indicate the dimension of heterogeneity studied. For each treatment, the last panel reports the p-value of a test of the null hypothesis that the effect of treatment is not heterogenous across the dimension under study. Robust standard errors reported in parenthesis. The standard errors reported in column (6) are bootstrapped to reflect the uncertainty in the estimation of the present value of the job. Sample used: all applicants.

Table A32: Decomposition of impact on cognitive ability

Heterogeneity by	Gender	Experience	Unemployed	Length unemployment	Age	Value job
	(1)	(2)	(3)	(4)	(5)	(6)
Compositional effect	0.03	0.04	-0.01	-0.06	-0.00	-0.00
Within-group effect for	Female	Low experience	Unemployed	Long spell	Young	High value
	0.95	0.97	1.13	0.73	0.79	0.94
Within-group effect for	Male	High experience	Employed	Short spell	Old	Low value
	0.03	-0.01	-0.12	0.33	0.21	0.06

Notes: Decomposition of the treatment effect reported in Table A27. We implement this decomposition as follows. There are six dimensions of heterogeneity, namely gender, experience, unemployment status, unemployment length, age, and job value. Each of these dimensions is split in two categories (male/female, etc.). We denote these with a vector of dummy variables $v_i \in \{0, 1\}$ for all dimensions $i = \{1, 2, ..., 6\}$. Further, we use $j \in \{0, 1\}$ to indicate the experimental group ($j = 0$ refers to the control group and $j = 1$ to the incentive group). Finally, we use p_i^j to indicate the share of applicants in group j for whom $v_i = 1$. For each dimension of heterogeneity i, we decompose the total effect on expected ability T into three components: (i) a compositional effect: $(p_i^1 - p_i^0) * (E[T|v_i = 1, j = 0] - E[T|v_i = 0, j = 0])$; (ii) a within-group effect for the first group of applicants: $p_i^1 * (E[T|v_i = 1, j = 1] - E[T|v_i = 1, j = 0])$; (iii) a within-group effect for the second group of applicants: $(1 - p_i^1) * (E[T|v_i = 0, j = 1] - E[T|v_i = 0, j = 0])$.

Table A33: Additional heterogeneity: first language

	Apply	Cognitive	Raven	GPA	No. other applications
	(1)	(2)	(3)	(4)	(5)
Impacts for			First language Amharic		
Incentive	0.123	0.127	0.708	0.087	0.075
	(0.021)	(0.129)	(0.734)	(0.030)	(0.101)
High Wage	0.210	0.138	0.251	0.024	-0.091
	(0.021)	(0.126)	(0.733)	(0.029)	(0.086)
Control Mean	0.390	0.384	40.140	2.908	1.618
Incentive = Wage (p)	0.000	0.920	0.474	0.019	0.067
Impacts for			First language not Amharic		
Incentive	0.098	0.360	1.527	-0.019	-0.188
	(0.032)	(0.202)	(1.101)	(0.043)	(0.125)
High Wage	0.136	0.196	0.811	-0.005	-0.099
	(0.031)	(0.200)	(1.045)	(0.041)	(0.130)
Control Mean	0.457	-0.680	35.861	3.006	1.482
Incentive = Wage (p)	0.231	0.372	0.492	0.733	0.437
No het. incentive (p)	0.524	0.330	0.536	0.044	0.101
No het. wage (p)	0.049	0.805	0.661	0.558	0.964
Obs.	4689	2386	2397	2285	4328

Notes: OLS regressions. The dependent variable is reported in the column headings. Robust standard errors reported in parenthesis. Sample used: baseline sample (column 1), all applicants (columns 2-4), second phone call sample (column 5).

Table A34: Time preferences and cost of effort

	β	δ	γ
	(1)	(2)	(3)
Control	0.817	0.740	5.842
	(0.083)	(0.090)	(0.784)
Incentive	0.794	0.891	5.032
	(0.057)	(0.064)	(0.500)
High Wage	0.811	0.971	5.233
	(0.053)	(0.070)	(0.511)
Incentive - Control (p)	0.825	0.420	0.384
Incentive - Wage (p)	0.833	0.400	0.778
Wage - Control (p)	0.953	0.044	0.515

Notes: Structural estimates of present bias (β), impatience (δ) and cost of effort (γ). The estimation technique is described in detail in Appendix C. Standard errors obtained with the delta method reported in parenthesis. The last three rows report the p-values of tests of the equality of the coefficients. Sample used: all applicants.

Table A35: Preferences and sophistication

	Present bias	Risk Preferences	Social Preferences	Sophistication
	(1)	(2)	(3)	
Incentive	0.022	0.054	-0.033	0.007
	(0.027)	(0.066)	(0.097)	(0.021)
High Wage	0.005	0.030	0.042	0.001
	(0.025)	(0.064)	(0.094)	(0.020)
Variable	β_i	Index	Index	Dummy
Control group mean	0.301	1.905	0.040	0.190
Control group st.dev.	0.459	1.170	1.762	0.393
Incentive = Wage (p)	0.466	0.684	0.397	0.751
Obs.	2053	2110	2193	2331

Notes: Estimates from OLS regressions. Robust standard errors reported in parenthesis. Present bias is a dummy for individuals with $\beta < 0.99$. Sophistication is a dummy for individuals with $k \geq 2$. The tasks used to elicit these variables are described in detail in Appendix C. Sample size changes because of missing responses and because we are not able to estimate a β coefficient for all choice patterns. Sample used: all applicants.

Table A36: Test effort

	Mistake		Unfinished	
	(1)	(2)	(3)	(4)
Incentive	0.096	0.033	-0.019	-0.005
	(0.082)	(0.025)	(0.038)	(0.007)
High Wage	0.067	0.021	-0.028	-0.002
	(0.078)	(0.024)	(0.034)	(0.007)
Measure	continuous	dummy	continuous	dummy
Control group mean	0.711	0.292	0.081	0.018
Incentive = Wage (p)	0.712	0.570	0.782	0.701
Obs.	2316	2316	2332	2332

Notes: OLS regression. In column (1), the dependent variable 'Mistake' is the number of strings transcribed incorrectly. In column (2), the dependent variable is a dummy capturing whether any string was transcribed incorrectly. In column (3), the dependent variable 'Unfinished' is the number of strings that the applicant has failed to transcribe. In column (4), the dependent variable is a dummy capturing whether the applicant has failed to transcribe any string. Robust standard errors reported in parenthesis. The second to last row reports the p-value of a test of the null hypothesis that the treatments have the same effect. Sample used: all applicants.

Table A37: Salience of the position

	Correct answer		Absolute mistake	
	(1)	(2)	(3)	(4)
Incentive	0.026	0.025	-29.578	-26.355
	(0.017)	(0.017)	(20.603)	(20.959)
High Wage	0.039	0.037	-44.615	-38.755
	(0.017)	(0.017)	(20.034)	(21.187)
Control for application	no	yes	no	yes
Control group mean	0.686	0.686	167.303	167.303
Incentive = Wage (p)	0.476	0.497	0.399	0.493
Obs.	4375	4375	3634	3634

Notes: OLS regression. The dependent variable is indicated in the column headings. 'Correct answer' is a dummy capturing whether the respondent recalled the wage offered correctly. 'Absolute mistake' is the absolute difference between the wage recalled by the respondent and the wage actually offered. The number of observation changes because some individuals report that they do not remember the wage offered. These individuals are included in the regressions reported in columns (1) and (2), but not in the regressions reported in columns (3) and (4). Columns (2) and (4) include a control for whether the respondent has applied for the experiment's job. Robust standard errors reported in parenthesis. The second to last row reports the p-value of a test of the null hypothesis that the treatments have the same effect. Sample used: second phone call sample.

Table A38: Beliefs about labour-market prospects

	Weeks unemployment	Wage (Ethiopian Birr)
	(1)	(2)
Incentive	-0.036	95.746
	(0.546)	(196.659)
High Wage	-0.635	449.245
	(0.602)	(145.498)
Control group mean	8.690	5192.291
Incentive = Wage (p)	0.293	0.076
Obs.	3849	3817

Notes: OLS regression. 'Weeks unemployment' captures the number of weeks that the respondent expect she or he would need in order to be offered a job they would be willing to work at. 'Wage' captures the wage that the respondent expects this job will pay. Beliefs about the wage are elicited through the method of Attanasio and Kaufmann (2009), as explained in footnote **??**. For both questions the respondent was asked to consider an hypothetical job search spell starting on the day following their interview. Thus, the answers to these questions do not refer to the experiment's job. The second to last row reports the p-value of a test of the null hypothesis that the treatments have the same effect. Robust standard errors reported in parenthesis. Sample used: second phone call sample.

Table A39: Beliefs about the probability of getting the experiment's job

	Subjective forecast		
	(1)	(2)	(3)
Incentive	1.130	0.619	0.105
	(1.119)	(1.095)	(1.071)
High Wage	1.461	1.453	0.648
	(1.112)	(1.090)	(1.070)
Forecasts implying certainty	included	excluded	excluded
Truncation	no	no	yes
Control group mean	57.065	53.083	48.544
Incentive = Wage (p)	0.769	0.450	0.617
Obs.	4325	3893	3446

Notes: OLS regression. The dependent variable is the subject's subjective forecast of the probability of being offered the experiment's job. We elicit this forecast retrospectively, by asking the following question in the second phone call: "How confident were you of getting an offer for this position at the time when you decided whether to apply or not? In order to quantify this, you can think of applying to 100 positions like this one. How many offers would you get?". In the first column, we report the raw data. In the second and third column, we drop forecasts that imply certainty, that is, forecasts of 0 or 100. In the third column, we truncate the variable at the 95th percentile. The second to last row reports the p-value of a test of the null hypothesis that the treatments have the same effect. Robust standard errors reported in parenthesis. Sample used: second phone call sample.

Table A40: Beliefs about the attributes of the job

	Holidays	Overtime	Satisfaction	Autonomy	Career	Opportunities	New Skills
	(1)	(2)	(3)	(4)	(5)	(6)	(8)
Incentive	0.025	0.019	0.006	-0.013	0.034	0.006	-0.005
	(0.013)	(0.018)	(0.011)	(0.019)	(0.014)	(0.011)	(0.005)
High Wage	0.017	0.046	0.028	-0.006	0.043	0.004	0.001
	(0.012)	(0.018)	(0.010)	(0.019)	(0.014)	(0.011)	(0.004)
Control group mean	0.127	0.411	0.904	0.486	0.810	0.908	0.986
Incentive = Wage (p)	0.529	0.151	0.034	0.692	0.506	0.861	0.158
Obs.	4366	4362	4364	4361	4363	4364	4368

Notes: OLS regression. The dependent variable is indicated in the column headings. 'Holiday' is a dummy variable capturing whether the respondent believes the job has more than four days of holiday per month. 'Overtime' is a dummy variable capturing whether the respondent believes the job will require work in the evenings. 'Satisfaction' is a dummy variable capturing whether the respondent believes the job will be satisfying. 'Autonomy' is a dummy variable capturing whether the respondent believes he or she will have freedom to organise their own schedule at work. 'Career' is a dummy variable capturing whether the respondent believes the experience in this job will help them find other jobs in the future. 'Opportunity' is a dummy variable capturing whether the respondent believes there will be further work opportunities with the employer. 'New Skills' is a dummy variable capturing whether the respondent believes they will learn new skills in this job. The second to last row reports the p-value of a test of the null hypothesis that the treatments have the same effect. Robust standard errors reported in parenthesis. Sample used: second phone call sample.

