Job Search and Hiring with Two-sided Limited Information about Workseekers' Skills*

Eliana Carranza† Robert Garlick‡ Kate Orkin§ Neil Rankin¶

January 11, 2019

Abstract

Firms typically make hiring decisions with limited information about workseekers' skills and productivity. Similarly, workseekers make job search and employment decisions with limited information about their own skills and productivity. These information frictions on both the demand and supply sides of the labor market can distort job search and hiring decisions, lowering total employment, total earnings, and firm productivity. We test for the existence of information frictions using a series of randomized controlled trials in urban South Africa. We directly assess workseekers' skills in six domains and experimentally vary whether workseekers receive information about their own skills and whether they can signal this information to firms. We find that giving information to workseekers has large effects on their beliefs about their skills, no effects on their job search behavior, and limited effects on their employment outcomes. Helping workseekers signal their skills to firms substantially increases their employment rate and earnings. We conclude that information frictions exist on both sides of the labor market but only the demand-side frictions appear important for labor market outcomes.

JEL codes: J23, J24, J31, J41, O15, O17

*We are grateful for exceptional work by research managers Emmanuel Bakirdjian and Nilmini Herath; research assistants Shelby Carvalho, Allegra Cockburn, Lukas Hensel, Wim Louw, Svetlana Pimkina, Ashley Pople, Sri Ramesh, Caitlan Russell, Wendy Trott, and Laurel Wheeler; multiple staff at the Harambee Youth Employment Accelerator; the Abdul Latif Jameel Latif Poverty Action Lab Africa Office (J-PAL Africa); and administrative staff at the Universities of Cape Town (UCT), Duke, Oxford, and Stellenbosch, the Centre for Study of African Economies, and the Africa Gender Innovation Lab and Jobs Practice at the World Bank. The project is conducted in collaboration with the World Bank Africa Gender Innovation Lab, Jobs Practice and Jobs Trust Fund, and is partly funded by the Private Enterprise in Low Income Countries Consortium. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors, and do not necessarily represent the views of the World Bank, its Executive Directors, or the governments of the countries they represent. This study has been approved by the Commerce Faculty Ethics in Research Committee at UCT, the Duke Institutional Review Board for the Protection of Human Subjects in Non-Medical Research (protocol # D0368) and the Department of Economics Research Ethics Committee at Oxford (protocol # ECONCIA15-055). Our pre-analysis plan is available at https://www.socialscienceregistry.org/trials/1631. This is a preliminary draft of the paper that does not yet include all pre-specified analyses.

†World Bank, ecarranza@worldbank.org

‡Duke University, robert.garlick@duke.edu

§University of Oxford, kate.orkin@merton.ox.ac.uk

¶University of Stellenbosch, neilrankin@sun.ac.za

1 Introduction

Many countries face high and rising unemployment and underemployment, particularly for young people (World Bank, 2018; ILO, 2016). This challenge occurs across both developed and developing countries in Latin America, the Middle East, North Africa, South Europe, and sub-Saharan Africa. This poses major economic, political and social challenges: periods of unemployment can permanently harm youths' employment trajectories and contribute to crime and political instability (Altonji, Kahn, and Speer, 2015; Beaudry and DiNardo, 1991; Blattman and Annan 2016; Freeman, 1999). At the same time, firms in many developing countries have low labor productivity and low growth rates and report that they cannot find and hire suitable workers (Hardy and McCasland, 2017; Hsieh and Olken, 2014).

These patterns are consistent limited information about workseekers' skills that prevents firm-worker matching.[1] On the demand side, firms who observe noisy proxies for workseekers' skills may hire fewer workers if the firms are risk-averse or their output is a concave function of skill-adjusted labor use. This employment reduction will be larger if firms face minimum wages or hiring/firing costs, which are a feature of many labor markets around the world. On the supply side, noise in workseekers' beliefs about their own skills will distort their job search behavior and their potentially employment outcomes, though the direction of these distortions is theoretically ambiguous. Both the demand- and supply-side distortions can reduce the quality of firm-worker matches, reducing output and earnings even if there are no employment effects. Both the demand- and supply-side distortions will be larger when formal education is weakly correlated with skills, as documented by Pritchett (2013), and for young workseekers who cannot use prior work experience to certify their skills. The latter effect can lead to inefficiently low hiring of young workseekers from a social welfare perspective, as firms do not want to incur the cost of making first-time hires and allowing the market to learn about workseekers' skills through experience.

We use a series of supply- and demand-side experiments in urban South Africa to demonstrate that information frictions exist, distort labor market outcomes, and potentially contribute to unemployment. We randomly vary the extent of supply- and demand-side information frictions by randomly varying the information that workseekers and firms observe about workseekers' skills.[2] We measure workseekers' skills in six different dimensions using validated psychometric tests.

On the supply side, we assess roughly 7,000 workseekers' skills and randomly assign workseekers to a control group or a private treatment group. Workseekers in the private treatment group observe their assessment results through printed reports and a briefing with a psychologist. Workseekers in

[1] These patterns are also consistent with other explanations, including the possibility of aggregate mismatches between skills supply and demand. We take aggregate demand for and supply of skills as fixed in this work.

[2] We examine only frictions due to firms and workseekers having limited information about workseekers' skills. We do not explicitly study other types of information frictions in this proposal, such as limitations on workseekers' ability to observe firm characteristics or on firms' ability to observe workers' on-the-job effort.

the control group do not. We measure workseekers' beliefs about their own skills, beliefs about the labor market, search behavior, and employment outcomes before the treatment, four months after the treatment, and one year after the treatment. After four months, workseekers in the treatment group have substantially more accurate beliefs about their skills relative to other workseekers. The treatment does not have robust effects on search behavior and has only small effects on employment that are typically not significantly different from zero. We conclude that supply-side information frictions exist but are not quantitatively important for labor market outcomes.[3]

On the demand side, we conduct two experiments. First, we randomly assign some workseekers to a public treatment group, whose members observe their assessment results and can inform prospective employers about these results. This reduces demand-side information frictions if workseekers make the endogenous decisions to share the results. The information is provided by a large, well-known assessment agency and is likely to be credible to firms. After four months, workseekers in the treatment group have higher employment, higher earnings, and more accurate beliefs about their skills relative to other workseekers. Treatment does not robustly change their search behavior. We conclude that demand-side information frictions exist and are quantitatively important for labor market outcomes.[4]

Second, we conduct an ongoing audit study using real resumes from workseekers in our sample. We randomise whether the resumes include reports on workseekers' skills, like those provided in the public treatment group.[5] Comparing callback rates across resumes with and without assessment results directly identifies the effect of demand-side frictions. This complements the experiment with workseekers: the audit study delivers more information to firms without conditioning on workseekers' decisions to share the results but only measures effects on callbacks. Both the workseeker and audit studies treat firm-level labor demand as fixed: firms are unlikely to make more hires because a few workseekers have more reliable information about their skills. In ongoing work, we randomize access to information at the firm level to test if firms with access to more information about a large pool of workseekers hire more and/or better-matched workseekers.

We build on a growing literature on information frictions in labor markets. Spence (1973) developed the canonical model of limited information, in which workseekers invest in education to signal their skills to firms. A large literature, mostly from the US, documents patterns consistent

[3] This draft only includes results measured four months after treatment. Results from the one-year follow-up survey will be included in July 2018.

[4] The public treatment effects on employment are significantly larger than the private treatment effects. This difference is consistent with demand-side information frictions, as we argue, or with other mechanisms. On the supply side, workseekers might search more when they know they can convey their assessment results to firms. But search behavior is not different in the public and private treatment groups. On the demand side, firms might respond to the existence or branding of the assessment results, rather than the actual results. We test the latter explanation with a placebo treatment in which workseekers can inform firms about the assessments and assessment process but not their own assessment results. Employment outcomes in the placebo group are closer to outcomes in the control group than the public treatment group.

[5] The audit study is currently in progress. Results will be included in this paper from August 2018.

with information frictions: the relationship between wages and skills that are observed by the econometrician but unobserved by firms becomes stronger with tenure, consistent with employers learning about skills over time (Altonji and Pierret, 2001; Farber and Gibbons, 1996; Kahn and Lange, 2014). Similarly, MacLeod et al. (2017) show that the association between college reputation and employment outcomes in Colombia is attenuated when they condition on individual skills, suggesting firms use education as a proxy for hard-to-observe skills. Arcidiaono et al. (2010) show that information frictions are more salient for workers without college education and suggest this reflects a weak skill:education correlation at the high school level. This helps to motivate our focus in this paper on a sample with relatively low education levels.

Much of this literature tests for information frictions using earnings or wage data conditional on employment. This is less informative in labor markets with high unemployment, informal employment, and self-employment. This can understate the importance of information frictions if frictions inhibit hiring. More recent work, mostly in developing countries, focuses on the employment margin by providing workseekers with more information about their skills. This changes their job search behaviour and, in some cases, their employment outcomes (Abebe et al., 2016; Abel et al., 2016; Bassi and Nansamba, 2017; Pallais, 2014). Related research shows that firms use apprenticeship programmes or wage subsidies as screening mechanisms to identify high-skilled workers (Hardy and McCasland, 2017; Levinsohn et al., 2013). Similarly, Alfonsi et al. (2017) argue that vocational training helps workseekers more than on-the-job training because it carries a credential employers can observe.

Another relevant strand of the literature studies firm-level hiring behavior and finds evidence consistent with information frictions. Specifically, firms using skills assessments or algorithmic recommendations to make hiring decisions can raise fill rates, raise productivity and lower turnover (Autor and Scarborough, 2008; Hoffman et al., 2015; Horton, 2017). This helps motivate the idea that information frictions can reduce the quality of firm-worker matches in addition to any employment effects.

We contribute to the literature in three ways. First, we separately vary workseekers' information about their own skills and their ability to signal this to firms, which allows us to separate supply- and demand-side information frictions. Existing work, other than Bassi and Nansamba (2017), studies only one side of the market. Second, we observe and measure the full decision process for workseekers, from job search through hiring to retention. Existing work studies downstream decisions and may miss the effect of information frictions on upstream decisions. For example, studying information friction at the employment margin misses effects of information frictions on search behavior.[6] Third, we collect more detailed measures of job search and employment quality for a larger sample than most research in this area. This allows us to study additional margins,

[6]For example, Bassi and Nansamba (2017) study the effect of supply- and demand-side information frictions in interviews. This means they cannot speak job search, application, posting, or interview invitations.

particularly job characteristics such as hours and earnings.

With these advances, we provide a more comprehensive description of labor markets with information frictions. Our finding that demand-side information frictions can be quantitatively important can help to inform the design of policies to improve labor market functioning. The class of policies we study are relatively cheap and easy to implement. Most other policies which try to reduce youth unemployment, such as wage subsidies, vocational training, entrepreneurship training, and public works programs are either expensive or do not robustly increase employment (Blattman and Ralson, 2015; Card et al., 2015; McKenzie, 2017).

We work in a developing country context with high unemployment and with young workseekers. We argue this is a particularly salient context and population to study. Formal education is weakly correlated with measured skills and there are many small firms that cannot afford complex in-house skills assessment systems (Hsieh and Olken, 2014; Pritchett, 2013). Information frictions may also have particularly serious impacts for young workseekers, who cannot use past work experience and reference letters as signals of employment (World Bank, 2018). We do not claim our results generalize to all labor markets. But the indirect evidence of information frictions documented in the US suggests that our results are relevant even in markets with lower unemployment.

We divide this paper into five substantive sections. We present a simple conceptual description of a labor market with limited information in section 2. We describe our randomized controlled trial in section 3. We lay out the data and analytical methods in section 4. We report results in section 5 and focus on the relationship between the level of skill and information frictions in section 6.

2 Conceptual Framework

We consider a labor market facing *two-sided limited information*: neither workseekers nor firms perfectly observe workseekers' skills. We argue that these information frictions can distort job search decisions and job offer decisions, ultimately distorting labor supply, labor demand, workseekers' earnings, and firms' revenue. In this draft of the paper we analyze the supply and demand sides of the labor market separately and do not yet introduce general equilibrium considerations.

Assume each workseeker i has fixed skill S_i and fixed time endowment T_i.[7] She allocates time between search for formal work W_i and other activities O_i. Job search yields utility $U^W(S_i, W_i)$ while time spent on other activities yields utility $U^O(S_i, O_i)$. These utility functions include pecuniary and non-pecuniary benefits and costs such as wages, monetary costs of job search, and leisure.[8] In a full information world, she chooses W_i^* and $O_i^* = T_i - W_i^*$ to equate the marginal

[7] We focus on single-dimensional skills. But the core features of the model are unchanged if skills are multidimensional, provided at least one dimension is imperfectly observed.