Table A41: Survey experiment on job-attribute beliefs: Balance

	Mean		StDev	N	Balance test (p)
	Control	Incentive			
	(1)	(2)	(3)	(4)	(5)
Female	0.49	0.48	0.50	724	0.66
Age	23.95	24.13	2.21	724	0.28
Born in Addis Ababa	0.20	0.17	0.40	724	0.22
Wage work experience (dummy)	0.40	0.43	0.49	724	0.45
Wage work experience (months)	20.98	20.90	23.04	235	0.98
Currently unemployed	0.77	0.82	0.42	724	0.08
Overall balance					0.22

Notes: In this Table, we present summary and balance statistics for the sample of individuals that participated in the new survey fielded in 2019/2020. We present summary statistics in columns 1-4. In column 5, we report the p-value of a test of covariate balance. We first report balance tests for single covariates and then, in the last row, report a joint test of orthogonality (following the recent literature, e.g. McKenzie (2017)). To perform the joint test of orthogonality we regress the treatment variable on all covariates and we then test the joint hypothesis that all covariates have a zero coefficient. Sample used: 2020 survey sample.

Table A42: Survey experiment on job-attribute beliefs: Results

	Expectations					
	Holidays	Overtime	Satisfaction	Autonomy	Career	Enjoyable
	(1)	(2)	(3)	(4)	(5)	(6)
Incentive	-0.019	0.011	-0.017	-0.017	0.033	-0.003
	(0.037)	(0.034)	(0.037)	(0.029)	(0.034)	(0.037)
Control mean	0.461	0.279	0.459	0.202	0.691	0.544
Obs.	724	724	724	724	724	724

Notes: OLS regression. The second to last row reports the p-value of a test of the null hypothesis that the treatments have the same effect. 'Holiday' is a dummy variable capturing whether the respondent believes the job has more than four days of holiday per month. 'Overtime' is a dummy variable capturing whether the respondent believes the job will require work in the evenings. 'Satisfaction' is a dummy variable capturing whether the respondent believes the job will be satisfying. 'Autonomy' is a dummy variable capturing whether the respondent believes he or she will have freedom to organise their own schedule at work. 'Career' is a dummy variable capturing whether the respondent believes the experience in this job will help them find other jobs in the future. 'Enjoyable' is a dummy variable capturing whether the respondent believes the work environment will be pleasant. Robust standard errors reported in parenthesis. Sample used: 2020 survey sample.

Table A43: Impacts on individual psychological traits

	Conscientiousness	Neuroticism	Openness	Extraversion	Agreeableness	Grit	Locus of control	Core self-evaluaton	Self-esteem
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Incentive	0.005	0.020	0.020	-0.011	0.041	-0.033	-0.114	-0.152	-0.040
	(0.026)	(0.033)	(0.022)	(0.025)	(0.024)	(0.025)	(0.518)	(0.185)	(0.028)
High Wage	0.032	-0.032	0.030	0.018	0.041	0.029	-0.816	-0.010	0.019
	(0.024)	(0.031)	(0.021)	(0.024)	(0.023)	(0.024)	(0.490)	(0.183)	(0.026)
Control Mean	4.539	1.762	4.115	3.467	4.244	3.862	35.309	23.431	4.121
St. dev.	0.488	0.620	0.435	0.464	0.481	0.475	9.817	3.545	0.520
Incentive = Wage (p)	0.234	0.076	0.573	0.193	0.995	0.006	0.121	0.396	0.015
Obs.	2385	2385	2385	2385	2385	2373	2378	2377	2363

Notes: OLS regressions. The dependent variable is reported in the column headings. The first five columns report impacts on each of the big five personality traits (John and Srivastava, 1999). Columns (6)-(9) report impacts on grit (Duckworth et al., 2007), core self evaluation (Gardner and Pierce, 2010), locus of control scale (Lefcourt, 1991) and self-esteem (Rosenberg, 1986). The second to last row reports the p-value of a test of the null hypothesis that the treatments have the same effect. Robust standard errors reported in parenthesis. Sample used: all applicants.

Table A44: Impacts on specific work experience

	Routine	Physical	Managerial	Problem-solving	Math	Reading	Client
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Incentive	-0.172	0.599	-0.816	-1.320	-1.157	0.450	0.041
	(1.865)	(1.122)	(1.173)	(1.732)	(1.585)	(1.831)	(0.026)
High Wage	-0.715	2.870	-0.159	-0.939	-0.564	-0.096	0.027
	(1.732)	(2.145)	(1.179)	(1.725)	(1.578)	(1.661)	(0.025)
Control Mean	13.029	6.480	8.566	11.328	8.534	10.904	0.347
St. dev.	36.669	19.374	24.005	37.380	34.780	35.069	0.477
Incentive = Wage (p)	0.707	0.291	0.485	0.748	0.563	0.706	0.558
Obs.	2311	2310	2311	2311	2311	2311	2310

Notes: OLS regressions. The dependent variable is reported in the column headings. All dependent variables refer to specific dimensions of job experience, measured using the questionnaire developed by Autor and Handel (2013). The variables reported in columns (1)-(6) measure the number of months of experience in a job where the task indicated in the variable name was regularly performed. 'Routine' refers to short, repetitive tasks; 'managerial' refers to tasks such as managing and supervising other workers; 'physical' refers to physical tasks such as standing, handling objects, operating machinery or vehicles, or making or fixing things with one's hands; 'problem-solving' refers to tasks where the worker is faced with a new or difficult situation where they have to think for at least 30 minutes to find a good solution; 'math' refers to advanced mathematics such as algebra, geometry, trigonometry, probability, or calculus; 'reading' refers to tasks that require reading documents that are longer than ten pages. Column (7), 'client-service', is a dummy variable that takes a value of one if the respondent has substantial experience interacting face-to-face with clients, customers, suppliers or contractors. Robust standard errors reported in parenthesis. Sample used: all applicants.

Table A45: Identification (noisy-ability case)

Structural parameters (13)		Moments (14/16)
Quality and costs:	$\Leftrightarrow$	$\mathrm{E}[T \mid apply, B = b_z, \texttt{control}]$,
μ_{T_l}, σ_{T_l}, μ_{C_l},		$\mathrm{SD}[T \mid apply, B = b_z, \texttt{control}]$,
μ_{T_h}, σ_{T_h}, μ_{C_h}		$\mathrm{Pr}[apply \mid B = b_z, \texttt{control}]$
		for $z \in \{l, h\}$
Shocks and st. dev. of costs:	$\Leftrightarrow$	ΔApplications$[B = b_z, \texttt{incentive}]$
σ_{C_l}, σ_{C_h}, τ, τ^w		ΔApplications$[B = b_z, \texttt{wage}]$
		for $z \in \{l, h\}$
Covariance and selectivity:	$\Leftrightarrow$	ΔApplicantAbility$[B = b_z, \texttt{incentive}]$,
σ_{TC_l}, σ_{TC_h}, a		ΔApplicantAbility$[B = b_z, \texttt{wage}]$
		$\mathrm{E}[\mathrm{Pr}[T > a \mid B = b_z, C = c]]$
		for $z \in \{l, h\}$

Table A46: Identification (noisy-selection case)

Structural parameters (14)		Moments (14/16)
Quality and costs:	$\Leftrightarrow$	$\text{E}[T \vert apply, B = b_z, \texttt{control}]$,
μ_{T_l}, σ_{T_l}, μ_{C_l},		$\text{SD}[T \vert apply, B = b_z, \texttt{control}]$,
μ_{T_h}, σ_{T_h}, μ_{C_h}		$\Pr[apply \vert B = b_z, \texttt{control}]$
		for $z \in \{l, h\}$
Shocks and st. dev. of costs:	$\Leftrightarrow$	ΔApplications$[B = b_z, \texttt{incentive}]$
σ_{C_l}, σ_{C_h}, τ, τ^w		ΔApplications$[B = b_z, \texttt{wage}]$
		for $z \in \{l, h\}$
Covariance and selectivity:	$\Leftrightarrow$	ΔApplicantAbility$[B = b_z, \texttt{incentive}]$,
σ_{TC_l}, σ_{TC_h}, μ_a, σ_a		ΔApplicantAbility$[B = b_z, \texttt{wage}]$
		$\text{E}[\Pr[T > a]]$
		for $z \in \{l, h\}$

Table A47: Additional parameter estimates

	(1)	(2)	(3)	(4)
	Low B			
μ_T	45.605	45.697	45.373	45.357
	(1.11)	(0.97)	(2.76)	(3.15)
σ_T	13.785	13.818	13.674	13.667
	(0.76)	(0.73)	(1.58)	(1.75)
	High B			
μ_T	46.365	46.457	46.885	46.905
	(1.03)	(1.03)	(3.64)	(3.59)
σ_T	14.293	14.348	14.531	14.541
	(0.87)	(0.87)	(2.26)	(2.21)
Information	Noisy ability		Noisy selectivity	
Moments	14	16	14	16

Notes: The table shows the additional parameters of the marginal distribution of ability that we did not report in Table 4. We report both estimates for the noisy-ability case (columns 1 and 2) and the noisy-selection case (columns 3 and 4). Estimation is based on minimum distance estimation. Column (1) and (3) use 14 moments (reported in Table A48 and Table A50). Column (2) and (4) use 16 moments (reported in Table A49 and Table A51). Standard errors obtained through a bootstrap of the structural estimation reported in parenthesis. The bootstrap includes the estimation of B and the demediation procedure.

Table A48: Fit between empirical and simulated moments
(noisy-ability case, core moments)

Moment	Empirical	Simulated
	Low B	
Pr[$apply$|$B = b_l$, $\texttt{control}$]	47.648	47.651
E[T|$apply, B = b_l$, $\texttt{control}$]	38.225	38.234
SD[T|$apply, B = b_l$, $\texttt{control}$]	11.811	11.812
ΔApplications[$B = b_l$, $\texttt{incentive}$]	10.952	11.165
ΔApplications[$B = b_l$, $\texttt{wage}$]	14.286	13.713
ΔAbility[$B = b_l$, $\texttt{incentive}$]	1.973	1.536
ΔAbility[$B = b_l$, $\texttt{wage}$]	1.462	1.872
	High B	
Pr[$apply$|$B = b_h$, $\texttt{control}$]	49.667	49.662
E[T|$apply, B = b_h$, $\texttt{control}$]	39.812	39.803
SD[T|$apply, B = b_h$, $\texttt{control}$]	12.715	12.715
ΔApplications[$B = b_h$, $\texttt{incentive}$]	8.337	8.749
ΔApplications[$B = b_h$, $\texttt{wage}$]	12.033	11.291
ΔAbility[$B = b_h$, $\texttt{incentive}$]	2.332	1.107
ΔAbility[$B = b_h$, $\texttt{wage}$]	0.581	1.418

Notes: The table shows the empirical and simulated moments for the structural estimates reported in column (1) of Table 4 (noisy-ability case, core moments). Pr[$apply$|$B = b_l$, $\texttt{control}$ is the application rate for low-B jobseekers in the control group. E[T|$apply, B = b_l$, $\texttt{control}$] is the average ability among control applicants from the low-B group. SD[T|$apply, B = b_l$, $\texttt{control}$] is the standard deviation of ability among control applicants from the low-B group. ΔApplications[$B = b_l$, $\texttt{incentive}$] is the change in application rates generated by the incentive intervention among low-B jobseekers. ΔApplications[$B = b_l$, $\texttt{wage}$] is the change in application rates generated by the high wage intervention among low-B jobseekers. ΔAbility[$B = b_l$, $\texttt{incentive}$] is the change in average applicant ability generated by the incentive intervention among low-B jobseekers. ΔAbility[$B = b_l$, $\texttt{wage}$] is the change in average applicant ability generated by the high wage intervention among low-B jobseekers. Moments for high-B jobseekers are defined in a similar way. To generate the two groups, we first drop observations with a negative estimated value of B. We then split the remaining observations at the median value of B.

Table A49: Fit between empirical and simulated moments (noisy-ability case, core moments + beliefs)

Moment	Empirical	Simulated
Low B		
Pr[$apply \| B = b_l, \texttt{control}$]	47.648	47.626
E[$T \| apply, B = b_l, \texttt{control}$]	38.225	38.270
SD[$T \| apply, B = b_l, \texttt{control}$]	11.811	11.820
ΔApplications[$B = b_l, \texttt{incentive}$]	10.952	11.117
ΔApplications[$B = b_l, \texttt{wage}$]	14.286	13.785
ΔAbility[$B = b_l, \texttt{incentive}$]	1.973	1.540
ΔAbility[$B = b_l, \texttt{wage}$]	1.462	1.895
E[$Pr[T > a] \| C = c, B = b_l$]	0.482	0.464
High B		
Pr[$apply \| B = b_h, \texttt{control}$]	49.667	49.689
E[$T \| apply, B = b_h, \texttt{control}$]	39.812	39.762
SD[$T \| apply, B = b_h, \texttt{control}$]	12.715	12.708
ΔApplications[$B = b_h, \texttt{incentive}$]	8.337	8.811
ΔApplications[$B = b_h, \texttt{wage}$]	12.033	11.119
ΔAbility[$B = b_h, \texttt{incentive}$]	2.332	1.138
ΔAbility[$B = b_h, \texttt{wage}$]	0.581	1.426
E[$Pr[T > a] \| C = c, B = b_h$]	0.468	0.487

Notes: The table shows the empirical and simulated moments for the structural estimates reported in column (2) of Table 4 (noisy-ability case, core moments + beliefs). Pr[$apply|B = b_l, \texttt{control}$ is the application rate for low-B jobseekers in the control group. E[$T|apply, B = b_l, \texttt{control}$] is the average ability among control applicants from the low-B group. SD[$T|apply, B = b_l, \texttt{control}$] is the standard deviation of ability among control applicants from the low-B group. ΔApplications[$B = b_l, \texttt{incentive}$] is the change in application rates generated by the incentive intervention among low-B jobseekers. ΔApplications[$B = b_l, \texttt{wage}$] is the change in application rates generated by the high wage intervention among low-B jobseekers. ΔAbility[$B = b_l, \texttt{incentive}$] is the change in average applicant ability generated by the incentive intervention among low-B jobseekers. ΔAbility[$B = b_l, \texttt{wage}$] is the change in average applicant ability generated by the high wage intervention among low-B jobseekers. $E[Pr[T > a]|C = c, B = b_l]$ is the average forecast of the probability of being offered the job among low-B jobseekers. Moments for high-B jobseekers are defined in a similar way. To generate the two groups, we first drop observations with a negative estimated value of B. We then split the remaining observations at the median value of B.

Table A50: Fit between empirical and simulated moments (noisy-selection case, core moments)

Moment	Empirical	Simulated
Low B		
$\Pr[apply \mid B = b_l, \texttt{control}]$	47.648	47.647
$\mathrm{E}[T \mid apply, B = b_l, \texttt{control}]$	38.225	38.225
$\mathrm{SD}[T \mid apply, B = b_l, \texttt{control}]$	11.811	11.811
ΔApplications$[B = b_l, \texttt{incentive}]$	10.952	11.128
ΔApplications$[B = b_l, \texttt{wage}]$	14.286	13.878
ΔAbility$[B = b_l, \texttt{incentive}]$	1.973	1.484
ΔAbility$[B = b_l, \texttt{wage}]$	1.462	1.837
High B		
$\Pr[apply \mid B = b_h, \texttt{control}]$	49.667	49.666
$\mathrm{E}[T \mid apply, B = b_h, \texttt{control}]$	39.812	39.813
$\mathrm{SD}[T \mid apply, B = b_h, \texttt{control}]$	12.715	12.715
ΔApplications$[B = b_h, \texttt{incentive}]$	8.337	8.817
ΔApplications$[B = b_h, \texttt{wage}]$	12.033	11.006
ΔAbility$[B = b_h, \texttt{incentive}]$	2.332	1.202
ΔAbility$[B = b_h, \texttt{wage}]$	0.581	1.492

Notes: The table shows the empirical and simulated moments for the structural estimates reported in column (3) of Table 4 (noisy-selection case, core moments). $\Pr[apply|B = b_l, \texttt{control}$ is the application rate for low-B jobseekers in the control group. $\mathrm{E}[T|apply, B = b_l, \texttt{control}]$ is the average ability among control applicants from the low-B group. $\mathrm{SD}[T|apply, B = b_l, \texttt{control}]$ is the standard deviation of ability among control applicants from the low-B group. ΔApplications$[B = b_l, \texttt{incentive}]$ is the change in application rates generated by the incentive intervention among low-B jobseekers. ΔApplications$[B = b_l, \texttt{wage}]$ is the change in application rates generated by the high wage intervention among low-B jobseekers. ΔAbility$[B = b_l, \texttt{incentive}]$ is the change in average applicant ability generated by the incentive intervention among low-B jobseekers. ΔAbility$[B = b_l, \texttt{wage}]$ is the change in average applicant ability generated by the high wage intervention among low-B jobseekers. Moments for high-B jobseekers are defined in a similar way. To generate the two groups, we first drop observations with a negative estimated value of B. We then split the remaining observations at the median value of B.

Table A51: Fit between empirical and simulated moments (noisy-selection case, core moments + beliefs)

Moment	Empirical	Simulated
Low B		
Pr[$apply \vert B = b_l, \texttt{control}$]	47.648	47.647
E[$T \vert apply, B = b_l, \texttt{control}$]	38.225	38.225
SD[$T \vert apply, B = b_l, \texttt{control}$]	11.811	11.811
ΔApplications[$B = b_l, \texttt{incentive}$]	10.952	11.132
ΔApplications[$B = b_l, \texttt{wage}$]	14.286	13.881
ΔAbility[$B = b_l, \texttt{incentive}$]	1.973	1.482
ΔAbility[$B = b_l, \texttt{wage}$]	1.462	1.834
E[$Pr[T > a]$]	0.482	0.478
High B		
Pr[$apply \vert B = b_h, \texttt{control}$]	49.667	49.666
E[$T \vert apply, B = b_h, \texttt{control}$]	39.812	39.812
SD[$T \vert apply, B = b_h, \texttt{control}$]	12.715	12.715
ΔApplications[$B = b_h, \texttt{incentive}$]	8.337	8.812
ΔApplications[$B = b_h, \texttt{wage}$]	12.033	10.999
ΔAbility[$B = b_h, \texttt{incentive}$]	2.332	1.205
ΔAbility[$B = b_h, \texttt{wage}$]	0.581	1.496
E[$Pr[T > a]$]	0.468	0.478

Notes: The table shows the empirical and simulated moments for the structural estimates reported in column (4) of Table 4 (noisy-selection case, core moments + beliefs). Pr[$apply|B = b_l, \texttt{control}$ is the application rate for low-B jobseekers in the control group. E[$T|apply, B = b_l, \texttt{control}$] is the average ability among control applicants from the low-B group. SD[$T|apply, B = b_l, \texttt{control}$] is the standard deviation of ability among control applicants from the low-B group. ΔApplications[$B = b_l, \texttt{incentive}$] is the change in application rates generated by the incentive intervention among low-B jobseekers. ΔApplications[$B = b_l, \texttt{wage}$] is the change in application rates generated by the high wage intervention among low-B jobseekers. ΔAbility[$B = b_l, \texttt{incentive}$] is the change in average applicant ability generated by the incentive intervention among low-B jobseekers. ΔAbility[$B = b_l, \texttt{wage}$] is the change in average applicant ability generated by the high wage intervention among low-B jobseekers. $E[Pr[T > a]|B = b_l]$ is the average forecast of the probability of being offered the job among low-B jobseekers. Moments for high-B jobseekers are defined in a similar way. To generate the two groups, we first drop observations with a negative estimated value of B. We then split the remaining observations at the median value of B.

Table A52: Elasticity of simulated moments
(noisy-ability case, core moments)

	μ_T	σ_T	μ_C	σ_C	σ_{TC}	a	τ	τ^w
E$[T \mid apply, B = b_h, \texttt{control}]$	2.487	0.163	0.317	0.219	0.226	1.339	0.000	0.000
SD$[T \mid apply, B = b_h, \texttt{control}]$	0.472	1.321	0.110	0.275	0.283	0.440	0.000	0.000
Pr$[apply \mid B = b_h, \texttt{control}]$	8.246	1.005	1.929	0.346	0.364	8.048	0.000	0.000
ΔApplications$[B = b_h, \texttt{incentive}]$	3.095	2.190	0.672	5.204	2.726	6.111	1.021	0.000
ΔApplications$[B = b_h, \texttt{wage}]$	15.570	1.612	2.595	6.191	3.233	18.728	0.000	1.249
ΔAbility$[B = b_h, \texttt{incentive}]$	0.506	2.466	0.027	6.124	3.802	2.953	0.993	0.000
ΔAbility$[B = b_h, \texttt{wage}]$	12.593	1.869	1.890	7.037	4.287	15.682	0.000	1.213

Notes: The table reports the moment elasticity for the high-B group and the noisy-ability case, estimated using fourteen moments. The corresponding parameter estimates are reported in column (1) of Table 4. As in Kaboski and Townsend (2011) and Lagakos et al. (2017), we first compute all moments using the structural estimates of the parameters. We then shock by one percent the value of each parameter at a time, and compute the percent change in the simulated moments.