[8] The utility function $U^W(\cdot, \cdot)$ can be interpreted as the reduced-form of three structural functions: a job production function that maps skills and search time into a probability of securing a formal job, a wage function that maps skills

return from job search and the marginal return from other activities

$$\frac{\partial U^W\left(S_i, W_i^*(S_i)\right)}{\partial W_i} = \frac{\partial U^O\left(S_i, T_i - W_i^*(S_i)\right)}{\partial W_i}. \tag{1}$$

$W_i^*(\cdot)$ will be a non-degenerate function of S_i if skill and time are non-separable in the utility functions $U^W(\cdot,\cdot)$ and/or $U^O(\cdot,\cdot)$. For example, the return to job search may higher for workseekers with higher skills, in which case the optimal time allocation to searching for formal work will be increasing in skill.

We now introduce supply-side information frictions. Workseekers observe a noisy proxy $\tilde{S}_i = F(S_i, \epsilon_i)$, where ϵ_i captures the information friction. Workseekers allocate time between job search and other activities based on $\tilde{S}_i$ but their payoffs from job search depend on S_i.[9] Hence, the new indifference condition is

$$\mathbb{E}\left[\frac{\partial U^W\left(S_i, W_i^*(S_i)\right)}{\partial W_i}\right] = \frac{\partial U^O\left(S_i, T_i - W_i^*(S_i)\right)}{\partial W_i} \tag{2}$$

where the expectation is taken over the distribution of ϵ_i. The optimal time allocations in conditions (1) and (2) will be equal if both utility functions are linear in skill and time. Otherwise, the time allocations will generally be different. If, for example, utility from job search is a concave function of time allocated to job search, then information frictions will lower the optimal time allocation to job search. This will in turn lower the formal employment rate and lower earnings.

This framework illustrates that the effect of information frictions will depend on the structure of the utility functions. The framework does not generate an unconditional prediction about which types of workseekers (in terms of skills) will be affected most by information frictions. The relative effects on workseekers with different skills depend on the nature of information friction, captured by $F(\cdot, \epsilon_i)$, and on the structure of the utility functions. Even when the information friction is identical for all workseekers, $\tilde{S}_i = S_i + \epsilon_i$ with $\mathbb{E}[\epsilon_i | S_i] = 0$, a nonlinear utility function means that the information friction may distort decisions differently at different levels of skills.

This structure still applies if skills are multidimensional, provded at least one skill is unobserved. This structure still applies if workseekers can search in multiple sectors or using multiple strategies: information frictions will then distort time allocations between search and other activities *and* between different types of search.

Labor demand is determined by the decisions of profit-maximizing firms.We assume firm f faces produces output Q_f using labor L_f and other inputs K_f with a constant elasticity of substitution (CES) production function and firm-specific technology A_f. Absent information frictions, firms'

into wages conditional on securing a formal job, and a utility function over wages. But this does not change the core ideas of the framework.

[9] With two-sided information frictions their payoff from job search will depend on $\tilde{\tilde{S}}_i = G(S_i, \nu_i)$, where ν_i captures the information friction facing the firm. The wedge between search-relevant and wage-relevant skills remains, except in the special case where firms and workseekers face identical information frictions.

optimal choice of labor satisfies $p \times \frac{\partial Q_f(L^*, K^*)}{\partial L} = w$, where w is the wage.[10] We model heterogeneous skills by specifying output as a function of effective labor $L^E = \sum_{i \in f} S_i$, which is the sum of the skills of all workers employed by the firm. Workers are still paid their marginal revenue product, so workers with higher skills receive higher wages. This model gives rise to a threshold rule, where each firm hires workers in a skill band that depends on the firm-specific technology A_f.

Demand-side information frictions occur when firms cannot perfectly observe workseekers' skills. The concavity of the CES production means that the expected marginal revenue product from each workseeker is lower than their expected skills:

$$\mathbb{E}\left[p \times \frac{\partial Q_f\left(L^{E*}, K^*\right)}{\partial L^{E*}}\right] < p \times \frac{\partial Q_f\left(\mathbb{E}\left[L^{E*}\right], K^*\right)}{\partial L^{E*}}. \tag{3}$$

If wages are completely flexible and all firms face information frictions, then wages will fall at each level of skills. If there is some floor on wages then the lower wages may not fully offset the fall in expected revenue and firms will reduce total labor demand.[11][12] Concavity of the production function is sufficient but not necessary to generate this result. The same result holds if firms are risk-averse, which may be an appropriate assumption for small firms that cannot easily insure against negative shocks from low-skilled workers damaging equipment or alienating customers.

This simple framework illustrates that information frictions on either the demand or supply side of the labor market can generate lower employment and earnings. The framework also illustrates an important distinction. Eliminating demand-side information frictions can change employment and earnings without any change in workseekers' behavior. Eliminating supply-side information frictions can change employment and earnings only by changing workseekers' behavior.[13] Our experiment and data collection are designed both to test for the existence of information frictions and to differentiate demand- and supply-side frictions.

This model matches many features of countries with high youth unemployment. Formal school qualifications are weakly correlated with measured skills and productivity (Fedderke, 2005; Pritchett, 2013; Soderbom & Teal, 2004) so workseekers do not learn about their skills during their schooling and cannot signal skills to firms through schooling qualifications. Firms rely more on other signals, such as reference letters and referrals, which exclude young workseekers who lack references from previous work and people without access to referral networks (Beaman et al., 2017;

[10] We assume firms have no market power in the labor or product market so p and w do not depend on the firm's input choices. This assumption does not affect the core insights of the framework.

[11] This wage floor may reflect minimum wages or positive reservation wages from workseekers' outside options. Even a wage floor at zero may distort hiring decisions if information frictions are very large and there is a positive probability that some workseekers will have negative marginal revenue products.

[12] This framwork can be extended to a multiperiod model where firms observe workers' skills after hiring them, as in Altonji and Pierret (2001). As long as revelation is not instantaneous or firms incur hiring or firing costs, the predictions of the framework are unchanged.

[13] Labor supply can also be distorted if both firms and workseekers perfectly observe skills but workseekers *incorrectly believe* that firms face information frictions. We do not pursue this case in the current paper draft.

Calvo-Armengol & Jackson, 2004; Dustmann et al., 2015; Ioannides & Loury, 2004). Small and new firms, which cannot easily afford complex workseeker assessment systems, risk making mistakes during hiring and hence grow more slowly and struggle to challenge incumbents (Hsieh & Olken, 2014).

3 Experiment and Setting

We implement a three-arm randomized controlled trial with 6,895 workseekers in South Africa's Gauteng province. The trial is directly designed to test predictions from the conceptual framework in section 2.

We work with the Harambee Youth Employment Accelerator, a large South African NGO that provides assessment and placement services to 200 large corporates, to recruit participants and conduct the skills assessments. Harambee recruits young workseekers, assesses them, and selects a fraction of assessed workseekers for further training and placement at corporate partners. Workseekers are eligible for recruitment and assessment if they are age 18-29, have no criminal record, are South African citizens or have legal permission to work in South Africa, have graduated from a high school in a disadvantaged neighborhood, have not been formally employed in the past 12 months, and are actively searching for work. The last three criteria are intended to target people who do not have strong employment prospects through channels like university education and parents' connections but are difficult to verify and are unlikely to be strictly enforced. Harambee recruits workseekers through advertising in traditional and social media, referrals, and some door-to-door recruitment. In practice, our sample will consist of young, unemployed and underemployed high school graduates from low-income homes. Our sample will exclude people with less than high school, discouraged workseekers, and people aged 30 or older.[14] Our baseline survey is designed to allow us to compare the distribution of age, gender, education, employment status, and job search in our sample to the population of the Gauteng province.

We conduct six skills assessments with workseekers (three cognitive and three non-cognitive) and randomly divide workseekers into four groups:

1. A *control group* of 2,276 workseekers, who receive no feedback on their assessment results.

2. A *private treatment group* of 2,116 workseekers, who receive feedback on their assessment results in two ways: a roughly one-hour group briefing with a psychologist and printed reports showing their assessment results and briefly explaining the assessments. These reports contain no identifying information about the workseeker and no branding (figure 1). This treatment

[14] Workseekers are still eligible to participate if they have completed tertiary education. Many people in our sample have completed some tertiary education, such as short certificate course. But very few have a completed university degree.

is designed to improve workseekers' information about their own skills and hence test for supply-side information frictions.

3. A *public treatment group* of 2,248 workseekers, who receive the same treatment as the private group and 30 color-printed reports that contain the workseeker's name and national identity number and the brand of the agency and the World Bank (figure 2). This treatment is designed to improve workseekers' information about their own skills and their ability to inform prospective employers about these skills by sharing the reports and hence test for both supply- and demand-side information frictions. Workseekers in this group can scan or photocopy their reports and contact the implementing agency to receive additional reports. They can also direct prospective employers to https://www.assessmentreport.info/, which provides more information about the skills assessments, sample questions, and research linking these assessments to workplace productivity.

4. A *placebo treatment group* of 255 workseekers, who receive the same reports as members of the public group but with their own assessment results omitted (figure 3). This treatment is designed to test if any public treatment effect is driven by the information in the skills assessments or some other feature of the reports. For example, employers may be more likely to hire workseekers with reports because the branding makes applications more salient or because just attending the assessment process conveys a positive signal. This treatment group is deliberately small because it is designed to test a competing mechanism rather than a core prediction of the conceptual framework. We do not discuss this treatment group in this draft of the paper.

All reports place workseekers in terciles, relative to the population of assessed workseekers, rather than reporting their absolute scores. We piloted versions of the report with only absolute scores and only relative scores. The absolute scores provided readers with less information, as they could not easily anchor absolute scores to real outcomes. Providing both the absolute and relative scores led to complicated reports that overwhelmed and confused users.

We assess workseekers' skills in three cognitive and three noncognitive domains. Detailed information on all six assessments is available at https://www.assessmentreport.info, including sample questions.

1. *Numeracy*, focusing on practical arithmetic and pattern recognition. We calculate a single numeracy score using the inverse variance-weighted average of two numeracy assessment scores. One assessment is developed by a South Africa-based adult education agency and is registered with the South African Qualifications Authority. The other assessment is developed by a large retail chain and used in their applicant screening process.

Figure 1: Sample Private Report

REPORT ON CANDIDATE COMPETENCIES
-Personal Copy-

This report contains results from the assessments you took at Harambee in Phase 1 and Phase 2. These results can help you learn about some of your strengths and weaknesses and inform your job search.

You completed assessments on English Communication (listening, reading and comprehension) and Numeracy today in Phase 2. In Phase 1, you completed a Concept Formation assessment which asked you to identify patterns.

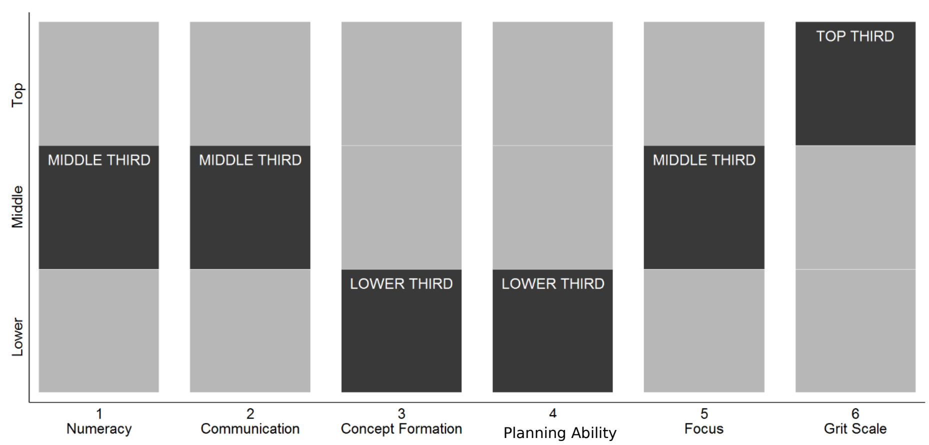
1. The Numeracy tests measure various maths abilities. Your score is the average of the two maths tests you did today at Harambee.
2. The Communication test measures English language ability through listening, reading and comprehension.
3. The Concept Formation test measures the ability to understand and solve problems. Candidates with high scores can generally solve complex problems, while lower scores show an ability to solve less complex problems.

You also did some games and questionnaires to measure your soft skills:

4. The Planning Ability Test measures how you plan your actions in multi-step problems. Candidates with high scores generally plan one or more steps ahead in solving complex problems.
5. The Focus Test looks at your ability to pick out which information is important in confusing environments. Candidates with high scores are able to focus on tasks in distracting situations.
6. The Grit Scale measures candidates' determination when working on difficult problems. Candidates with high scores spend more time working on the problems rather than choosing to pursue different problems.