Table A53: Elasticity of simulated moments
(noisy-ability case, core moments + beliefs)

	μ_T	σ_T	μ_C	σ_C	σ_{TC}	a	τ	τ^w
E$[T\|apply, B = b_h, \texttt{control}]$	2.907	0.030	0.546	0.184	0.181	1.798	0.000	0.000
SD$[T\|apply, B = b_h, \texttt{control}]$	0.637	1.283	0.189	0.275	0.283	0.598	0.000	0.000
Pr$[apply\|B = b_h, \texttt{control}]$	10.652	0.177	3.256	0.091	0.083	10.529	0.000	0.000
ΔApplications$[B = b_h, \texttt{incentive}]$	5.070	3.353	0.463	6.325	3.655	0.598	0.974	0.000
ΔApplications$[B = b_h, \texttt{wage}]$	2.725	3.831	1.547	7.240	4.317	10.806	0.000	1.106
ΔAbility$[B = b_h, \texttt{incentive}]$	7.858	3.322	1.626	7.120	4.615	3.674	0.940	0.000
ΔAbility$[B = b_h, \texttt{wage}]$	0.196	3.773	0.407	7.980	5.245	6.781	0.000	1.073
$E[Pr[T > a]\|C = c, B = b_h]$	1.326	0.013	0.000	0.000	0.000	2.790	0.000	0.000

Notes: The table reports the moment elasticity for the high-B group and the noisy-ability case, estimated using sixteen moments. The corresponding parameter estimates are reported in column (2) of Table 4. As in Kaboski and Townsend (2011) and Lagakos et al. (2017), we first compute all moments using the structural estimates of the parameters. We then shock by one percent the value of each parameter at a time, and compute the percent change in the simulated moments.

Table A54: Elasticity of simulated moments
(noisy-selection case, core moments)

	μ_T	σ_T	μ_C	σ_C	σ_{TC}	μ_a	σ_a	τ	τ^w
E$[T \mid apply, B = b_h, \texttt{control}]$	1.178	0.000	0.211	0.176	0.178	0.043	0.045	0.000	0.000
SD$[T \mid apply, B = b_h, \texttt{control}]$	0.000	1.306	0.079	0.299	0.307	0.016	0.016	0.000	0.000
Pr$[apply \mid B = b_h, \texttt{control}]$	0.002	0.000	1.184	0.006	0.000	0.246	0.248	0.000	0.000
ΔApplications$[B = b_h, \texttt{incentive}]$	0.000	0.000	0.141	0.978	0.000	0.032	0.031	0.984	0.000
ΔApplications$[B = b_h, \texttt{wage}]$	0.000	0.000	0.182	0.963	0.000	0.164	0.164	0.000	0.981
ΔAbility$[B = b_h, \texttt{incentive}]$	0.000	0.000	0.549	1.938	1.006	0.116	0.116	0.957	0.000
ΔAbility$[B = b_h, \texttt{wage}]$	0.000	0.000	0.563	1.917	1.005	0.080	0.080	0.000	0.945

Notes: The table reports the moment elasticity for the high-B group and the noisy-selection case, estimated using fourteen moments. The corresponding parameter estimates are reported in column (3) of Table 4. As in Kaboski and Townsend (2011) and Lagakos et al. (2017), we first compute all moments using the structural estimates of the parameters. We then shock by one percent the value of each parameter at a time, and compute the percent change in the simulated moments.

Table A55: Elasticity of simulated moments
(noisy-selection case, core moments + beliefs)

	μ_T	σ_T	μ_C	σ_C	σ_{TC}	μ_a	σ_a	τ	τ^w
E$[T \mid apply, B = b_h, \texttt{control}]$	1.181	0.003	0.153	0.178	0.178	0.008	0.008	0.000	0.000
SD$[T \mid apply, B = b_h, \texttt{control}]$	0.000	1.306	0.055	0.299	0.315	0.008	0.000	0.000	0.000
Pr$[apply \mid B = b_h, \texttt{control}]$	0.002	0.000	0.874	0.006	0.000	0.042	0.040	0.000	0.000
ΔApplications$[B = b_h, \texttt{incentive}]$	0.000	0.000	0.107	0.976	0.001	0.006	0.005	0.985	0.000
ΔApplications$[B = b_h, \texttt{wage}]$	0.000	0.000	0.136	0.973	0.000	0.045	0.036	0.000	0.973
ΔAbility$[B = b_h, \texttt{incentive}]$	0.000	0.000	0.407	1.933	1.004	0.017	0.017	0.954	0.000
ΔAbility$[B = b_h, \texttt{wage}]$	0.007	0.000	0.421	1.919	1.003	0.027	0.027	0.000	0.949
$E[Pr[T > a]]$	0.001	0.000	0.000	0.000	0.000	0.049	0.046	0.000	0.000

Notes: The table reports the moment elasticity for the high-B group and the noisy-selection case, estimated using sixteen moments. he corresponding parameter estimates are reported in column (4) of Table 4. As in Kaboski and Townsend (2011) and Lagakos et al. (2017), we first compute all moments using the structural estimates of the parameters. We then shock by one percent the value of each parameter at a time, and compute the percent change in the simulated moments.

Table A56: Robustness of parameter estimates (noisy-ability case)

	(1)	(2)	(3)	(4)	(5)	(6)
			Low B			Low B
μ_C	112.270	137.790	136.100	136.260	115.130	154.300
σ_C	190.750	193.170	192.980	192.750	244.900	178.580
ρ	0.733	0.639	0.640	0.640	0.673	0.482
			High B			Medium B
μ_C	206.860	216.340	217.510	217.350	133.090	220.890
σ_C	213.610	244.670	244.930	245.220	203.250	212.190
ρ	0.587	0.572	0.572	0.571	0.506	0.765
						High B
μ_C						270.270
σ_C						213.910
ρ						0.629
a	2.479	50.301	50.431	50.429	50.239	48.829
		50.504				
τ	10.889	19.045	19.092	18.971	39.747	7.153
				19.191		
τ^w	34.077	55.481	56.187	55.677	127.140	9.941
			55.225			
Goodness of fit	.34147	2.2614	2.2618	2.2626	4.8354	3.0429

Notes: The table shows parameter estimates for the noisy-ability model. Estimation is based on minimum distance estimation. The model in column (1) uses moments based on the cognitive ability score (as opposed to the Raven test score). The model in columns (2)-(4) use the 14 moments reported in Table A48. These models let, in turn, a, τ^w, and τ differ by B group. For each model, we first report the value of the parameter for the low-B group and, in the row below, we report the value for the high-B group. The model in column (5) uses moments obtained by predicting B using an OLS model instead of the post-LASSO estimator. The model in column (6) allows for three types of B. Costs are expressed in Ethiopian Birr.

Table A57: Parameter estimates: heterogeneity by age

	Old	Young
μ_T	45.475	45.653
σ_T	13.136	14.338
μ_C	200.670	262.760
σ_C	223.350	241.070
ρ	0.578	0.580
a	46.315	
τ	23.487	
τ^w	54.258	
Goodness of fit	4.7357	

Notes: Estimates from classical minimum distance estimator. Noisy-ability case. Empirical moments obtained by splitting the sample by age. Empirical and simulated moments reported in Table A59. Costs are expressed in Ethiopian Birr.

Table A58: Parameter estimates: heterogeneity by gender

	Men	Women
μ_T	46.508	44.355
σ_T	13.628	14.911
μ_C	241.340	211.490
σ_C	225.580	219.840
ρ	0.616	0.598
a	46.923	
τ	17.424	
τ^w	44.299	
Goodness of fit	18.535	

Notes: Estimates from classical minimum distance estimator. Noisy-ability case. Empirical moments obtained by splitting the sample by age. Empirical and simulated moments reported in Table A60. Costs are expressed in Ethiopian Birr.

Table A59: Fit between empirical and simulated moments: heterogeneity by age (noisy-ability case, core moments + beliefs)

Moment	Empirical	Simulated
Old		
Pr[$apply$|$\texttt{old}$, $\texttt{control}$]	48.611	48.665
E[T|$apply$, $\texttt{old}$, $\texttt{control}$]	39.132	39.253
SD[T|$apply$, $\texttt{old}$, $\texttt{control}$]	11.656	11.637
ΔApplications[$\texttt{old}$, $\texttt{incentive}$]	8.213	8.910
ΔApplications[$\texttt{old}$, $\texttt{wage}$]	13.606	11.031
ΔAbility[$\texttt{old}$, $\texttt{incentive}$]	0.618	1.055
ΔAbility[$\texttt{old}$, $\texttt{wage}$]	0.265	1.298
E[$Pr[T > a]$|$C = c$, $\texttt{old}$]	0.474	0.475
Young		
Pr[$apply$|$\texttt{young}$, $\texttt{control}$]	48.621	48.580
E[T|$apply$, $\texttt{young}$, $\texttt{control}$]	38.929	38.825
SD[T|$apply$, $\texttt{young}$, $\texttt{control}$]	12.666	12.688
ΔApplications[$\texttt{young}$, $\texttt{incentive}$]	10.645	10.718
ΔApplications[$\texttt{young}$, $\texttt{wage}$]	12.919	13.903
ΔAbility[$\texttt{young}$, $\texttt{incentive}$]	2.974	1.383
ΔAbility[$\texttt{young}$, $\texttt{wage}$]	1.540	1.778
E[$Pr[T > a]$|$C = c$, $\texttt{young}$]	0.475	0.482

Notes: The table shows the empirical and simulated moments for the structural estimates reported in Table A57.

Table A60: Fit between empirical and simulated moments: heterogeneity by gender (noisy-ability case, core moments + beliefs)

Moment	Empirical	Simulated
Men		
Pr[$apply$|$\texttt{male}, \texttt{control}$]	48.028	48.083
E[T|$apply, \texttt{male}, \texttt{control}$]	39.656	39.554
SD[T|$apply, \texttt{male}, \texttt{control}$]	11.992	11.839
ΔApplications[$\texttt{male}, \texttt{incentive}$]	10.677	9.498
ΔApplications[$\texttt{male}, \texttt{wage}$]	14.409	13.749
ΔAbility[$\texttt{male}, \texttt{incentive}$]	0.812	1.246
ΔAbility[$\texttt{male}, \texttt{wage}$]	0.668	1.780
E[$Pr[T > a]$|$C = c, \texttt{male}$]	0.475	0.488
Women		
Pr[$apply$|$\texttt{female}, \texttt{control}$]	50.435	50.335
E[T|$apply, \texttt{female}, \texttt{control}$]	37.087	37.286
SD[T|$apply, \texttt{female}, \texttt{control}$]	12.945	13.108
ΔApplications[$\texttt{female}, \texttt{incentive}$]	6.769	9.981
ΔApplications[$\texttt{female}, \texttt{wage}$]	13.872	13.044
ΔAbility[$\texttt{female}, \texttt{incentive}$]	5.898	1.368
ΔAbility[$\texttt{female}, \texttt{wage}$]	2.035	1.773
E[$Pr[T > a]$|$C = c, \texttt{female}$]	0.473	0.432

Notes: The table shows the empirical and simulated moments for the structural estimates reported in Table A58.

B Proofs

Proposition 2 *Suppose (T, B, C) are observable and distributed according to Assumptions 1 and 2. Further assume jobseekers anticipate that the threshold necessary to get the job is $a \sim \mathcal{N}(\mu_a, \sigma_a)$. Then it follows that for each $B = b_z > 0$, the application incentive (i) increases application rates, and (ii) increases the average ability of applicants, whenever*

$$\frac{\sigma_{T_z}}{\sigma_a \sigma_{C_z}} \frac{b_z}{\sqrt{2\pi}} \leq \rho_z$$

Proof. The application incentive is modeled as a shock that lowers application costs, shifting the distribution of C_z by an amount τ for all B-types; so the proof requires exploring the conditions under which application rates and expected applicant ability are increasing in τ. Under Assumptions 1 and 2, this exploration is identical for each of the B-types, so we drop here the subscripts without loss of generality.

(i) **Increase in application rates.** Using $\mathbb{I}_E$ to denote the indicator function for the application event, we can write the application rate as

$$\begin{aligned}
&\Pr\left(C \leq \Phi\left(\frac{T-\mu_a}{\sigma_a}\right) b + \tau\right) \\
&= \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathbb{I}_E \cdot f(c|t) f(t)\, dc\, dt\right) \\
&= \left(\int_{-\infty}^{\infty} \Phi\left(\frac{\Phi\left(\frac{t-\mu_a}{\sigma_a}\right) b + \tau - E(C|T=t)}{\sigma_{C|T}}\right) f(t)\, dt\right)
\end{aligned}$$

Hence, differentiating with respect to τ gives

$$\frac{1}{\sigma_{C|T}}\left(\int_{-\infty}^{\infty} \phi\left(\frac{\Phi\left(\frac{t-\mu_a}{\sigma_a}\right) b + \tau - E(C|T=t)}{\sigma_{C|T}}\right) f(t)\, dt\right)$$

which is strictly positive.

(ii) **Increase in the average ability of applicants.** Using once again the same notation, we can write the expected ability of applicants as

$$E\left(T \mid C \leq \Phi\left(\frac{T-\mu_a}{\sigma_a}\right)b+\tau\right)$$
$$= \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} t\cdot \mathbb{I}_E \cdot f(c|t)f(t)\,dc\,dt}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathbb{I}_E \cdot f(c|t)f(t)\,dc\,dt}$$
$$= \frac{\int_{-\infty}^{\infty} t\cdot \Phi\left(Y(t;\tau)\right)f(t)\,dt}{\int_{-\infty}^{\infty} \Phi\left(Y(t;\tau)\right)f(t)\,dt}$$

where $Y(t;\tau) = \frac{\Phi\left(\frac{t-\mu_a}{\sigma_a}\right)b+\tau-E(C|T=t)}{\sigma_{C|T}}$

Hence, differentiating with respect to τ gives

$$K\left(\int_{-\infty}^{\infty} t\cdot\phi\left(Y(t;\tau)\right)f(t)\,dt - \frac{\int_{-\infty}^{\infty} t\cdot\Phi\left(Y(t;\tau)\right)f(t)\,dt}{\int_{-\infty}^{\infty}\Phi\left(Y(t;\tau)\right)f(t)\,dt}\int_{-\infty}^{\infty}\phi\left(Y(t;\tau)\right)f(t)\,dt\right)$$

where $K \equiv \frac{1}{\sigma_{C|T}}\left(\int_{-\infty}^{\infty}\Phi\left(Y(t;\tau)\right)f(t)\,dt\right)^{-1}$

Since the term K is strictly positive, the derivative of expected applicant ability with respect to τ will be positive whenever the ratio $\frac{\phi(Y(t;\tau))}{\Phi(Y(t;\tau))}$ is increasing in t.[1]

In general, the Inverse Mills Ratio $\frac{\phi(x)}{\Phi(x)}$ is decreasing in x; so a sufficient condition for a positive derivative is that $Y(t;\tau)$ is decreasing in t. That is

$$\frac{b}{\sigma_a}\phi\left(\frac{t-\mu_a}{\sigma_a}\right) \leq \frac{\partial E(C|T=t)}{\partial t}$$

Since $E(C|T=t) = \mu_C + \frac{\sigma_C}{\sigma_T}\rho(t-\mu_T)$ for the conditional bivariate normal distribution, the condition becomes

$$\frac{\sigma_T}{\sigma_C}\frac{b}{\sigma_a}\phi\left(\frac{t-\mu_a}{\sigma_a}\right) \leq \rho$$

and since the standard normal density $\phi(\cdot)$ has a maximum, this is achieved when

$$\frac{\sigma_T}{\sigma_a\sigma_C}\frac{b}{\sqrt{2\pi}} \leq \rho$$

[1]Notice that whenever the ratio $\lambda = \frac{\phi(Y(t;\tau))}{\Phi(Y(t;\tau))}$ is increasing in t, then t and λ are positively correlated, so

$$\int_{-\infty}^{\infty} t\cdot\lambda\left(Y(t;\tau)\right)\left(\frac{\Phi\left(Y(t;\tau)\right)f(t)}{\int_{-\infty}^{\infty}\Phi\left(Y(t;\tau)\right)f(t)\,dt}\right)dt \geq$$
$$\int_{-\infty}^{\infty} t\left(\frac{\Phi\left(Y(t;\tau)\right)f(t)}{\int_{-\infty}^{\infty}\Phi\left(Y(t;\tau)\right)f(t)\,dt}\right)dt\int_{-\infty}^{\infty}\lambda\left(Y(t;\tau)\right)\left(\frac{\Phi\left(Y(t;\tau)\right)f(t)}{\int_{-\infty}^{\infty}\Phi\left(Y(t;\tau)\right)f(t)\,dt}\right)dt$$

□

Proposition 3 *Suppose (T, B, C) are distributed according to Assumptions 1 and 2. Assume job-seekers are confident about the selection threshold a, but they only observe C and B, which they can use to update their beliefs about the probability that they will pass the recruitment test $T > a$. Then for each $B = b_z > 0$, the application incentive (i) increases application rates, and (ii) increases the average ability of applicants, if and only if*

$$0 < \rho_z < \frac{\sqrt{2\pi}\sqrt{1-\rho_z^2}\sigma_{C_z}}{b_z}$$

Proof. The proof proceeds in 4 steps. In each step, the reasoning applies to all B-types; so we drop again the subscripts without loss of generality.

1. **Cut-off existence.** Let us define $H(c) = \Pr(T_z > a | C_z = c_z) - \frac{c_z}{b_z}$.

   Since $H(0+\epsilon) > 0$ for some small positive ϵ, and $H(b - \epsilon') < 0$ for some positive ϵ', it must be the case that $H(c) = 0$ at least once as c traverses the interval $(0, b)$.

2. **Cut-off uniqueness.** Given cut-off existence, to show that the threshold c^* is unique it suffices to show that $H(c)$ is decreasing in c. Using a standard result from the conditional bivariate normal distribution, we have

   $$H(c) = 1 - \Phi\left(\frac{a - \mu_T - \dfrac{\sigma_T}{\sigma_C}\rho(c - \mu_C)}{\sqrt{1-\rho^2}\sigma_T}\right) - \frac{c}{b}$$

   Hence, differentiating with respect to c gives

   $$\frac{\rho}{\sqrt{2\pi}\sqrt{1-\rho^2}\sigma_C}\exp\left\{-\frac{\left[a - \mu_T - \dfrac{\sigma_T}{\sigma_C}\rho(c-\mu_C)\right]^2}{2(1-\rho^2)\sigma_T^2}\right\} - \frac{1}{b}$$

   From this expression it is easy to check that:

   (a) when $\rho < 0$, the derivative is always negative so $H(c)$ has at least one root, which by monotonicity we know is unique;

   (b) when $\rho = 0$, $\alpha(c)$ is horizontal; so a similar argument applies, and the root is unique; and

(c) when $\rho > 0$, the derivative is negative whenever $\rho < \frac{\sqrt{2\pi}\sqrt{1-\rho^2}\sigma_C}{b}$.

3. **Treatment effect on applications.** For the treatment group the threshold c^* is defined as the level of costs for which $H(c^*;\tau') = Pr(T_z > a|C_z = c_z^*) - \frac{c_z^* - \tau}{b_z} = 0$.

   Hence, using implicit differentiation gives

$$\frac{dc^*}{d\tau'} = -\frac{\partial H(c^*;\tau')/\partial \tau'}{\partial H(c^*;\tau')/\partial c^*}$$

   which is strictly positive.

   Clearly, since the application threshold is increasing in τ, the share of applicants with costs lower than this threshold will also be increasing in τ.

4. **Treatment effect on the average ability of applicants.** Using the law of iterated expectations, we have that

$$\begin{aligned} E(T|C < c^*) &= E(E(T|C)|C < c^*) \\ &= E\left(\mu_T + \frac{\sigma_T}{\sigma_C}\rho(C - \mu_C) \mid C < c^*\right) \\ &= \mu_T - \frac{\sigma_T}{\sigma_C}\rho\ \left(\mu_c - E\left(C \mid C < c^*\right)\right) \\ &= \mu_T - \rho\,\sigma_T\,\frac{\phi\left(\frac{c^*-\mu_C}{\sigma_C}\right)}{\Phi\left(\frac{c^*-\mu_C}{\sigma_C}\right)} \\ &= \mu_T - \rho\,\sigma_T\,\lambda\left(\frac{c^* - \mu_C}{\sigma_C}\right) \end{aligned}$$

   and differentiating with respect to τ gives

$$\frac{d}{d\tau}E(T|C < c^*(\tau)) = -\rho\frac{\sigma_T}{\sigma_C}\frac{dc^*}{d\tau}\frac{d\lambda(c)}{dc}$$

   Since the Inverse Mills Ratio $\lambda(c)$ is decreasing in c, and we have shown that c^* is increasing in τ, the derivative is positive if and only if ρ is positive.

$\square$

C Measures of ability, preferences and sophistication

C.A Cognitive ability

We administer a Raven test and a Stroop test. The Raven test consists of 60 questions (Raven, 2000). Participants are given basic instructions about the test from an instructor and then have to complete the test in 60 minutes. To measure performance on this test, we use the number of correct answers.