Your results have been compared to a large group of young South African job seekers who have a matric certificate, are from socially disadvantaged backgrounds and have been assessed by Harambee.

You scored in the MIDDLE THIRD of candidates assessed by Harambee for Numeracy, MIDDLE THIRD for Communication, LOWER THIRD for Concept Formation, LOWER THIRD for Planning Ability, MIDDLE THIRD for Focus and TOP THIRD for the Grit Scale.



DISCLAIMER
Please note that this is a confidential assessment report and is intended for use by the person specified above. Assessment results are not infallible and may not be entirely accurate.

Notes: This figure shows an example of the reports given to workseekers in the private treatment arm. The reports contain the workseeker's assessment results but no identifying information and no branding.

Figure 2: Sample Public Report

THE WORLD BANK

harambee
YOUTH EMPLOYMENT ACCELERATOR
WORK FOR WORK

REPORT ON CANDIDATE COMPETENCIES

name.. surname..
ID No. id..

This report provides information on assessments conducted by Harambee Youth Employment Accelerator (harambee.co.za), a South African organisation that connects employers looking for entry-level talent to young, high-potential work-seekers with a matric or equivalent. Harambee has conducted more than 1 million assessments and placed candidates with over 250 top companies in retail, hospitality, financial services and other sectors. Assessments are designed by psychologists and predict candidates' productivity and success in the workplace. This report was designed and funded in collaboration with the World Bank. You can find more information about this report, the assessments and contact details at www.assessmentreport.info. «name» was assessed at Harambee on 13 September, 2016.

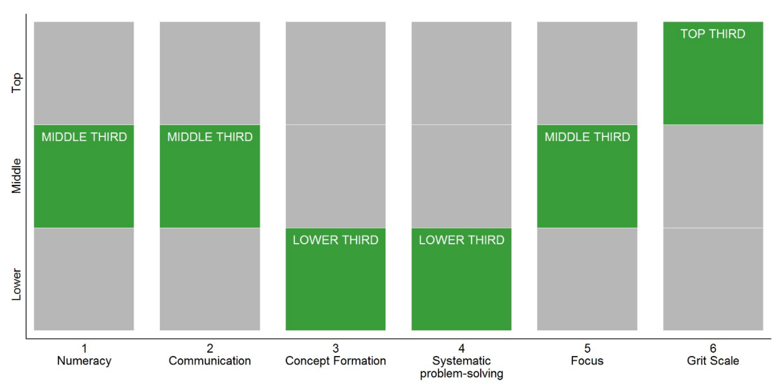
«**name**» completed assessments on English Communication (listening, reading, comprehension), Numeracy, and Concept Formation:

1. The Numeracy tests measure candidates' ability to apply numerical concepts at a National Qualifications Framework (NQF) level, such as working with fractions, ratios, money, percentages and units, and performing calculations with time and area. This score is an average of two numeracy tests the candidate completed.
2. The Communication test measures a candidate's grasp of the English language through listening, reading and comprehension. It assesses at an NQF level, for example measuring the ability to recognise and recall literal and non-literal text.
3. The Concept Formation Test is a non-verbal measure that evaluates candidates' ability to understand and solve problems. Those with high scores are generally able to solve complex problems, while lower scores indicate an ability to solve less complex problems.

«**name**» also completed tasks and questionnaires to assess their soft skills:

4. The Planning Ability Test measures how candidates plan their actions in multi-step problems. Candidates with high scores generally plan one or more steps ahead in solving complex problems.
5. The Focus Test assesses a candidate's ability to distinguish relevant from irrelevant information in potentially confusing environments. Candidates with high scores are generally able to focus on tasks in distracting surroundings, while candidates with lower scores are more easily distracted by irrelevant information.
6. The Grit Scale measures whether candidates show determination when working on challenging problems. Those with high scores generally spend more time working on challenging problems, while those with low scores choose to pursue different problems.

«name»'s results have been compared to a large benchmark group of young (age 18-34) South Africans assessed by Harambee. All candidates have a matric certificate and are from socially disadvantaged backgrounds. The benchmark group is 5,000 for cognitive skills and 400 for soft skills.
«name» scored in the «tercile_num» THIRD of candidates assessed by Harambee for Numeracy, «tercile_lit» THIRD for Communication, «tercile_cft» THIRD for Concept Formation, «tercile_tol» THIRD for Planning Ability, «tercile_troop» THIRD for Focus and «tercile_grit» THIRD for the Grit Scale.



DISCLAIMER: This is a confidential assessment report for use by the person specified above. The information in the report should only be disclosed on a "need to know basis" with the prior understanding of the candidate. Assessment results are not infallible and may not be entirely accurate. Best practice indicates that any organisation's career management decisions should depend on factors in addition to these assessment results. Harambee cannot accept responsibility for decisions made based on the information contained in this report and cannot be held liable for the consequences of those decisions.

Notes: This figure shows an example of the reports given to workseekers in the public treatment arm. The reports contain the assessment results, the workseekers's name and national identity number, and the logo of the World Bank and the implementing agency. Each workseeker received 20 of these reports and guidelines on how to request more reports.

Figure 3: Sample Placebo Report

THE WORLD BANK

harambee
YOUTH EMPLOYMENT ACCELERATOR
— WORK FOR WORK —

REPORT ON ASSESSMENT PROCESS

name.. surname..
ID No. id..

This report provides information on assessments conducted by Harambee Youth Employment Accelerator (harambee.co.za), a South African organisation that connects employers looking for entry-level talent to young, high-potential work-seekers with a matric or equivalent. Harambee has conducted more than 1 million assessments and placed candidates with over 250 top companies in retail, hospitality, financial services and other sectors. Assessments are designed by psychologists and predict candidates' productivity and success in the workplace. This report was designed and funded in collaboration with the World Bank. You can find more information about this report, the assessments and contact details at www.assessmentreport.info. «name» was assessed at Harambee on «date».

«**name**» completed assessments on English Communication (listening, reading, comprehension), Numeracy, and Concept Formation:

1. The Numeracy tests measure candidates' ability to apply numerical concepts at a National Qualifications Framework (NQF) level, such as working with fractions, ratios, money, percentages and units, and performing calculations with time and area. This score is an average of two numeracy tests the candidate completed.
2. The Communication test measures a candidate's grasp of the English language through listening, reading and comprehension. It assesses at an NQF level, for example measuring the ability to recognise and recall literal and non-literal text.
3. The Concept Formation Test is a non-verbal measure that evaluates candidates' ability to understand and solve problems. Those with high scores are generally able to solve complex problems, while lower scores indicate an ability to solve less complex problems.

«**name**» also completed tasks and questionnaires to assess their soft skills:

4. The Planning Ability Test measures how candidates plan their actions in multi-step problems. Candidates with high scores generally plan one or more steps ahead in solving complex problems.
5. The Focus Test assesses a candidate's ability to distinguish relevant from irrelevant information in potentially confusing environments. Candidates with high scores are generally able to focus on tasks in distracting surroundings, while candidates with lower scores are more easily distracted by irrelevant information.
6. The Grit Scale measures whether candidates show determination when working on challenging problems. Those with high scores generally spend more time working on challenging problems, while those with low scores choose to pursue different problems.

DISCLAIMER: This is a confidential assessment report for use by the person specified above. The information in the report should only be disclosed on a "need to know basis" with the prior understanding of the candidate. Harambee cannot accept responsibility for decisions made based on the information contained in this report and cannot be held liable for the consequences of those decisions.

Notes: This figure shows an example of the reports given to workseekers in the placebo treatment arm. The reports contain no assessment results, the workseekers's name and national identity number, and the logo of the World Bank and the implementing agency. Each workseeker received 20 of these reports and guidelines on how to request more reports.

2. *Literacy and communication*, focusing on comprehension of spoken and written passages. The assessment is developed by a South Africa-based adult education agency and is registered with the South African Qualifications Authority.

3. *Concept formation*, which measures fluid intelligence and has a similar design to the Ravens Progressive Matrices test. The assessment is developed by a South Africa-based adult education agency and is registered with the Health Professionals Council of South Africa.

4. *Grit*, a self-reported measure of how much determination candidates show when they work on challenging problems. Our assessment is adapted from the survey-based measure discussed in Duckworth et al. (2007).

5. *Focus*, a measure of a candidate's ability to distinguish relevant from irrelevant information in potentially confusing environments. Our assessment is a shortened and computerized version of the widely-used Stroop Test (Stroop, 1935).

6. *Planning*, a measure of how candidates behave when faced with complex, multi-step problems. Our assessment is adapted from a test proposed by Gneezy et al. (2010).

The three cognitive assessments were developed and validated by South African psychologists. We developed the three noncognitive assessments by piloting measures used by psychologists and behavioral economists in other contexts, administering the assessments with a sample of more than 100 respondents from our target population, verifying that the individual items are appropriately positively correlated with each other, and shortening the test by selecting the items that best predict the average score.[15] All assessments are conducted on desktop computers, so the assessment results will be driven in part by candidates' computer skills.

This design yields six cardinal skills measures. We divide candidates into terciles separately for each skills measure and display these terciles on the reports. We use both the terciles and cardinal scores in analysis, which we discuss in section 4.

We interact with workseekers over multiple stages:

1. Workseekers apply on date T_1 to be assessed by Harambee, after being recruited.

2. Workseekers are invited to a first assessment on date T_2. They complete the concept formation assessment, a career placement tool (see https://www.shadowmatch.com for details), and receive some basic job search training (e.g. how to write a CV). Time $T_2 - T_1$ varies across workseekers and may be several months. Surveys are completed at a computer workstation.

[15] For the first 23% of the workseekers we assessed, we used self-reported measures of *control* and *flexibility* instead of the focus and planning assessments. The control scale is a self-reported measure of how candidates react when confronted with new and confusing problems. The flexibility scale is a self-reported measure of how candidates tackle problems with several potential solution approaches. We adapted both scales from an instrument developed by Hepner and Petersen (1982) and validated them for our target population.

Candidates receive some basic training on how to use a mouse and keyboard, but the computer skills required are very minimal.

3. Workseekers are invited to a second assessment on date T_3. They complete the other five assessments and a self-administered baseline survey completed at a computer workstation. This asks questions about their demographics, employment, job search, beliefs about their skills, self-esteem, beliefs about the labor market returns to skills and search effort, risk preferences, and time preferences. Time $T_3 - T_2$ varies across workseekers and may be several months.

4. At the end of day T_3, workseekers receive any treatment to which they have been assigned.

5. On day $T_3 + 2$ or $T_3 + 3$, workseekers are invited to participate in a short text message survey. This surveys asks about perceived numeracy skills and their self-esteem.

6. Starting 2 months after T_3, workseekers complete a detailed follow-up survey by phone. This survey asks about their current and past employment, job search, beliefs about their skills, and beliefs about the labor market returns to search effort. The date that we first attempt to survey workseekers is randomized between 2 and 4 months.

7. Starting 8 months after T_3, workseekers complete a detailed follow-up survey by phone. This survey asks about their current and past employment, job search, beliefs about their skills, and beliefs about the labor market returns to search effort. The date that we first attempt to survey workseekers is randomized between 8 and 12 months.

The multiple rounds of follow-up surveys are designed to track treatment effects over one year. The time lags between T_1, T_2, and T_3 are determined by Harambee's operations. The assessment dates are based on the enterprise's perception about labor demand from corporate partners in specific neighborhoods. The time lags will be smaller if workseekers apply during periods of high labor demand or live in neighborhoods where there is high labor demand. This recruitment schedule means that some workseekers will complete stage 1 but not proceed to stage 2 or stage 3. Hence our sample at T_3, relative to T_1, is likely to contain fewer workseekers who are at the margin of employment and unemployment (they get jobs in the meantime) and who are at the margin between searching and discouragement (they withdraw from search). Our randomization is based on workseekers who reach stage 3 so this attrition may affect the generalizability of our results but not their internal validity.

We implemented the treatment for 84 days between 22 September 2016 and 13 April 2017. We conducted the first phone follow-up survey between 30 November 2016 and 5 September 2017. The second phone follow-up survey began in August 2017 and will finish in early 2018.

Table 1: Balance Test Statistics

Baseline variable	F-test statistic	p-value From clustered SEs	p-value From permutation test
Hours searched in past 7 days	0.01	0.988	0.98
Worked in past 7 days	1.75	0.181	0.226
Male	1.47	0.236	0.306
Grit assessment score	2.91	0.061	0.096
Flexibility assessment score	1.02	0.382	0.55
Control assessment score	2.58	0.107	0.232
Concept formation score	0.00	0.999	0.996
Literacy assessment score	0.83	0.442	0.504
Numeracy assessment score	0.51	0.602	0.64
Self-esteem scale	2.51	0.088	0.124

Notes: This table shows balance tests constructed by regressing each of the selected baseline characteristics on indicators for the public and private treatment groups and testing if these coefficients are jointly zero. The tests are conducted using heteroskedasticity-robust standard errors clustered by treatment date and by permuting the treatment assignments within blocks.

Harambee invites some workseekers after date T_3 for further training and placement at a corporate partner. These workseekers are selected based on strong performance in the concept formation and numeracy assessments and their results in the career placement tool. Only a small fraction of workseekers are invited for further training and placement and these invitations are independent of treatment assignment.

We randomize treatment at the day level using a sequential blocking algorithm:

1. We divide the first 18 days of the trial intro 2 blocks of 9 days. We then randomly assign 3 days from each block to control, private treatment, and pubic treatment.

2. Near the end of block 2, we repeat the randomization process for a third block. We then repeat this process near the end of blocks 3 to 9 for blocks 4 to 10.

3. If we cannot implement one day of the intervention (e.g. due to power failures or network problems that mean the baseline cannot be administered) in block B, then we add another day to block B+1. We then randomly assign one extra day from block B+1 to the same treatment status as the dropped day from block B.

The first 7 blocks each contain 9 or 10 days. The last 3 blocks contain 6, 13, and 7 days as the exact end date of the intervention was subject to ongoing negotiation with Harambee. All placebo treatment days are in block 9. Table 1 shows that the treatment groups are balanced on key baseline characteristics.

4 Data and Analytical Methods

We collect data in five stages: skills assessments, computerized in-person baseline survey, text message survey 2 days after treatment, phone survey 2-5 months after treatment, phone survey 8-12 months after treatment. We report summary statistics for key baseline and follow-up variables in table 2.

The attrition rates for the text message and first phone surveys are respectively 18.6 and 4.1 percent. Table 3 shows that attrition is balanced across treatment arms. The largest pairwise difference in attrition between treatment groups is 1.1 percent for the text message survey and 0.4 percent for the first phone survey. None of the differences are statistically significant. Table 4 show that attrition is higher for male respondents and, in the text message survey, negatively related to numeracy and concept formation skills. These differences by baseline characteristics are jointly significant but not large: the baseline characteristics explain less than 2 percent of the variation in attrition. We conclude that attrition is not a threat to identification of the average treatment effects.

We also report treatment effects on indices that summarize the outcomes within each family following Anderson (2008). To construct the indices, we standardize each primary outcome in primary family to have mean zero and standard deviation one, recode it if necessary to standardize the "direction" of the outcome, and replace missing values with zeroes. We then calculate a weighted average of the recoded, demeaned outcomes in each family, where the weights equal the diagonal elements of the inverse covariance matrix. This gives us seven indices: beliefs about own skills, beliefs about search returns to search effort, beliefs about future earnings, search effort, search effectiveness, employment status, and employment quality. Each index has mean zero but their standard deviations vary; indices whose components are more strongly correlated will have lower variance.

We report treatment effects from three models. We first estimate average treatment effects using the regression model

$$Y_{id} = Private_{id} \cdot \beta_1 + Public_{id} \cdot \beta_2 + \mathbf{S}_d + \mathbf{X}_{id} \cdot \mu + \epsilon_{id} \tag{4}$$

for workseeker i who completed the intervention on date d. $Private_{id}$ and $Public_{id}$ are treatment group indicators, $\mathbf{S}_{id}$ is a vector of randomization block fixed effects, and $\mathbf{X}_{id}$ is a vector of prespecified control variables from the baseline survey and assessment process.[16] We report

[16]We prespecified baseline controls for numeracy score, literacy score, concept formation score, grit score, first principal component of the other two non-cognitive assessment scores, expected numeracy score, expected tercile of the literacy score, expected tercile of the concept formation score, score on the 7-point self-esteem scale, age, gender, an indicator for having worked in the previous week, and an indicator for having education above the sample median. All control variables are measured in the baseline survey or assessment process. We selected these controls from a larger list by running a LASSO algorithm for each pre-specified outcome and including all controls that were selected by the LASSO for at least 50 percent of the outcome measures (Tibshirani, 1996). Education, cognitive

Table 2: Summary Statistics for Selected Baseline and Endline Variables

Baseline Variable	# observations	Mean	Std dev.	Minimum	10^{th} pctile	90^{th} pctile	Maximum
Age	6895	23.649	3.301	18.001	19.8	28.3	35.277
Male	6895	0.382	0.486	0	0	0	1
Completed diploma or degree	6895	0.167	0.373	0	0	0	1
Completed post-high school certificate	6895	0.212	0.409	0	0	0	1
Completed high school	6895	0.61	0.488	0	0	0	1
Numeracy assessment score	6895	0	1	-2.536	-1.25	1.38	3.045
Literacy/communications assessment score	6895	0	1	-3.258	-1.15	1.66	3.061
Concept formation assessment score	6895	0	1	-2.697	-1.58	1.22	2.345
Grit assessment score	6895	0	1	-4.779	-1.35	1.26	2.074
Average score on other two noncognitive assessments	6895	0	1.071	-5.605	-1.3	1.32	2.958
Expected numeracy tercile equals actual tercile	6811	0.382	0.486	0	0	0	1
Expected literacy/communications tercile equals actual tercile	6811	0.382	0.486	0	0	0	1
Expected concept formation tercile equals actual tercile	6811	0.382	0.486	0	0	0	1
Self-esteem scale (1-7 points)	6895	0	1	-3.624	-1.37	1.23	1.408
# of job applications submitted in past 30 days	6819	9.295	12.92	0	2	20	90
Hours spent searching for work in past 7 days *	6703	16.717	19.595	0	2	48	96
Money spent searching for work in past 7 days *	6150	188.774	214.494	0	30	400	1500
# of job applications expected to submit in next 30 days	6844	17.673	24.808	0	4	36	180
Worked in past 7 days	6895	0.378	0.485	0	0	0	1
Total earnings from all jobs in past 7 days *	2117	561.501	715.813	0	100	1500	4000
Hourly wages from all jobs in past 7 days *	2077	64.552	85.421	0	7.5	167	500

Notes: This table shows summary statistics for selected baseline variables. Variables marked with * have been winsorized at the 99^{th} percentile.

Endline Variable	# observations	Mean	Std dev.	Minimum	10^{th} pctile	90^{th} pctile	Maximum
Expected grit tercile equals actual tercile	6598	0.454	0.498	0	0	0	1
Expected literacy tercile equals actual tercile	6603	0.513	0.5	0	0	0	1
Expected concept formation tercile equals actual tercile	6603	0.4	0.49	0	0	0	1
Expected numeracy tercile equals actual tercile	6605	0.531	0.499	0	0	0	1
# of cognitive assessments where expected tercile equals actual tercile	6610	1.442	0.988	0	0	3	3
# of cognitive assessments where expected tercile is higher than actual tercile	6610	1.201	0.994	0	0	3	3
# of cognitive assessments where expected tercile is lower than actual tercile	6610	0.354	0.593	0	0	1	3
# of noncognitive assessments where expected tercile equals actual tercile	6611	1.421	0.954	0	0	3	3
# of noncognitive assessments where expected tercile is higher than actual tercile	6611	0.868	0.836	0	0	2	3
# of noncognitive assessments where expected tercile is lower than actual tercile	6611	0.702	0.75	0	0	2	3
Self-esteem scale	6613	-0.076	0.86	-3.624	-1.19	0.887	1.408
Any job search activity in last 30 days	6612	0.918	0.274	0	0	0	1
Any job search activity in last 7 days	6612	0.692	0.462	0	0	0	1
# job applications submitted in last 30 days	6581	12.26	17.572	0	1	27	120
Hours spent searching for work in past 7 days *	6605	9.808	13.445	0	0	25	72
Money spent searching for work in past 7 days *	6603	112.64	138.836	0	0	300	700
Worked in past 7 days	6611	0.323	0.468	0	0	0	1
Worked since treatment	6611	0.69	0.462	0	0	0	1
Worked in first calendar month after treatment	6608	0.482	0.5	0	0	0	1
Worked in second calendar month after treatment	6611	0.454	0.498	0	0	0	1
Hours worked for all jobs in past 7 days *	2117	28.377	21.119	0	4	56	91
Total earnings from all jobs in past 7 days *	1630	704.608	693.067	0	50	1500	3800
Hourly wages from all jobs in past 7 days *	1617	43.305	69.002	0	3.57	100	437.5
Have written permanent job contract	2101	0.035	0.183	0	0	0	1
Wage employment	2101	0.563	0.496	0	0	0	1
Self employment	2101	0.251	0.434	0	0	0	1

Notes: This table shows summary statistics for selected endline variables. Variables marked with * have been winsorized at the 99^{th} percentile.

Table 3: Post-Treatment Survey Attrition Rate by Treatment Group

	(1) Missed Text Message Survey	(2) Missed First Phone Survey
Control group	0.170	0.040
	(0.013)	(0.006)
Public treatment group	0.177	0.039
	(0.011)	(0.004)
Private treatment group	0.181	0.043
	(0.010)	(0.004)
p: Public = Private	0.794	0.392
p: Public = Control	0.667	0.857
p: Private = Control	0.484	0.635
p: Public = Private = Control	0.782	0.686
# observations	6640	6640
# clusters	81	81

Notes: This table shows the fraction of each treatment group that completes the text message survey 2-3 days after treatment and the first phone survey 2-5 months after treatment. Respondents who end a survey before completing it are defined as non-completers. Heteroskdasticity-robust standard errors clustered by treatment date are shown in parentheses.

heteroskedasticity-robust standard errors clustered by intervention date, the unit at which treatment is assigned. Building on the conceptual framework in section 2, we test if the average effects of the public and private treatment are separately equal to zero ($\beta_1 = 0$ and $\beta_2 = 0$), equal to each other ($\beta_1 = \beta_2$), and jointly equal to zero ($\beta_1 = \beta_2 = 0$). We report p-values for these tests using the cluster- and heteroskedasticity-robust standard errors. We also report the corresponding q-values that control the family-wise error rate across multiple outcomes within each outcome family, following Benjamini et al. (2006).

5 Treatment Effects

5.1 Summary of Results

Table 5.1 reports the public and private treatment effects on all seven outcome indices.[17] The public treatment increases the employment quantity index by 0.13 standard deviations, while the private treatment has smaller and insignificant effect of 0.05 standard deviations (column 7). The two effects are close to being statistically different (p=0.108). Neither treatment has effects on the employment quality index (column 8): as we discuss below, earnings are higher in both treatment

and non-cognitive test scores, grit, and baseline beliefs about scores were selected frequently across all categories of endline variable. Baseline self-esteem was most predictive of endline beliefs about wages and job search and of endline self-esteem. Age was most predictive of job search behavior and endline beliefs. Gender was most predictive of employment outcomes and beliefs about wages. Employment status was most predictive of endline employment and beliefs about job search and about the future.

[17]We report only results with prespecified controls for the sake of brevity. As in table 5.1, including or excluding the pre-specified controls makes little difference to these results.

Table 4: Post-Treatment Survey Attrition Rate by Baseline Characteristics

	(1) Missed Text Message Survey	(2) Missed First Phone Survey
Baseline numeracy score	-0.031	0.002
	(0.006)	(0.003)
Baseline literacy score	0.008	0.003
	(0.006)	(0.003)
Baseline concept formation score	-0.019	0.004
	(0.006)	(0.003)
Baseline grit score	-0.000	-0.007
	(0.006)	(0.003)
Baseline score on remaining noncog. assessments	0.000	-0.001
	(0.004)	(0.003)
Baseline belief about numeracy score	-0.000	-0.000
	(0.000)	(0.000)
Baseline belief about literacy score	0.016	-0.003
	(0.010)	(0.005)
Baseline belief about concept formation score	0.008	-0.002
	(0.009)	(0.005)
Baseline self-esteem index	0.005	0.002
	(0.004)	(0.002)
Completed high school	-0.011	-0.005
	(0.012)	(0.005)
Age	-0.002	0.001
	(0.001)	(0.001)
Male	0.049	0.012
	(0.010)	(0.005)
Worked in last 7 days	-0.001	-0.001
	(0.009)	(0.005)
Constant	0.180	0.021
	(0.057)	(0.029)
F: All coefficients jointly zero	6.89	1.86
p: All coefficients jointly zero	0.000	0.049
Adj R2	0.015	0.004
# observations	6380	6380
# clusters	79	79

Notes: This table shows the relationship between pre-specified baseline covariates and completion of the text message survey 2-3 days after treatment and the first phone survey 2-5 months after treatment. Respondents who end a survey before completing it are defined as non-completers. Heteroskdasticity-robust standard errors clustered by treatment date are shown in parentheses.

groups than in the control group, but the treated groups are no more likely to have permanent or written contracts. Broadly, the result suggests that demand-side information frictions may distort hiring decisions and reduce employment and earnings. If firms are provided with better information on a given workseeker's abilities, potentially reducing uncertainty about the workseekers; likely skills or performance, they are more likely to employ the worker. In contrast, supply-side information frictions result in more limited distortions in workseekers' job search decisions: simply making workseekers better informed about their own ability has limited effects on the likelihood they find employment.