We administer the Stroop test proposed by Mani et al. (2013). In this test, the instructor shows a string of digits and then test-taker has to report the number of digits shown. For example, if the string is '44', the correct answer is 'two'. Individuals are shown 75 strings in total. There are two measures of performance: the number of mistakes and (ii) the time taken to complete all strings (which is measured by the instructor using a stopwatch).

C.B Non-cognitive ability

Our main measures of non-cognitive ability are derived from two standard scales: the big five inventory (BFI-44) and the 12-item grit scale (John and Srivastava, 1999; Duckworth et al., 2007). Further, we administer the 12-item core self evaluation scale (Gardner and Pierce, 2010), a 16-item locus of control scale (Lefcourt, 1991), Rosenberg's 10-item self-esteem scale (Rosenberg, 1986). Participants are told that their answers to these psychometric questions are not going to be used to select the workers for the position. We included this feature to maximise truthful reporting.

C.C Time preferences

We measure time preferences over the allocation of effort using a design proposed by Augenblick et al. (2015). We administer this task after the main measures of ability are collected. Applicants are informed that with a certain probability they will be invited to complete a small job, for which they will receive a financial remuneration. This job consists of transcribing 60 pages of text. The job has to be completed in two separate sessions, one week apart from each other. Participants have to transcribe at least five pages per session, but are free to allocate the remaining 50 pages across the two sessions. They are informed that this job is unrelated to the main position and that the effort allocation decisions is not going to be used in the selection process for the main position.

We ask individuals to make allocation decisions for ten different scenarios, one of which will be randomly drawn and implemented. In the first five scenarios, the near work session is on the day following the allocation decision and the late work session is seven days after that. In the last five scenarios, the near work session is two weeks after the allocation decision, and the late work session is seven days after that.[2] Across scenarios, we vary the relative cost of allocating work to early and late sessions. Pages allocated to the near work session always have four sentences. Pages allocated to the late work sessions have $x =$ 6,5,4,3 or 2 sentences, depending on the scenario. $R = \frac{4}{x}$ is thus the rate of exchange of effort between the late and the early work session.

Consider an individual with beta-delta preferences and a cost of effort function given by $(e + \omega)^\gamma$ (where e is the effort chosen and ω is background effort, in our case 5 pages). Augenblick et al. (2015) show that, for each scenario d, the allocation of effort between the near work session at time t and the late work session at time $t + k$ is given by:

(A1)
$$\log \frac{e_{d,t} + \omega}{e_{d,t+k} + \omega} = \frac{\log(\beta)}{\gamma - 1} \text{Early}_d + \frac{\log(\delta)}{\gamma - 1} \text{k}_d + \frac{1}{\gamma - 1} \log(\text{R})_d$$

where Early_d is a dummy for scenarios where the near date is on the following day. We estimate equation (A1) for each experimental group using a two-limit tobit estimator and obtain estimates of β, γ and δ through non-linear combination of the coefficients. We obtain standard errors and test hypothesis about the equality of the coefficients using the delta method. We also obtain individual estimates of each parameter by estimating model (A1) for each individual. As these estimates are less stable, we windsorize the estimates of β_i and classify as present biased any individual with $\beta < .99$.

C.D Risk and social preferences

We measure risk preferences using the following questions adapted from the Global Preferences Survey (Falk et al., 2016b):

[2] Differences in allocation decisions in near and far time horizons enable us to identify present bias. This is a common strategy in the literature on time preferences (Dohmen et al., 2012). Augenblick et al. (2015), on the other hand, elicit allocation decisions for a single time horizon and enable subjects to revise this decision just before the start of the first work session. This feature of the design of Augenblick et al. (2015) would have been difficult to replicate in our setting.

1. How do you see yourself: are you a person who is generally willing to take risks, or do you try to avoid taking risks? (On a scale from 0 to 10).

2. Please imagine the following situation: You can choose between a sure payment and a lottery. The lottery gives you a 50 percent chance of receiving 500 Birr. With an equally high chance you receive nothing. Now imagine you had to choose between the lottery and a sure payment. We will present to you 2 different situations. The lottery is the same in all situations. The sure payment is different in every situation.[3]

3. Please imagine the following situation: you have won a prize in a contest. Now you can choose between two different payment methods, either a lottery or a sure payment. If you choose the lottery there is a 50 percent chance that you receive 1700 Birr and an equally high chance that you receive nothing. Please consider: what would the sure payment need to be in order for you to prefer the sure payment over playing the lottery?

We measure social preferences using the following questions adapted from the Global Preferences Survey (Falk et al., 2016b):

1. How would you assess your willingness to share with others without expecting anything in return, for example your willingness to give to charity? (On a scale from 0 to 10).

2. Imagine the following situation: Today you unexpectedly received 1700 Birr. How much of this amount would you donate to charity?

3. How well does the following statement describe you as a person? I do not understand why people spend a lifetime fighting for a cause that is not beneficial to them. (On a scale from 0 to 10).

C.E Strategic sophistication

We measure strategic sophistication with a simplified beauty contest game (Nagel, 1995). In this game, participants hypothetically play with four other players. Each player reports a

[3]The first sure payment is 170 ETB. If the person chooses the sure payment, the next decisions has a sure payment of 80 Birr. If the person chooses the lottery, the next decisions has a sure payment of 260 Birr. These two choices enable us to bound the CRRA coefficient of the respondent. We assign to each respondent the midpoint of the interval of risk aversion consistent with his or her decisions.

number from zero to six. To win the task, the player has to choose a number that is equal to the average of the numbers chosen by the other players minus one. This simple task enables us to identify different types of strategic reasoning:

- If a player thinks that other subjects choose numbers at random, then he or she expects the average number chosen by the other players to be three. The optimal strategy is then choose number two. This corresponds to $k = 1$ behaviour.
- If a player thinks that other subjects are $k = 1$, then he or she expects the average number to be two. The optimal strategy is thus to choose number one. This corresponds to $k = 2$ behaviour.
- Finally, if a player thinks that other subjects are $k = 2$, then they he or she expects the average number to be one. The optimal strategy is thus to choose number zero. This corresponds to $k = 3$ behaviour.

D The Estimation of the Structural Model

D.A The analytical expressions of the moments

D.A.1 Noisy selection

Application rates $\Pr\left(C_z \leq \Phi\left(\frac{T_z-\mu_a}{\sigma_a}\right)(b_z+\tau^w)+\tau\right)$**:**

$$\int_{-\infty}^{\infty} \Phi\left(\frac{\Phi\left(\frac{t_z-\mu_a}{\sigma_a}\right)(b_z+\tau^w)+\tau-E(C_z|T_z=t_z)}{\sigma_{C_z|T_z}}\right) f(t_z)\,dt_z \tag{A2}$$

Expected applicant ability $E\left(T_z \mid C_z \leq \Phi\left(\frac{T_z-\mu_a}{\sigma_a}\right)(b_z+\tau^w)+\tau\right)$**:**

$$\frac{\int_{-\infty}^{\infty} t_z \cdot \Phi\left(Y(t_z)\right) f(t_z)\,dt_z}{\int_{-\infty}^{\infty} \Phi\left(Y(t_z)\right) f(t_z)\,dt_z} \tag{A3}$$

where

$$Y(t_z) = \frac{\Phi\left(\frac{t_z-\mu_a}{\sigma_a}\right)(b_z+\tau^w)+\tau-E(C_z|T_z=t_z)}{\sigma_{C_z|T_z}}$$

Dispersion in applicant ability $Var\left(T_z \mid C_z \leq \Phi\left(\frac{T_z-\mu_a}{\sigma_a}\right)(b_z+\tau^w)+\tau\right)$**:**

$$\frac{\int_{-\infty}^{\infty}\left(t_z - E\left(T_z \mid C_z \leq \Phi\left(\frac{T_z-\mu_a}{\sigma_a}\right)(b_z+\tau^w)+\tau\right)\right)^2 \cdot \Phi\left(Y(t_z)\right) f(t_z)\,dt_z}{\int_{-\infty}^{\infty} \Phi\left(Y(t_z)\right) f(t_z)\,dt_z} \tag{A4}$$

where

$$Y(t_z) = \frac{\Phi\left(\frac{t_z-\mu_a}{\sigma_a}\right)(b_z+\tau^w)+\tau-E(C_z|T_z=t_z)}{\sigma_{C_z|T_z}}$$

Expected recruitment probability $E\left(Pr\left(a \leq T_z\right)\right)$**:**

$$\int_{-\infty}^{\infty} \Phi\left(\frac{t_z-\mu_a}{\sigma_a}\right) f(t_z)\,dt_z \tag{A5}$$

D.A.2 Noisy ability

Application threshold c_z^*:

$$\text{(A6)} \qquad \frac{c_z^* - \tau^w}{b_z + \tau} = 1 - \Phi\left(\frac{a - \mu_{T_z} - \dfrac{\sigma_{T_z}}{\sigma_{C_z}}\rho_z(c_z^* - \mu_{C_z})}{\sqrt{1-\rho_z^2}\sigma_{T_z}}\right)$$

Application rates $\Pr\left(C_z < c_z^*\right)$:

$$\text{(A7)} \qquad \Phi\left(\frac{c_z^* - \mu_{C_z}}{\sigma_{C_z}}\right)$$

Expected applicant ability $E(T_z|C_z < c_z^*)$:

$$\text{(A8)} \qquad \mu_{T_z} - \rho_z\,\sigma_{T_z}\,\frac{\phi\left(\frac{c_z^* - \mu_{C_z}}{\sigma_{C_z}}\right)}{\Phi\left(\frac{c_z^* - \mu_{C_z}}{\sigma_{C_z}}\right)}$$

Dispersion in applicant ability $Var(T_z|C_z < c_z^*)$:

$$\text{(A9)} \qquad \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(t_z - E(T_z|C_z < c_z^*)\right)^2 \cdot \mathbb{I}_E \cdot f(c_z|t_z)f(t_z)\; dc_z\, dt_z}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathbb{I}_E \cdot f(c_z|t_z)f(t_z)\; dc_z\, dt_z}$$

Expected recruitment probability $E\left(Pr\left(T_z \geq a \mid C_z = c_z\right)\right)$:

$$\text{(A10)} \qquad \int\limits_{-\infty}^{\infty}\left(1 - \Phi\left(\frac{a - E\left(T_z \mid C_z = c_z\right)}{\sigma_{T_z|C_z}}\right)\right) f(c_z)\, dc_z$$

D.A.3 Parameter estimation and standard errors

We use the formulas reported above to calculate simulated moments for different draws of parameters, and then compute the value of the loss function (**??**). We minimize this function using MATLAB's unconditional minimizer `fminunc`.

To calculate standard errors, we produce 100 draws of bootstrapped moments (the boostrap procedure includes the estimation of B and the demediation of application rates). We then estimate each of the four versions of the model reported in Table **??** for each set of bootstrapped moments.[4] The standard error of a parameter is given by the standard deviation of that parameter over these replications.