Table 5: Effects of Public and Private Treatments on Outcome Indices

	Beliefs about ...			Search ...		Employment ...	
	Skill	Search	Earnings	Effort	Effective	Status	Quality
Public	0.706***	0.079**	0.002	0.011	0.030	0.132***	0.018
	(0.036)	(0.025)	(0.029)	(0.031)	(0.034)	(0.025)	(0.027)
Private	0.558***	0.045*	0.012	0.013	0.031	0.047	0.014
	(0.033)	(0.026)	(0.027)	(0.033)	(0.031)	(0.029)	(0.025)
# observations	6367	6367	6367	6367	6367	6367	6367
# clusters	81	81	81	81	81	81	81
R-squared	0.293	0.055	0.306	0.083	0.046	0.115	0.015
Control mean	0.000	0.000	-0.000	-0.000	0.000	0.000	0.000
p: Pub = 0	0.000	0.003	0.950	0.739	0.369	0.000	0.515
p: Pvt = 0	0.000	0.091	0.655	0.698	0.311	0.108	0.563
p: Pub = Pvt	0.000	0.189	0.672	0.936	0.977	0.001	0.900
p: Pub = Pvt = 0	0.000	0.010	0.871	0.918	0.527	0.000	0.769

Outcomes are expressed in standard deviations of family-level indices (Anderson, 2008). All regressions include randomization block fixed effects and baseline values of the outcome; education; all assessment results; beliefs about numeracy, communications & concept formation results; age; gender; and employment status. Heteroskedasticity-robust standard errors clustered by treatment day are shown in parentheses. Lower rows of the table show p-values for linear combinations of coefficients.

We explore the mechanisms through which effects on employment occur. The larger public treatment effect on employment is consistent with two mechanisms: i) the public treatment has a larger effect on beliefs, which drives the change in employment, potentially through changes in workseekers' job search behavior or performance in interviews or ii) the public treatment may change firms' hiring decisions. Our experiment is not designed to cleanly separate these mechanisms.

We find evidence that some of the effect works through changes in beliefs and search behavior. Both treatments increase the accuracy of workseekers' beliefs about their own skills, by approximately 0.71 standard deviations for the public and 0.56 for the private treatment (columns 1 and 2). Both effects are precisely estimated and statistically significantly different to zero and each other. Workseekers believe the reports to be credible, are influenced by the information contained in them, and remember this information even three to four months after treatment.

Both treatments also slightly increase workseekers' perceived return to job search but the effects are smaller – 0.08 and 0.05 standard deviations for respectively the public and private treatments (column 3), which we interpret further below. Treatment effects on expected wages are small and insignificant, potentially because entry-level wages for workseekers with limited education and little work experience may be fairly fixed in this labor market regardless of workseeker ability. [18]

We find some evidence that these stronger changes in beliefs result in differential changes in search behavior between the public and private treatment groups. On average, in table 5.1, effects on search effort and effectiveness of search effort are small and insignificant on average and the coefficients are very similar in both treatment groups. However, for the public treatment group, the timing of our measures of search effort is not ideal and we have only an imperfect test of the effect of the treatment on search. By the time of endline, some treated workseekers have already found work and may have stopped searching altogether or be searching less as a result. Below, we show suggestive evidence that the public treatment did increase search effort or effectiveness for workseekers randomly surveyed sooner after treatment, compared to those surveyed longer after treatment. Effects on search decrease over time in this group as workseekers find work.

For the private treatment group, however, there are no strong employment effects, so we have a fairly good test of whether search effort or effectiveness changes. We find a similar pattern in search behavior to the public treatment group: the private treatment does increase search effort sooner after treatment, although it does not affect search effectiveness. Importantly, changes in search behavior do not affect employment, supporting our interpretation that supply-side information frictions are not binding constraints on employment. Workseekers' beliefs about their ability do alter job search behavior, but without an effective technology to signal their skills, these changes in search do not change likelihood of employment.

Importantly, the experiment does not establish that changes in particular beliefs cause changes in employment, or that changes in beliefs cause changes in search, or that changes in search cause changes in employment, without further assumptions. The experiment does show that changes beliefs and in behavior in the short run can be attributed relatively cleanly to exposure to the intervention, not to other changes in opportunities and constraints for treated workseekers. We use measures of other beliefs, such as about likely wages, to argue that some channels do not appear to be important. But otherwise, we rely on our model to suggest how these changes are related.

Importantly, we do not argue that all changes in employment arise solely through changes in workseeker behavior. We have some suggestive evidence that some of the effect on employment arises from changes firms' information sets as well as changing workseekers' beliefs and behavior. The ratio of changes in the employment quantity index from the public and private treatments,

[18] Some preliminary analysis not shown in the paper suggests that workseekers' expectations prove correct: although wages in both treatment groups increase, these arise because more workers are employed in sum, not because workers employed as a result of the treatment receive higher wages.

2.8, is much larger than the ratio of changes in the beliefs about own skills index, 1.3, and than the ratio of changes in search, 0.84.

The workseeker experiment is not ideally designed to identify the pure effect of alleviating pure demand-side information friction. The comparison of public and private arms gives the additional effect of alleviating pure demand-side information frictions, conditional on the supply-side decision to share information. In the public treatment group, workseekers may endogenously choose whether to share reports with firms, based both on their ability (as signalled by the reports) and their beliefs about whether firms will value the reports.[19]

We thus also conduct an audit study (currently in the field) where we submit multiple applications to approximately 1,200 jobs, randomly vary whether applications are accompanied by assessment reports and measure the share of applications that receive interview invitations and job offers. This cleanly identifies the effect of alleviating pure demand-side information frictions. Together, the two studies give a comprehensive description of how information frictions affect both sides of the labor market.

5.2 Effects on Employment

We discuss effects on the variables which make up the employment quantity index in table 6. The public treatment robustly increases employment on all dimensions in the employment quantity index. Respondents are 5 percentage points more likely to have worked in the 7 days before the phone survey and to have worked at any points since the baseline, off bases of 30 and 67 percent (columns 1 and 2). They also worked 16 percent more hours in the preceding 7 days. Workseekers appear to have found work relatively quickly after treatment: they are more likely to have worked in the first or second month after treatment (columns 5 and 6).[20] The private treatment effects are uniformly positive but substantially and often significantly smaller.

Table 6 shows effects on the components of the employment quality index. The positive public treatment effect on the employment quality index is largely driven by a 32 percent increase in total earnings in the preceding week and 20 percent increase in hourly wages (columns 1 and 2). Neither treatment shifts the share of workers with written contracts or permanent contracts (columns 3 and 4). There is suggestive evidence that the employment increase is driven by a 2.4 percentage point increase in wage employment rather than self-employment (columns 5 and 6): column 5 shows effects on a binary variable equal to one if the respondent has wage employment and zero if the respondent has self employment, family employment, or no employment; while in column 6 the

[19] For example, 73% of the top-achieving workseekers provide reports to employers, whereas only 55% of the poorer performers do so.

[20] Given differences in the length of time elapsing between baseline and endline, we only observe 3230 workseekers in the sample four months after treatment, enabling us to calculate effects on employment one, two and three months after treatment. For the remainder, endline is three months after baseline, which only enables us to calculate effects on employment after one and two months. The effects after three months are similar in size to those after one and two months but are not statistically significantly different from zero because the sample is smaller.

Table 6: Effects of Public and Private Treatments on Employment Quantity

	Worked …		Hours in	Worked in … month		
	past 7 d	since baseline	past 7 d	First	Second	Third
Public	0.052***	0.052***	1.339*	0.035***	0.058***	0.038
	(0.012)	(0.010)	(0.571)	(0.011)	(0.014)	(0.022)
Private	0.011	0.031*	0.525	0.027	0.009	0.010
	(0.012)	(0.013)	(0.513)	(0.013)	(0.015)	(0.021)
# observations	6365	6365	6356	6362	6365	3230
# clusters	81	81	81	81	81	81
R-squared	0.041	0.089	0.034	0.095	0.074	0.057
Control group mean	0.309	0.671	8.794	0.465	0.437	0.403
Control mean | employed			28.672			
p: Pub = 0	0.000	0.000	0.021	0.002	0.000	0.082
q: Pub = 0	0.001	0.001	0.014	0.002	0.001	0.031
p: Pvt = 0	0.348	0.021	0.309	0.035	0.565	0.618
q: Pvt = 0	0.771	0.142	0.771	0.142	0.792	0.792
p: Pub = Pvt	0.002	0.097	0.152	0.457	0.001	0.210
q: Pub = Pvt	0.005	0.193	0.236	0.354	0.004	0.267
p: Pub = Pvt = 0	0.000	0.000	0.070	0.006	0.000	0.204
q: Pub = Pvt = 0	0.001	0.001	0.044	0.007	0.001	0.109

All regressions include randomization block fixed effects and baseline values of the outcome; education; all assessment results; beliefs about numeracy, communications & concept formation results; age; gender; and employment status. Heteroskedasticity-robust standard errors clustered by treatment day are shown in parentheses. Lower rows of the table show p- and q-values for linear combinations of coefficients. The q-values adjust for multiple testing over all outcomes in the same family following Benjamini *et al.* (2006).

outcome is equal to one if the respondent has self employment.

Table 7: Effects of Public and Private Treatments on Employment Quality

	In past week ...		Contract type ...		Employment type	
	earnings	wages	written	permanent	wage	self
Public	49.635***	1.782*	0.020	0.004	0.024*	0.010
	(17.892)	(0.830)	(0.010)	(0.003)	(0.012)	(0.008)
Private	43.855***	1.009	0.017	0.005	0.022	-0.012
	(18.183)	(1.013)	(0.009)	(0.003)	(0.011)	(0.007)
# observations	6347	6333	6333	6333	6333	6333
# clusters	81	81	81	81	81	81
R-squared	0.027	0.022	0.014	0.004	0.023	0.021
Control mean	155.163	8.731	0.120	0.009	0.172	0.081
Control mean | employed	504.858	28.645	0.392	0.029	0.562	0.265
p: Pub = 0	0.007	0.035	0.055	0.266	0.049	0.187
q: Pub = 0	0.116	0.211	0.211	0.500	0.211	0.454
p: Pvt = 0	0.018	0.322	0.072	0.148	0.043	0.080
q: Pvt = 0	0.158	0.210	0.158	0.177	0.158	0.158
p: Pub = Pvt	0.789	0.465	0.771	0.710	0.903	0.001
q: Pub = Pvt	1.000	1.000	1.000	1.000	1.000	0.014
p: Pub = Pvt = 0	0.007	0.104	0.091	0.307	0.069	0.004
q: Pub = Pvt = 0	0.038	0.167	0.166	0.241	0.152	0.038

All regressions include randomization block fixed effects and baseline values of the outcome; education; all assessment results; beliefs about numeracy, communications & concept formation results; age; gender; and employment status. Heteroskedasticity-robust standard errors clustered by treatment day are shown in parentheses. Lower rows of the table show p- and q-values for linear combinations of coefficients. The q-values adjust for multiple testing over all outcomes in the same family following Benjamini *et al.* (2006).

5.3 Effects on Beliefs

We now turn to the mechanisms behind these effects. We first report treatment effects on measures of respondents' beliefs about their ability in table 8. At baseline, we elicit respondents' beliefs about their relative performance on tests after they have taken the tests but before they receive any test results or reports. In table 8, we show effects on respondents' beliefs about their performance at endline, three to four months after treatment. We examine effects on the proportion of respondents who place themselves in the correct tercile on three cognitive skills measures when their beliefs are compared to their actual scores, a measure which ranges from 0 for those who are inaccurate on all cognitive skills to 3 for those who are accurate on all skills. We construct the same measure for non-cognitive skills.

Both treatment increase the number of domains in which respondents place themselves in the correct tercile for cognitive and noncognitive assessments (columns 1 and 4) and decrease the number of domains in which they place themselves in a higher tercile (columns 2 and 5) or lower tercile (columns 3 and 6). The public effects are uniformly and significantly larger than the private effects. Effects on over-confidence are larger than on under-confidence, but this may be driven by the higher levels of over- than underconfidence at baseline. These results show robust, large increases in the accuracy of respondents beliefs about their skills that are slightly larger for the public treatment.