[4] To make the bootstrap computationally manageable we truncate the minimisation procedure at 1,500 function evaluations and set a slightly lower optimality threshold. The minimisation procedure is truncated in about 8 percent of the simulations. We only use simulations that have converged to a minimum.

D.B The Internal Rate of Return of the interventions

We calibrate the Internal Rate of Return (IRR) of the interventions to represent the returns to a typical a firm hiring a clerical worker in Addis Ababa. Table D1 below summarises all our key assumptions, which we describe in detail in what follows. First, we estimate the number of potential applicants. The average firm in our sample receives 50 job applications for a clerical post. On the basis of this, we assume that the pool of potential applicants is composed of 100 individuals.[5]

Second, we quantify the monthly benefit of the interventions using the following formula:

$$\texttt{monthly benefit} = \pi * \Delta\texttt{E[Raven|hire]} * \texttt{No hires} \tag{A11}$$

where π is the monetary return to an extra unit of performance on the Raven test, $\Delta\texttt{E[Raven|hire]}$ is the change in the expected Raven score of a hire, and `No hires` is the number of workers hired. To obtain a value for π, we regress wages on Raven test scores using the data of Abebe et al. (2020). We compute $\Delta\texttt{E[Raven|hire]}$ using the structural estimates from the noisy-ability case estimated with the core moments.[6] `No hires` is the average number of workers hired by a firm in our sample in a given hiring round. Finally, we assume that the monthly benefit accrues to the firm for 45 months. We obtain this number by taking the average separation rate reported by firms in our survey and calculating the expected duration of a match.

Third, we quantify the one-off cost of application incentives using the following formula:

$$\texttt{cost} = r * \Delta\texttt{E[No Applicants]} + (100 * \texttt{E[No applicants]}) \tag{A12}$$

where r is the cost of reviewing one additional application. We calculate r using firms' self-reports. In particular, firms report that it takes them about one hour of a manager's time to review an application. We price this hour at the median hourly salary of the HR staff who review applications in these firms. We obtain $\Delta\texttt{E[No Applicants]}$ (the expected change in the number of applicants compared to the control condition) and `E[No applicants]` (the expected total number of applicants in the incentive condition) using our structural estimates.

[5]This corresponds to an application rate of 50%, which is in line with the average application rate across treatment conditions on our experiment.

[6]To calculate `E[Raven|hire]`, we first need to calculate the selectivity threshold used by the employer. This is the threshold that ensures that the expected number of applicants with ability above the threshold is equal to the desired number of hires. We then calculate the expected change in Raven scores among applicants above this threshold.

Finally, the total cost of the wage intervention is given by (i) a one-off cost of $r * \Delta \texttt{E[No Applicants]}$ (as the employer needs to review additional applications), and (ii) a salary cost of 1,600 ETB for each hired workers, for three months.

We calculate the Internal Rate of Return (IRR) of the interventions using these calibrated costs and benefits, and applying MATLAB's `irr` function. We bootstrap the estimation procedure to obtain a confidence interval for the IRR: (i) for each new bootstrap sample, we solve the model again and obtain a new set of parameters, (ii) for each new set of parameters, we calculate the IRR of the interventions, (iii) we use this distribution of IRRs to compute the confidence interval. In step (1), for the incentive and wage intervention, we use the same bootstrapped parameter estimates that we obtained to calculate parameter standard errors. For the counterfactual interventions that target the incentive on the basis of gender or age, we run 100 additional simulations (50 for each model).

Table D1: Assumptions for the cost-benefit analysis

	Value
Number of potential applicants	100
Number of workers hired	3
Monthly return to one extra point on the Raven test (π)	22.8
Marginal cost of interviewing one more applicant (r)	38.8
Expected tenure on the job (no. months)	45

E Comparison with original plan

For the reduced-form analysis, we follow a registered pre-analysis plan. We have updated the plan in the following way:

1. Heterogeneity. We planned to study heterogeneous treatment effects with respect to individuals' savings, cash on hand, expenditure. These variables were collected during phone call number two. The plan was to ask these questions about the month preceding phone call number one. Due to miscommunication with the field team, this plan was not implemented. Instead, the questions were asked about the last completed month, which in most cases included several weeks after treatment. We are thus unable to use these variables to study heterogeneity. Further, we included three additional dimensions of heterogeneity measured during the second phone call: a measure of credit constraints and two questions on time preferences. The credit constraint question – a newly-designed question applying the logic of multiple price list to the measurement of credit constraints – was hard to understand for respondents according to the reports of the field team. Similarly, the two questions on time preferences were ultimately poorly formulated. These worries compound the fundamental problem that these variables were measured *after* treatment and are thus not suitable to study the heterogeneity of treatment effects. We thus chose not to use them for heterogeneity analysis in the paper.

2. Quantile regressions. Due to a lack of fine-grained variation in the scores at the top and at the bottom of the distribution in the individual tests, we have performed most the the quantile regression analysis on overall indices of ability computed over the pre-specified families of ability measures.

3. Outcomes. We are unable to report results related to the wage paid by and the location of the jobs that jobseekers apply for between the two phone interviews because of the large amount of missing data. For both variables, we have more than 50% missing data.

4. The variable 'value of the job'. This variable was not part of the original plan. This variable has a clear theoretical interpretation and we thus prefer it to splitting the sample based on endogenous stratification, as per our original plan.

5. Experiment two. We originally planned to use experiment number two to obtain the

weights that managers placed over the various dimensions of quality. However, we found this challenging to implement. We ran several pilot sessions of a design where managers were shown several pairs of candidates with different characteristics and had to pick one candidate in each pair. These pilot sessions suggested that managers place a larger weight on cognitive ability, compared to non-cognitive ability or experience. However, managers decisions were very hard to predict and to reconcile with our model of decision utility (our estimated model could predict decisions only modestly better than a random guess). We thus opted for the simpler and more transparent task which is described in the paper.

6. Robustness tests. We have also included several robustness tests that were not pre-specified, but were requested by referees, suggested by seminar audiences or motivated by the findings of the main pre-specified analysis.

F The value of the job

In this section, we describe how we calculate B, the value of the experiment's job. This value is given by the stream of utility that the worker obtains if they get the job, minus the stream of utility that the worker would have obtained otherwise:

$$b = \begin{cases} V(j) - V(u) & \text{if currently unemployed} \\ V(j) - V(e) & \text{if currently employed} \end{cases} \tag{A13}$$

where $V(j)$ is the gross value of the experiment's job, $V(u)$ is the value of being unemployed, and $V(e)$ is the value of being employed at the wage that the market currently pays for the worker's skills (we will refer to this as the 'market wage').

We proceed in two steps. First, we characterise $V(j)$, $V(u)$ and $V(e)$ as functions of the wage paid by the experiment's job, market wages, worker impatience and the probability of finding and losing a job. Second, we forecast the market wage of each worker using a Post-LASSO estimator (Belloni et al., 2014) and make informed assumptions about the other parameters. Throughout this section, we assume that time is discrete and measured in months. We also assume that workers have a time-separable, linear utility function of the following form:

$$U_t = \sum_{k=0}^{T} \delta^{t+k} E[w_{t+k}] \tag{A14}$$

We start by calculating the value of unemployment. We assume that the worker values non-work time at c (Mas and Pallais, 2017). This includes transfers, the value of leisure, etc. We assume that c is given to the worker at the end of the month. Further, we assume that the worker will find a job in the next period with probability p. The value of being in unemployment is thus given by:

$$\begin{aligned} V(u) &= \delta c + \delta^2 p V(e) + \delta^2 (1-p) V(u) \\ &= \frac{\delta c + \delta^2 p V(e)}{1 - \delta^2 (1-p)} \end{aligned} \tag{A15}$$

The value of being employed, on the other hand, is given by:

$$\begin{aligned} V(e) &= \delta w + \delta^2 (1-q) V(e) + \delta^2 q V(u) \\ &= \frac{\delta w + \delta^2 q V(u)}{1 - \delta^2 (1-q)} \end{aligned} \tag{A16}$$

where w is the market wage and q is the probability of losing the job in any given period of time. We can substitute $V(e)$ into (A15) to derive an expression that defines $V(u)$ only as a function of the parameters c, w, δ, p and q:

$$\begin{aligned} V(u) &= \frac{\delta c}{1-\delta^2(1-p)} + \frac{\delta^2 p}{1-\delta^2(1-p)} \frac{\delta w + \delta^2 q V(u)}{1-\delta^2(1-q)} \\ &= \left(1 - \frac{\delta^4 pq}{(1-\delta^2(1-p))\,(1-\delta^2(1-q))}\right)^{-1} \times \left(c' + \frac{\delta^2 p}{1-\delta^2(1-p)} w'\right) \end{aligned} \tag{A17}$$

where $c' = \frac{\delta c}{1-\delta^2(1-p)}$ and $w' = \frac{\delta w}{1-\delta^2(1-q)}$. The value of being employed can be obtained by substituting (A17) into (A16).

Finally, the gross value of getting the experiment's job for a worker in treatment group f is given by:

$$V(j) = \sum_{k=0}^{3} \delta^k w_f + \delta^4 \left(pV(e) + (1-p)V(u)\right)$$

The worker will obtain wage w_f for three consecutive months and will then return to unemployment. For simplicity, we assume that work experience in the experiment's job does not affect future wages and that the worker will only hear about new job opportunities in the last month of the job. These assumptions make our estimates of $V(j)$ conservative.

We can now write an expression for the value of the job for an unemployed person. This is given by:

$$V(j) - V(u) = \sum_{k=0}^{3} \delta^k w_f + \delta^4 \left(pV(e) + (1-p)V(u)\right) - V(u)$$

Further, the value of the job for an employed person is given by:

$$V(j) - V(e) = \sum_{k=0}^{3} \delta^k w_f + \delta^4 \left(pV(e) + (1-p)V(u)\right) - V(e)$$

In our second step we forecast market wages. To do this, we use the Post-LASSO estimator recommended by Belloni et al. (2014). This estimator is obtained in two stages. First, we regress individual wages on a large set of covariates, using the LASSO estimator and all observations of jobseekers who have a formal job. This allows us to select a sub-set of covariates that can be used for forecasting. Second, we run an OLS regression of wages on the covariates selected in the first stage (using only control group observations, to minimise distortions in reporting potentially induced by the interventions) and use the OLS coefficients to derive a forecast of w for each worker.