These results show the information treatment is highly effective. Respondents are able to link the information they receive about their test score on the report to the test they took at Harambee. They also seem to understand the format of presenting results in terciles. They update their beliefs according to the information. And they remember the information after three months.

Table 8: Effects of Public and Private Treatments on Beliefs about Skills

	Number of cog. domains ...			Number of non-cog. domains ...		
	Accurate	Over-	Under-	Accurate	Over-	Under-
	Accurate	confident	confident	Accurate	confident	confident
Public	0.497***	-0.346***	-0.151***	0.450***	-0.296***	-0.149***
	(0.029)	(0.023)	(0.017)	(0.032)	(0.020)	(0.027)
Private	0.423***	-0.300***	-0.125***	0.316***	-0.247***	-0.064*
	(0.028)	(0.023)	(0.016)	(0.033)	(0.023)	(0.028)
# observations	6349	6349	6349	6365	6365	6365
# clusters	81	81	81	81	81	81
R-squared	0.303	0.535	0.228	0.135	0.403	0.144
Control mean	1.163	1.383	0.451	1.172	1.061	0.756
p: Public = 0	0.000	0.000	0.000	0.000	0.000	0.000
p: Private = 0	0.000	0.000	0.000	0.000	0.000	0.024
p: Public = Private	0.005	0.040	0.080	0.000	0.034	0.003
p: Public = Private = 0	0.000	0.000	0.000	0.000	0.000	0.000

All regressions include randomization block fixed effects and baseline values of the outcome; education; all assessment results; beliefs about numeracy, communications & concept formation results; age; gender; and employment status. Heteroskedasticity-robust standard errors clustered by treatment day are shown in parentheses. Lower rows of the table show p- and q-values for linear combinations of coefficients. The q-values adjust for multiple testing over all outcomes in the same family following Benjamini *et al.* (2006).

There are a number of potential reasons that the public treatment has a larger effect than the private on beliefs. Workseekers may believe the information on the public treatment is more credible because it is branded and color-printed. Information on the public reports may be more salient as workseekers may continue to see the information on the reports as they attach them to job applications. Or the public treatment may increase likelihood of employment, providing workseekers with some evidence that the information on the report is accurate and reinforcing their beliefs. We show some evidence in favor of the last explanation in table 9. For one test score, numeracy, we collect the respondents' beliefs about which tercile they fall into immediately after treatment via sms, and then after three or four months. Immediately after treatment, in the text message surveys, the public and private treatment effects on beliefs are nearly identical, and we can reject that they are the same. The difference in treatment effects on beliefs between public and private treatments only arises after three to four months, suggesting that the difference may arise from reinforcement of effects on beliefs when workseekers find work.

We can also show that neither treatment significantly changes overall self-esteem in either the text message or phone survey (columns 3 and 4). Effects on beliefs about ability do not spill over onto general self-worth, either immediately after treatment or three to four months later. Qualitative interviews with workseekers suggest that they do not view their test scores as reflecting their broader worth, as captured by the self-esteem items.

Table 9: Reconciling Differential Belief Effects

	Correct numeracy belief in ...		Above-median self-esteem	
	Text survey	Phone survey	Text survey	Phone survey
Public treatment	0.317	0.231	-0.003	0.002
	(0.015)	(0.016)	(0.014)	(0.015)
Private treatment	0.333	0.203	0.010	-0.000
	(0.017)	(0.016)	(0.015)	(0.017)
# respondents	5088	5088	0.4861	0.4861
# clusters	81	81	81	81
Control mean	0.400	0.404	0.487	0.481
p: Pub = Pvt	0.219	0.039	0.414	0.896
p: Pub effects equal	0.000		0.777	
p: Pvt effects equal	0.000		0.585	
p: Differences equal	0.014		0.480	

All regressions include randomization block fixed effects and baseline values of the outcome; education; all assessment results; beliefs about numeracy, communications & concept formation results; age; gender; and employment status. Sample is restricted to respondents with valid responses from both survey rounds. Round-specific models are estimated as a system. Heteroskedasticity-robust standard errors clustered by treatment day are shown in parentheses. Lower rows of the table show p–values for linear combinations of coefficients.

We also find smaller but significant effects on respondents' beliefs about their likelihood of finding work, as shown in table 10. We ask candidates the number of applications they expect to submit in next 30 days (the control mean is 14.6) and how long they believe they will take to find a formal job if they submit that many applications (the control mean is 2.7 months). Then candidates were asked how long they believe they will take to find a formal job if they submit half and twice as many applications (the control mean is 3.1 and 2.1 months respectively). Candidates do not view the likelihood of finding a job as a linear function of the number of applications submitted, but see a diminishing marginal return to the number of applications submitted. We calculate the expected number of months the workseekers expects it will take for them to get a job, per job application.

Effects on beliefs about the likelihood of finding work are similar in both treatment groups. On the index, both public and private treatment effects are positive and significant, and the difference between the public and private effects is not statistically significantly different (p=0.189), although the effect on the public group is somewhat larger and the effect on the private group is only marginally significant. On the components of the index, treatment effects in the public group are significantly different from zero for larger numbers of applications (n and 2n applications), but we cannot quite reject that the public and private treatment effects are the same, and effects are of the same sign. We conclude that both groups may believe they are more likely to find work because they understand their own abilities better and can more easily search for work for which they are well-suited. There is some limited evidence that the public treatment group may also believe they are more likely to find work because they can signal their skills to employers. The treatment effect is unlikely to arise from workseekers' exposure to information about the labor market during the Harambee assessment process, as all groups including the control group receive the same information.

Table 10: Effects of Public and Private Treatments on Beliefs about Likelihood of Finding Work

	Expected number of months to get a job if workseeker submits . . .		
	n apps	n/2 apps	2n apps
	. . . divided by number of applications		
Public	-0.028**	-0.036	-0.017*
	(0.010)	(0.022)	(0.008)
Private	-0.015	0.001	-0.005
	(0.012)	(0.023)	(0.007)
# observations	6099	6137	6313
# clusters	81	81	81
R-squared	0.046	0.043	0.026
Control mean	0.391	0.787	0.194
p: Public = 0	0.006	0.105	0.048
p: Private = 0	0.198	0.960	0.448
p: Public = Private	0.224	0.103	0.133
p: Public = Private = 0	0.023	0.164	0.132

All regressions include randomization block fixed effects and baseline values of the outcome; education; all assessment results; beliefs about numeracy, communications & concept formation results; age; gender; and employment status. Heteroskedasticity-robust standard errors clustered by treatment day are shown in parentheses. Lower rows of the table show p- and q-values for linear combinations of coefficients. The q-values adjust for multiple testing over all outcomes in the same family following Benjamini *et al.* (2006).

5.4 Effects on Search

In tables 11 and 12, we examine effects on workseekers' job search effort and the effectiveness of that effort. There are no significant effects of either treatment on whether workseekers searched for work, partly because, even in the control group, 70 percent of workseekers are actively searching. There are also no significant effects on hours of search in the last week, spending on search in the last week, and the number of applications submitted in the last 30 days. Effects from the public treatment are larger than from the private treatment, although effects are not significantly different. Effects on all measures of search effectiveness are fairly precisely estimated zeros, including on the number of employer responses of any form, number of offers received, the ratio of employer responses to the number of applications, and the ratio of offers to number of applicants.

For the public treatment group, this is not a strong test the effect of the treatment on search, because by endline. In table 6, we show that the public treatment effect on employment occurs very soon, one month and potentially sooner, after treatment. By endline, some workseekers may have stopped searching altogether or be searching less as a result, even if they had searched harder or more effectively immediately after treatment. To test this idea, we exploit that we randomized whether workseekers were surveyed three or four months after endline. In table 13, we show that there are positive, although not significant, effects on indices of both search effort and search effectiveness for workseekers in the public group who do the endline survey sooner after treatment, compared to those who do it later. The differences between the early and late group are statistically significantly different (p¡0.01). This suggests that workseekers in the public treatment group were searching both more intensively and, potentially, more effectively, and that this enabled them to find jobs.

The private treatment group also search somewhat more intensively sooner after treatment. But their search is no more effective than the control group in getting responses or offers from employers. By four months after treatment, they have potentially become discouraged and search at the same rate as the control group. Together, the public and private results suggest that enabling workseekers to signal their skills to firms by providing them with certified reports of their test results substantially increases their employment rate and earnings, suggesting that demand-side information frictions may distort hiring decisions and reduce employment and earnings. Providing information on its own does change workseeker beliefs and job search behavior somewhat sooner after treatment, but this has limited to no effects on their employment outcomes. This suggests supply-side information frictions result in only limited distortions in job search decisions and likelihood of employment.

6 Heterogeneity by Skills

The conceptual framework in section 2 predicts that employment effects should differ by skill level. We test this hypothesis by constructing an index of the three cognitive skills assessments, again following Anderson (2008).[21] We then regress employment outcomes on the skills index using a

[21] We cannot replicate this exact exercise for the noncognitive skills assessments, as different candidates take different assessments. But we construct two separate indices, one for the each of the sets of noncognitive assessments,

Table 11: Effects of Public and Private Treatments on Search Effort

	Any search last 7 days	Hours of search last 7 days	Apps submitted last 30 days	Spending on search last 7 days
Public	-0.02	0.252	0.624	3.001
	(0.014)	(0.368)	(0.684)	(4.058)
Private	-0.006	-0.31	0.078	1.02
	(0.014)	(0.362)	(0.648)	(4.801)
# observations	6366	6350	6336	6350
# clusters	81	81	81	81
R-squared	0.031	0.024	0.062	0.036
Control mean	0.695	11.083	11.891	122.81
p: Public = 0	0.152	0.496	0.364	0.462
p: Private = 0	0.702	0.393	0.904	0.832
p: Public = Private	0.243	0.137	0.394	0.657
p: Public = Private = 0	0.293	0.326	0.599	0.748

All regressions include randomization block fixed effects and baseline values of the outcome; education; all assessment results; beliefs about numeracy, communications & concept formation results; age; gender; and employment status. Heteroskedasticity-robust standard errors clustered by treatment day are shown in parentheses. Lower rows of the table show p- and q-values for linear combinations of coefficients. The q-values adjust for multiple testing over all outcomes in the same family following Benjamini *et al.* (2006).

Table 12: Effects of Public and Private Treatments on Search Effectiveness

	Number of employer responses	Number of offers	Number of responses per app	Number of offers per apps submitted
Public	0.049	0.007	0.001	0.001
	(0.047)	(0.017)	(0.004)	(0.002)
Private	0.041	0.017	-0.006	0.002
	(0.042)	(0.017)	(0.004)	(0.003)
# observations	6351	6350	5732	5731
# clusters	81	81	81	81
R-squared	0.021	0.011	0.010	0.009
Control mean	0.772	0.182	0.097	0.026
p: Public = 0	0.298	0.677	0.779	0.624
p: Private = 0	0.298	0.677	0.779	0.624
p: Public = Private	0.863	0.568	0.095	0.772
p: Public = Private = 0	0.488	0.610	0.200	0.759

All variables are measured over the 30 days prior to endline. All regressions include randomization block fixed effects and baseline values of the outcome; education; all assessment results; beliefs about numeracy, communications & concept formation results; age; gender; and employment status. Heteroskedasticity-robust standard errors clustered by treatment day are shown in parentheses. Lower rows of the table show p- and q-values for linear combinations of coefficients. The q-values adjust for multiple testing over all outcomes in the same family following Benjamini *et al.* (2006).

Table 13: Dynamic Effects and Survey Timing

	(1) Search Effort	(2) Search Effective	(3) Work Status	(4) Worked in past 7 days
Public × Early	0.047	0.067	0.152***	0.052**
	(0.053)	(0.044)	(0.036)	(0.018)
Public × Late	-0.042	-0.017	0.112**	0.054**
	(0.042)	(0.036)	(0.040)	(0.018)
Private × Early	0.058	0.066	0.047	0.017
	(0.051)	(0.044)	(0.040)	(0.021)
Private × Late	-0.049	0.007	0.047	0.007
	(0.040)	(0.035)	(0.039)	(0.016)
Late Phone Survey	0.037	-0.020	0.018	0.012
	(0.048)	(0.033)	(0.042)	(0.018)
# respondents	6367	6367	6367	6365
# clusters	81	81	81	81
R2	0.041	0.017	0.115	0.041
p: Average Public = 0	0.453	0.191	0.000	0.000
p: Public × Late = 0	0.381	0.127	0.000	0.006
p: Public × Early = 0	0.328	0.644	0.007	0.003
p: Average Private = 0	0.230	0.326	0.272	0.622
p: Private × Early = 0	0.258	0.137	0.250	0.417
p: Private × Late = 0	0.230	0.849	0.240	0.670

local linear regression, separately by treatment group, and show the fitted values in figures 4, 5, and 6.

These results show a strong relationship between the level of cognitive skills and all employment status measures. The employment quantity index, hours worked in the preceding 7 days, employment in the preceding 7 days and employment since treatment are weakly increasing in cognitive skills throughout most of the distribution, though the relationship is sometimes decreasing in the tails, for all treatment groups. In addition, there are public treatment effects for all these measures: all are higher for the public group than the private or control groups for most of the cognitive skills distribution. In contrast, the relationship between the private and control groups varies over the distribution and is often fairly similar. On earnings, there is a less strong relationship between cognitive skills ability and earnings. This is consistent with a labor market where entry-level wages for workseekers with limited education and little work experience may be fairly fixed regardless of workseeker ability.

However, there is little evidence of heterogeneous treatment effects: the gap between the public and other groups is similar at all levels of the skills distribution. Similarly, regressing employment outcomes on the treatment group indicators and interactions between the indicators and the cog-

standardize them to equate the variances and pool the two indices. This index has a similar relationship to the employment outcomes as the cognitive skills index does.

nitive skill index does yield robustly steeper or flatter employment-skills relationship for the public treatment group. But these linear regression results are quite sensitive to the tails of the skills distribution.

Treatment effects may differ by skill for two reasons. First, workseekers with different skills receive different information about themselves and are able to send a different signal to firms. Second, workseekers with different skills may use information and technology differently: more able workers may be better at understanding the information, may convey information about themselves better in interviews, or use the reports more effectively.

Later drafts will explore the first issue in more depth. We present preliminary results here. Identifying effects of heterogeneous treatment requires workseekers with similar skills who receive different information, which the public group can also use to signal skills to firms. We use a regression discontinuity to compare workseekers at the margin of being placed in different terciles.[22] There are a number of complications in implementing this – we need to stack running variables with different scales; some skills measures are very coarse, and we have fairly low power. And, as with all discontinuity designs, we can only recover treatment effects around the thresholds between terciles.

Preliminary results are as expected. In the public group, who are able to signal skills to firms, people who score just above the threshold to be in the top tercile (vs the middle tercile), or in the middle tercile (vs the bottom tercile) are marginally more likely to be more likely to employed (on the employment status index) (p=0.08) or to have worked in the last 7 days (p=0.117). They have significantly higher earnings. There are no similar effects for the private group, suggesting that effects are driven by employers responding more favourably to people with more positive signals than through changes in beliefs.

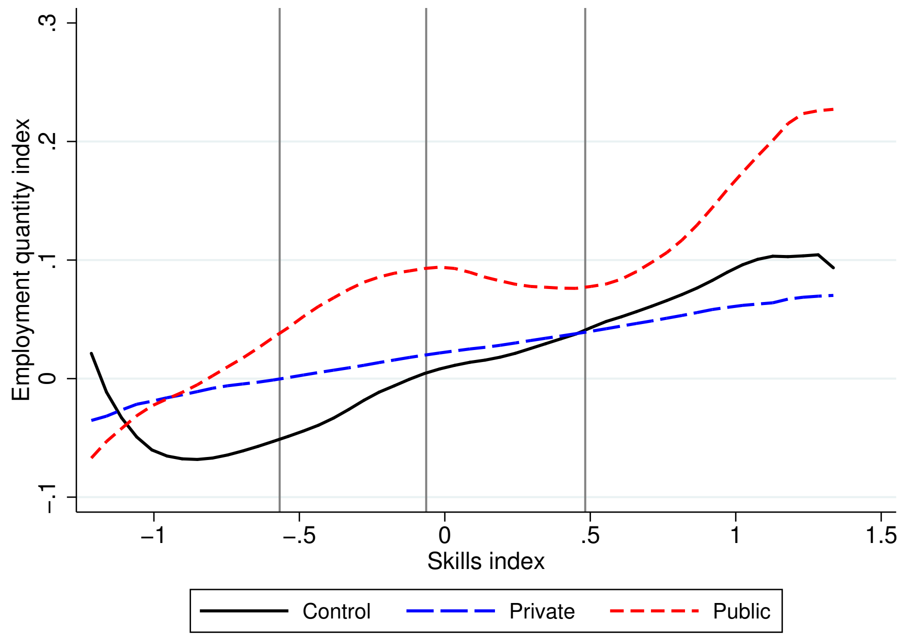
Second, we examine treatment effect heterogeneity by skill. Identifying heterogeneous treatment effects requires workseekers with different skills who receive the same information. We thus compare workseekers in the top and bottom half of each tercile, with the caveats that this only tests for local, not global, heterogeneity and has fairly low power. We focus only on numeracy scores, as this is the variable we analyse above. We find little evidence of either public or private treatment effect heterogeneity: there is no significant difference in treatment effects between workseekers with higher or lower skills, within a tercile. Broadly, the overall results of public treatment effect heterogeneity by skill we see are more consistent with *treatment heterogeneity* than *treatment effect heterogeneity*

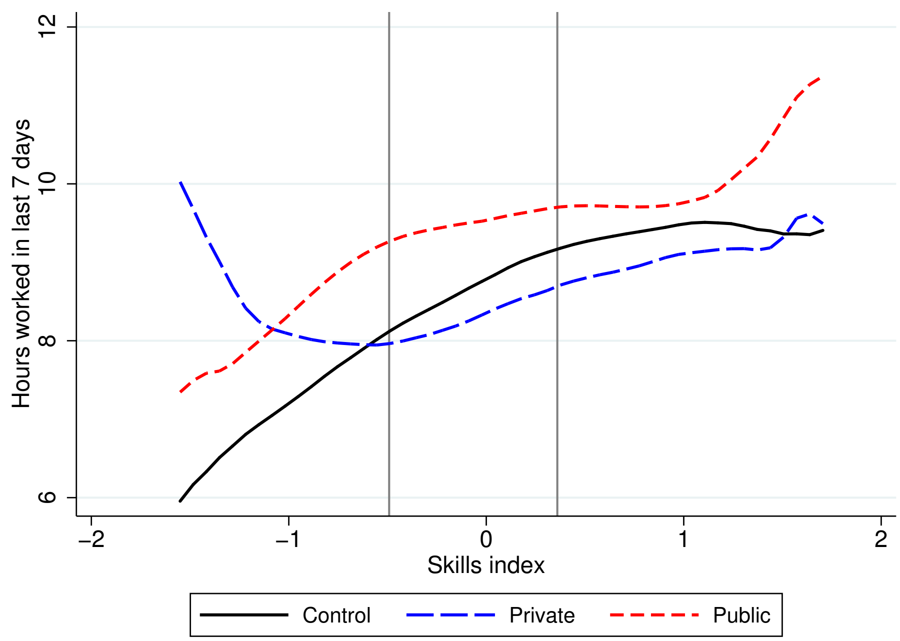
7 Conclusion

Firms typically make hiring decisions with limited information about workseekers' skills and productivity. Similarly, workseekers make job search and employment decisions with limited information

[22] We only use the numeracy assessment here because it is the only score with sufficiently continuous variation.

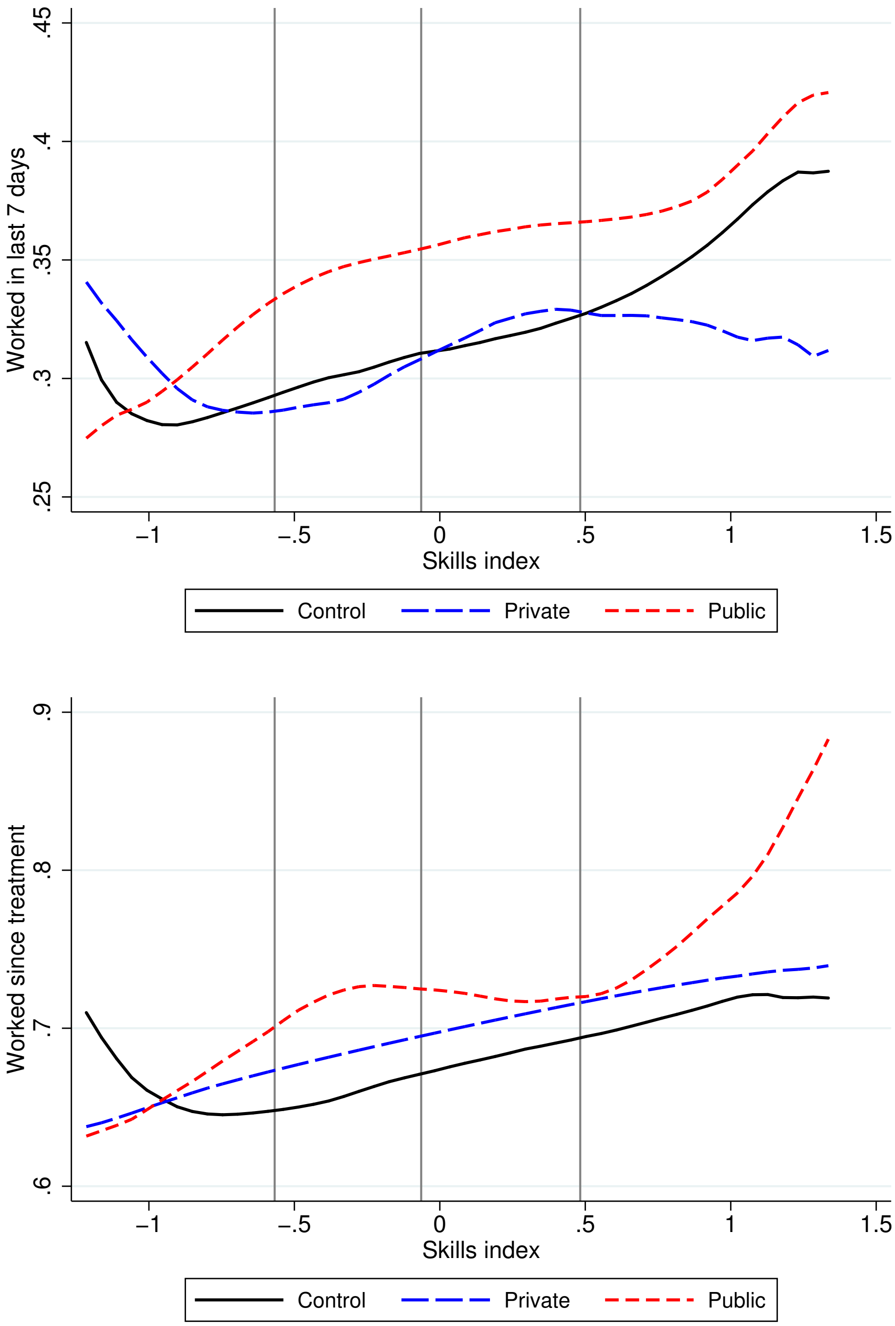
Figure 4: Effects of Public and Private Treatments on Employment by Cognitive Skills





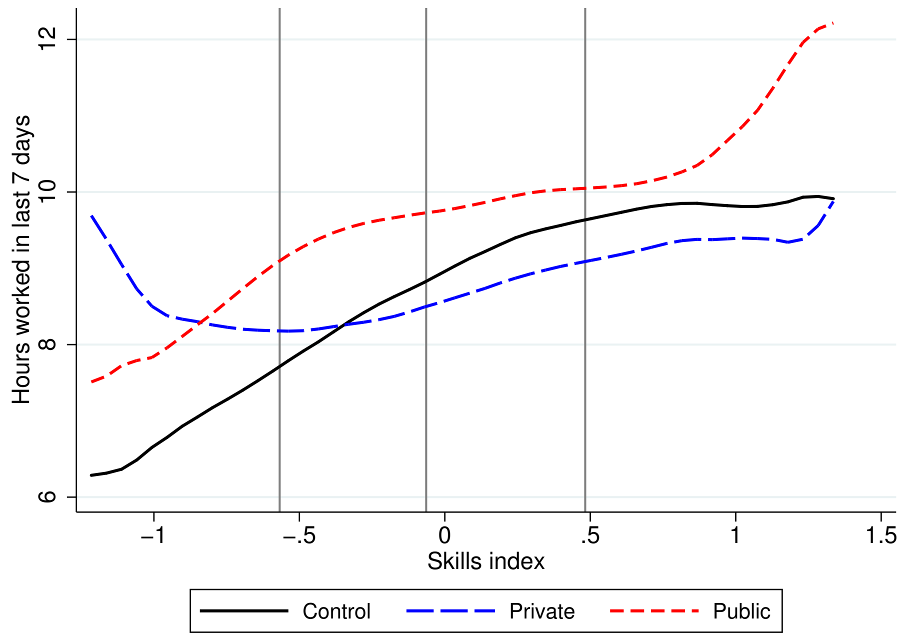
Notes: This figure shows the fitted values from local linear regressions of selected employment outcomes on the cognitive skills index. Regressions are truncated at the 5th and 95th percentiles of the skills index. Vertical lines show the 25th, 50th, and 75th percentiles of the skills index. Employment index construction is described in the text and follows Anderson (2008).