The Post-LASSO estimator is recommended to produce forecasts when a large number of potentially informative covariates are available. In these settings, estimators that maximise in-sample fit often have poor out-of-sample properties, as they tend to fit some of the noise in the data. The original LASSO estimator reduces over-fitting by imposing a penalty on non-zero coefficients. More precisely, for a canonical model:

$$y_i = \sum_{j=1}^{p} x_{i,j}\beta_j + u_i \tag{A18}$$

the LASSO estimator of the parameter vector β is obtained by minimising the following function:

$$\widehat{\beta} = \arg\min_{\beta} \sum^{n} \left(y_i - \sum_{j=1}^{p} x_{i,j}\beta_j \right)^2 - \lambda \sum_{j=1}^{p} |\beta_j|\gamma_j$$

where λ is a penalty parameter and γ_j are penalty loadings. One problem with this estimator is that the non-zero coefficients tend to be biased towards zero. The Post-LASSO estimator reduces this bias by re-estimating the coefficients with OLS.

We use a rich set of variables in order to forecast wages. These variables describe the socio-demographic characteristics of workers, their educational achievements, and their labour market experience. We report the full list of variables in table F1 below. To maximise the flexibility of our empirical model, we discretise continuous variables and include dummies for each possible discretised value of the variable. Finally, our measure of wages refers to the jobs that subjects held at the time of the first interview.[7] We report the coefficients estimates obtained with the Post-Lasso estimator in Table F2 below. The first column shows the estimates obtained by using the theoretically optimal penalty and the second column shows the estimate obtained with a manually-set lower penalty, which allows us to capture a number of additional plausible predictors. The predicted values we obtain from these two models are highly correlated. In what follows, we use the predicted values obtained with the optimal penalty.

We make the following assumptions on the remaining parameters. First, we assume that the monthly discount factor is $\delta = 0.786$. To determine this figure, we use the daily discounting factor estimated in a recent experiment in Nairobi (Balakrishnan et al., 2015). The

[7]We windsorise the forecast at the 5^{th} and 95^{th} percentiles so that we do not rely on extreme forecasts. Further, we adjust forecasted wages (by applying a simple location shift) to ensure that the mean of the forecast matches that of representative data for workers in Addis Ababa of comparable age and level of education.

estimates of Balakrishnan et al. (2015) suggest relatively high levels of impatience, which is consistent with the cross-country survey evidence reported by Falk et al. (2016a) for sub-Saharan Africa. Second, we set the probability of finding a job to 15.3 percent and the probability of losing a job to 11.6 percent, respectively. These figures reflect monthly transition rates from non-employment to employment, and vice-versa, which we calculate using the high-frequency panel data collected by Abebe et al. (2020). Finally, we assume that the value of c is 1,230 ETB. We calculate this figure by using estimates of the value of non-work time relative to mean wages from Mas and Pallais (2017), and combine this ratio with mean forecasted wages for our sample. This figure seems realistic in our context, as unemployed job-seekers report an average monthly expenditure of about 1,000 ETB.

We estimate that the position has positive value for about 61 percent of the individuals in our sample. To confirm that our estimates are informative, we regress the application dummy on our estimate of the value of the job. We find a large and significant correlation: a one standard deviation increase in the value of the job is associated with a 10 percentage points increase in application rates. We report the estimates in Table F3 below. Finally, in Table F4 we show that the measure of B that we obtain by using different predicted values of the wage (LASSO forecast with optimal penalty, LASSO forecast with manual penalty, and OLS) are highly correlated with each other.

Table F1: Variables used to forecast wages

Variable	Description
female	Female
age	Age
age_sq	Age squared
born_aa	Individual was born in Addis Ababa
newspaper	Individual has found out about the vacancy in the newspaper
amharic	First language is Amharic
oromo	First language is Oromifa
engineer	Engineering or hard science background
economics	Economics background
social_scientist	Degree in social science (other than economics)
GPA_dummy_	Dummies for GPA score (1 point intervals)
wexperience	Wage work experience (number of months)
wexperience_sq	Wage work experience squared
e_type_	Dummies for type of employer in last job
wage_dummy_	Dummies for wage earned in last job (2,000 ETB intervals)
sexperience	Individual has experience in self-employment
subcity_	Dummies for the subcity of residence of the respondent

Table F2: Post-LASSO regression of wages in Ethiopian Birr
(control group observations)

	Optimal penalty	Manual penalty
	(1)	(2)
Heard of job on newspaper	108.186	52.678
	(191.410)	(176.968)
Economics background	719.240	686.784
	(266.857)	(270.642)
Work experience (months)	18.188	16.703
	(3.101)	(3.015)
Worked for private foreign business		179.904
		(1058.206)
Age	45.814	23.520
	(37.495)	(35.126)
GPA dummy (2-3)		239.314
		(165.893)
Previous wage dummy (2000-4000)		995.650
		(519.302)
Previous wage dummy (4000-6000)	1862.386	2091.702
	(438.405)	(437.934)
Previous wage dummy (6000-8000)		4175.537
		(846.208)
Previous wage dummy (8000-10000)	3682.031	4012.018
	(594.769)	(664.741)
Obs.	361	361

Notes: Post-LASSO regressions to forecast market wages. In the first stage of the Post-LASSO procedure, we run a LASSO regression of wages on the set of covariates described in Table F1. In the second stage, we run an OLS regression of wages on the covariates selected by the LASSO estimator in the first stage. In column (1), we report estimates obtained by applying the optimal LASSO penalty parameter. In the column (2), we report estimates obtained by applying a manually-chosen, lower penalty parameter that enables us to select a larger number of covariates. Robust standard errors in parentheses. Sample used: control individuals employed at baseline.

Table F3: Regression of applications on the value of the job

	Applied to the experiment's job
	(1)
B (z score)	0.105
	(0.007)
Constant	0.512
	(0.007)
Obs.	4686

Notes: OLS regressions. The dependent variable is a dummy for whether the individual has applied to the experiment's job. The independent variable is the estimate of the value of the job B, obtained using the market wage forecast from the model of column 1 of Table F2. B is windsorised at the 10^{th} and 90^{th} percentiles. Robust standard errors in parentheses. Sample used: baseline sample.

Table F4: Correlation between measures of B obtained with different market wage forecasts

	Dep var: B (optimal LASSO)	
	(1)	(2)
B (OLS)	0.919	
	(0.009)	
B (manual LASSO)		0.955
		(0.010)
Constant	-89.205	-74.051
	(8.608)	(4.336)
R2	0.91	0.94
Obs.	3932	4686

Notes: OLS regressions. The dependent variable is the measure of B based on the market wage forecast produced by running the post LASSO estimator with the optimal penalty (reported in in column (1) of Table F2). This is the measure of B that is used in the rest of the paper. In the first column, we regress this measure of B on an alternative measure of B based on an OLS forecast of the market wage. The OLS estimator uses all the variables that are initially available to the LASSO estimator. In the second column, we regress our main measure of B on an alternative measure of B based on the post-LASSO market wage forecast reported in column (2) of Table F2. To obtain this second forecast, we impose on the post-LASSO estimator uses a manual penalty parameter. This results in the estimator relying on a larger number of covariates compared to the estimator that uses the optimal penalty parameter. All measures of B are windsorised at the 10^{th} and 90^{th} percentiles. Robust standard errors in parentheses. Sample used: baseline sample.

References

Abebe, Girum, Stefano Caria, Marcel Fafchamps, Paolo Falco, Simon Franklin, and Simon Quinn, "Anonymity or Distance? Job Search and Labour Market Exclusion in a Growing African City," *Unpublished*, 2020.

Attanasio, Orazio and Katja Kaufmann, "Educational Choices, Subjective Expectations, and Credit Constraints," *NBER Working Paper No. 15087*, 2009.

Augenblick, Ned, Muriel Niederle, and Charles Sprenger, "Working over Time: Dynamic Inconsistency in Real Effort Tasks," *The Quarterly Journal of Economics*, 2015, pp. 1067–1115.

Autor, David H and Michael J Handel, "Putting Tasks to the Test: Human Capital, Job Tasks, and Wages," *Journal of Labor Economics*, 2013, *31* (S1), S59–S96.

Balakrishnan, Uttara, Johannes Haushofer, and Pamela Jakiela, "How Soon Is Now? Evidence of Present Bias from Convex Time Budget Experiments," *Unpublished*, 2015.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen, "High-Dimensional Methods and Inference on Structural and Treatment Effects," *The Journal of Economic Perspectives*, 2014, *28* (2), 29–50.

Benjamini, Yoav, Abba M Krieger, and Daniel Yekutieli, "Adaptive Linear Step-up Procedures that Control the False Discovery Rate," *Biometrika*, 2006, *93* (3), 491–507.

Dohmen, Thomas J, Armin Falk, David Huffman, and Uwe Sunde, "Interpreting Time Horizon Effects in Inter-Temporal Choice," *Unpublished*, 2012.

Duckworth, Angela L, Christopher Peterson, Michael D Matthews, and Dennis R Kelly, "Grit: Perseverance and Passion for Long-Term Goals.," *Journal of Personality and Social Psychology*, 2007, *92* (6), 1087.

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde, "Global Evidence on Economic Preferences," *Unpublished*, 2016.

_ , _ , _ , David Huffman, and Uwe Sunde, "The Preference Survey Module: A Validated Instrument for Measuring Risk, Time, and Social Preferences," *Unpublished*, 2016.

Gardner, Donald G and Jon L Pierce, "The Core Self-Evaluation scale: Further Construct Validation Evidence," *Educational and Psychological Measurement*, 2010, *70* (2), 291–304.

John, Oliver P and Sanjay Srivastava, "The Big Five Trait Taxonomy: History, Measurement, and Theoretical Perspectives," *Handbook of Personality: Theory and Research*, 1999, *2* (1999), 102–138.

Kaboski, Joseph P and Robert M Townsend, "A Structural Evaluation of a Large-Scale Quasi-Experimental Microfinance Initiative," *Econometrica*, 2011, *79* (5), 1357–1406.

Laajaj, Rachid and Karen Macours, "Measuring Skills in Developing Countries," *Unpublished*, 2017.

Lagakos, David, Mushfiq Mobarak, and Michael E Waugh, "The Welfare Effects of Encouraging Rural-Urban Migration," *Unpublished*, 2017.

Lefcourt, Herbert M, *Locus of Control*, Academic Press, 1991.

Mani, Anandi, Sendhil Mullainathan, Eldar Shafir, and Jiaying Zhao, "Poverty Impedes Cognitive Function," *Science*, 2013, *341* (6149), 976–980.

Mas, Alexandre and Amanda Pallais, "Labor Supply and the Value of Non-Work Time: Experimental Estimates from the Field," *Unpublished*, 2017.

McKenzie, David, "Identifying and Spurring High-Growth Entrepreneurship: Experimental Evidence from a Business Plan Competition," *American Economic Review*, 2017, *107* (8), 2278–2307.

Nagel, Rosemarie, "Unraveling in Guessing Games: An Experimental Study," *The American Economic Review*, 1995, *85* (5), 1313–1326.

Raven, John, "The Raven's Progressive Matrices: Change and Stability over Culture and Time," *Cognitive psychology*, 2000, *41* (1), 1–48.

Rosenberg, Morris, *Conceiving the Self*, RE Krieger, 1986.