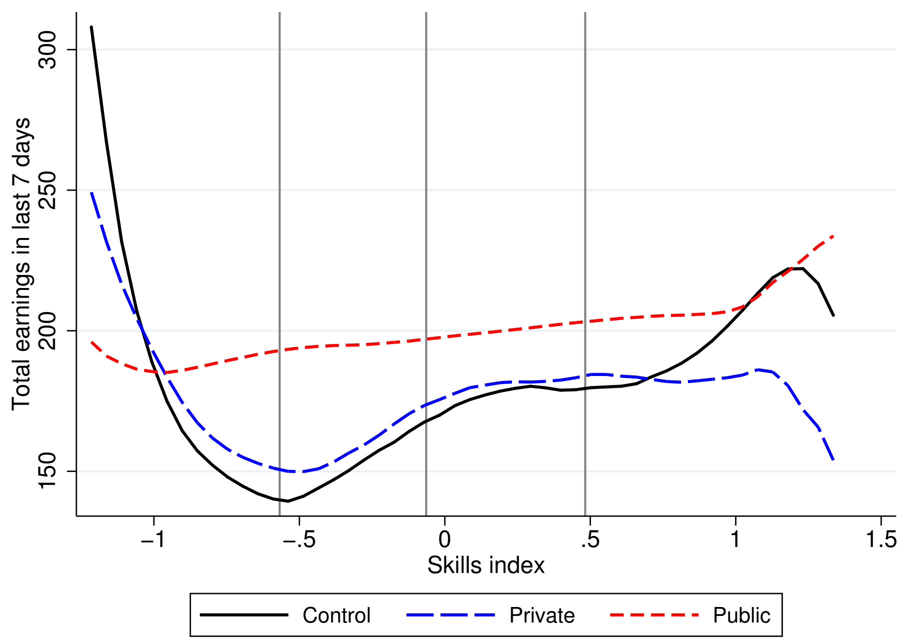
Figure 5: Effects of Public and Private Treatments on Employment by Cognitive Skills



Notes: This figure shows the fitted values from local linear regressions of selected employment outcomes on the cognitive skills index. Regressions are truncated at the 5th and 95th percentiles of the skills index. Vertical lines show the 25th, 50th, and 75th percentiles of the skills index.

Figure 6: Effects of Public and Private Treatments on Employment by Cognitive Skills





Notes: This figure shows the fitted values from local linear regressions of selected employment outcomes on the cognitive skills index. Regressions are truncated at the 5th and 95th percentiles of the skills index. Vertical lines show the 25th, 50th, and 75th percentiles of the skills index. Hours and earnings are both reported for the preceding seven days, asked separately about main and othe jobs and then combined, winsorized at their 99th percentiles, and set to zero for respondents who did not work in the preceding seven days.

Table 14: Effect of Treatment Heterogeneity by Skill

	Index of employment ...		In past 7 days ...	
	status	quality	worked	earnings
Public	0.147	-0.018	0.058	103.2
	(0.085)	(0.073)	(0.037)	(37.7)
p: Pub = 0	0.084	0.805	0.117	0.006
# observations	4496	4496	4322	4476
# clusters	81	81	81	81
Private	-0.016	-0.129	-0.042	-92.0
	(0.061)	(0.127)	(0.029)	(67.9)
p: Pvt = 0	0.793	0.310	0.148	0.176
# observations	4228	4228	4040	4216
# clusters	81	81	81	81
Control	-0.075	0.021	-0.004	-20.3
	(0.088)	(0.063)	(0.045)	(39.0)
p: Control = 0	0.394	0.739	0.929	0.603
# observations	4550	4550	4368	4542
# clusters	81	81	81	81

Outcomes in columns 1 - 2 are expressed in standard deviations of family-level indices (Anderson, 2008). Results are obtained by stacking the top-middle and middle-low tercile cutoffs for the numeracy assessment, recentering each cutoff at zero, and estimating discontinuities separately by treatment group. All results are generated using local linear regressions with triangular kernels and the robust bandwidth estimator from Calonico *et al.* (2014). Heteroskedasticity-robust standard errors clustered by treatment day are shown in parentheses. Lower rows of the table show p-values for linear combinations of coefficients.

Table 15: Treatment Effect Heterogeneity by Skill

	Index of employment ...		In past 7 days ...	
	status	quality	worked	earnings
Public	0.105	0.012	0.041	53.6
	(0.047)	(0.046)	(0.021)	(37.8)
Public * High	-0.002	0.002	-0.003	4.3
	(0.059)	(0.065)	(0.027)	(42.5)
Private	-0.024	-0.020	-0.024	13.8
	(0.050)	(0.042)	(0.022)	(35.2)
Private * High	0.067	0.056	0.040	60.9
	(0.056)	(0.067)	(0.026)	(45.0)
High	0.001	-0.005	0.007	-22.4
	(0.045)	(0.039)	(0.021)	(28.4)
# respondents	6367	6367	6365	6347
# clusters	81	81	81	81
R-squared	0.105	0.010	0.038	0.013
Control mean	0.000	-0.000	0.309	152.9
Control mean | employed				504.9
p: Pub = 0	0.025	0.775	0.062	0.127
p: Pub × High = 0	0.973	0.975	0.912	0.920
p: Pvt = 0	0.631	0.634	0.275	0.695
p: Pvt × High = 0	0.232	0.403	0.124	0.176

Outcomes in columns 1 - 2 are expressed in standard deviations of family-level indices (Anderson, 2008). All regressions include randomization block fixed effects and baseline values of the outcome; education; all assessment results; beliefs about numeracy, communications & concept formation results; age; gender; and employment status. Heteroskedasticity-robust standard errors clustered by treatment day are shown in parentheses. Lower rows of the table show p-values for linear combinations of coefficients.

about their own skills and productivity. These information frictions on both the demand and supply sides of the labor market can distort job search and hiring decisions, lowering total employment, total earnings, and firm productivity. We study information frictions using a randomized controlled trial and longitudinal survey in urban South Africa. South Africa is marked by extremely high unemployment, particularly amongst youths, and a weak education system. This suggests that information frictions may be particularly important in this market.

We directly assess workseekers' cognitive and noncognitive skills and experimentally vary whether workseekers receive information about their own skills and whether they can signal this information to firms. We find that giving information to workseekers has large effects on their beliefs about their skills, some early effects on their job search behavior (which do not persist) and limited effects on their employment outcomes. Helping workseekers signal their skills to firms substantially increases their employment rate and earnings, driven partly by increased search effort but also by workseekers being more able to signal skills to firms. We conclude that information frictions exist on both sides of the labor market but only the demand-side frictions appear important for labor market outcomes.

Our results contribute to a growing literature on information frictions in the labor market. We add to a small body of research that directly manipulates information about workseekers' skills. Relative to this literature, we separately vary workseekers' information about their own skills and their ability to signal this to firms, which allows us to separate supply- and demand-side information frictions. We provide more detailed measures of job search and employment quality than most research in this area, which allows us to show that changes in search behavior do not explain information-induced changes in employment outcomes. Our findings show that information frictions can substantially distort individual workseekers' employment outcomes. In ongoing work, we build on this to test if information frictions reduce total labor demand or simply help some workseekers at others' expense. This ongoing programme of work will more thoroughly characterize the employment and welfare effects of information frictions on both sides of the labor market.

References

ABEBE, G., S. CARIA, M. FAFCHAMPS, P. FALCO, S. FRANKLIN, AND S. QUINN (2016): "The Curse of Anonymity or Tyranny of Distance? The Impacts of Job-Search Support in Urban Ethiopia," Working Paper 22409, National Bureau of Economic Research.

ABEL, M., R. BURGER, AND P. PIRAINO (2016): "The Value of Reference Letters -Experimental Evidence from South Africa," Working Paper, Harvard University.

ALTONJI, J. AND C. PIERRET (2001): "Employer Learning and Statistical Discrimination," *Quarterly Journal of Economics*, 116, 313–335.

ANDERSON, M. (2008): "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association*, 103, 1481–1495.

ARCIDIAONO, P., P. BAYER, AND A. HIZMO (2010): "Beyond Signaling and Human Capital: Education and the Revelation of Ability," *American Economic Journal: Applied Economics*, 2, 76–104.

AUTOR, D. AND D. SCARBOROUGH (2008): "Does Job Testing Harm Minority Workers? Evidence from Retail Establishments," *Quarterly Journal of Economics*, 123, 219–277.

BASSI, V. AND A. NANSAMBA (2017): "Information Frictions in the Labor Market: Evidence from a Field Experiment in Uganda," Working Paper, University College London.

BENJAMINI, Y., A. KRIEGER, AND D. YEKUTIELI (2006): "Adaptive Linear Step-Up Procedures That Control the False Discovery Rate," *Biometrika*, 93, 491–507.

BLATTMAN, C. AND L. RALSON (2015): "Employment in Poor and Fragile States: Evidence from Labor Market and Entrepreneurship Programs," Manuscript, Columbia University.

CARD, D., J. KLUVE, AND A. WEBER (2015): "What Works? A Meta-Analysis of Recent Active Labor Market Program Evaluations," Working Paper 21431, National Bureau of Economic Research.

DUCKWORTH, A., C. PETERSON, M. MATTHEWS, AND D. KELLY (2007): "Grit: Perseverance and Passion for Long-term Goals," *Journal of Personality and Social Psychology*, 92, 1087–1101.

FARBER, H. AND R. GIBBONS (1996): "Learning and Wage Dynamics," *Quarterly Journal of Economics*, 111, 1007–1047.

GNEEZY, U., A. RUSTICHINI, AND A. VOSTROKNUTOV (2010): "Experience and Insight in The Race Game," *Journal of Economic Behavior and Organization*, 75, 144–155.

HARDY, M. AND J. MCCASLAND (2017): "Are Small Firms Labor Constrained? Experimental Evidence from Ghana," Working Paper, New York University - Abu Dhabi.

HEPNER, P. AND C. PETERSEN (1982): "The Development and Implications of a Personal Problem-Solving Inventory," *Journal of Counseling Psychology*, 29, 66–75.

HOFFMAN, M., L. KAHN, AND D. LI (2015): "Discretion in Hiring," Working Paper 21709, National Bureau of Economic Research.

Horton, J. (2017): "The Effects of Algorithmic Labor Market Recommendations: Evidence from a Field Experiment," *Journal of Labor Economics*, 35, 345–285.

Hsieh, C.-T. and B. Olken (2014): "The Missing 'Missing Middle'," *Journal of Economic Perspectives*, 28, 89–108.

Kahn, L. and F. Lange (2014): "Employer Learning, Productivity, and the Earnings Distribution: Evidence from Performance Measures," *Review of Economic Studies*, 84, 1575–1613.

Levinsohn, J., N. Rankin, G. Roberts, and V. Schoer (2013): "Wage Subsidies and Youth Employment in South Africa: Evidence from a Randomized Control Trial," Working Paper, University of Stellenbosch.

MacLeod, W. B., E. Riehl, J. Saavedra, and M. Urquiola (2017): "The Big Sort: College Reputation and Labor Market Outcomes," *American Economic Journal: Applied Economics*, 9, 223–261.

McKenzie, D. (2017): "How Effective Are Active Labor Market Policies in Developing Countries? A Critical Review of Recent Evidence," Working Paper 8011, World Bank Policy Research.

Pallais, A. (2014): "Inefficient Hiring in Entry-Level Labor Markets," *American Economic Review*, 104, 3565–3599.

Pritchett, L. (2013): *The Rebirth of Education: Schooling Ain't Learning*, Washington, DC: Center for Global Development.

Spence, M. (1973): "Job Market Signalling," *Quarterly Journal of Economics*, 87, 355–374.

Stroop, J. R. (1935): "Studies of Interference in Serial Verbal Reactions," *Journal of Experimental Psychology*, 18, 643–662.

Tibshirani, R. (1996): "Regression Shrinkage and Selection via the Lasso," *Journal of the Royal Statistical Society Series B (Methodological)*, 58, 267–288.

World Bank (2018): *World Development Report 2018: Learning to Realize Education's Promise*, World Bank.