INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

IFPRI

IFPRI Discussion Paper 01332

March 2014

Employment Risk and Job-Seeker Performance

Susan Godlonton

Markets, Trade and Institutions Division

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

The International Food Policy Research Institute (IFPRI), established in 1975, provides evidence-based policy solutions to sustainably end hunger and malnutrition and reduce poverty. The Institute conducts research, communicates results, optimizes partnerships, and builds capacity to ensure sustainable food production, promote healthy food systems, improve markets and trade, transform agriculture, build resilience, and strengthen institutions and governance. Gender is considered in all of the Institute's work. IFPRI collaborates with partners around the world, including development implementers, public institutions, the private sector, and farmers' organizations, to ensure that local, national, regional, and global food policies are based on evidence.

AUTHOR

Susan Godlonton (s.godlonton@cgiar.org) is an associate research fellow in the Markets, Trade and Institutions Division of the International Food Policy Research Institute, Washington, DC.

Notices

IFPRI Discussion Papers contain preliminary material and research results and are circulated in order to stimulate discussion and critical comment. They have not been subject to a formal external review via IFPRI's Publications Review Committee. Any opinions expressed are those of the author(s) and do not necessarily reflect the policies or opinions of IFPRI.

Copyright 2014 International Food Policy Research Institute. All rights reserved. Sections of this material may be reproduced for personal and not-for-profit use without the express written permission of but with acknowledgment to IFPRI. To reproduce the material contained herein for profit or commercial use requires express written permission. To obtain permission, contact the Communications Division at ifpri-copyright@cgiar.org.

Contents

Tables

Figures

ABSTRACT

High levels of job uncertainty in developing countries may have significant implications for job performance. This paper examines the relationship between employment risk and job-seeker performance. To induce exogenous variation in employment risk, the outside options for job seekers undergoing a real recruitment process were randomized by assigning them a 0, 1, 5, 50, 75, or 100 percent chance of real alternative employment of the same duration and wage as the jobs for which they were applying. The findings show that job-seeker performance is highest and effort is lowest among those assigned the lowest employment risk (a guaranteed alternative job), and performance is lowest and effort highest among those facing the highest employment risk (those without any job guarantee). Moreover, a nonlinear relationship exists between employment risk and performance.

The findings are consistent with a framework that ties together insights from economics and psychology—that is, performance is an increasing function of effort and an inverse U-shaped function of stress. The results are not driven by gift exchange, stereotype threat, or the nutritional efficiency wage hypothesis.

These performance improvements have significant welfare implications. In this study, job seekers assigned a high probability of receiving an outside option were twice as likely to be hired in the standard job recruitment process compared with those assigned a low probability of receiving an outside option. More broadly, these results suggest that stress-induced performance reductions are a potential mechanism through which exposure to high employment risk can perpetuate poverty and unemployment.

Keywords: employment risk, performance, recruitment, randomized controlled trial, stress

JEL Codes: O12, O15, J00, M51

ACKNOWLEDGMENTS

This project was supported with research grants from the Rackham Graduate School and the African Studies Center at the University of Michigan. I gratefully acknowledge use of the services and facilities of the Population Studies Center at the University of Michigan, funded by National Institute of Child Health and Human Development Center Grant R24 HD041028. I acknowledge Ernest Mlenga and his recruitment team for their willingness to provide their recruitment data as well as enabling the integration of this research into their recruitment process. I also acknowledge Kelvin Balakasi for his field and research assistance in Malawi and Soo Ji Ha for her research assistance in Michigan. I thank Sarah Burgard, David Lam, Jeff Smith, Rebecca Thornton, Jessica Goldberg, Emily Beam, and Caroline Theoharides for valuable comments. This paper also benefited from feedback from seminar participants at the Center for Global Development, Mathematica, University of Cape Town, University of Connecticut, University of Michigan, University of Notre Dame, University of Rochester, IFPRI, Williams College, and the World Bank.

1. INTRODUCTION

Typically, behavioral responses are studied either in reaction to the realization of a risky process or in response to underlying risk, which usually take the form of risk coping or risk mitigation, respectively. The same framework characterizes most studies of employment risk: The emphasis is either on understanding either the consequences of a job gain or loss or on coping strategies for mitigating income uncertainty.[1] However, little is known about how risk might affect performance in the job-seeking process or the ultimate chance of securing employment.[2] Given high and rising employment uncertainty (ILO 2012), it is important to understand the process through which risk may affect employment, as well as the extent to which risk has heterogeneous effects across job seekers. This paper explicitly examines the relationship between employment risk and job-seeker performance and employment.

Typically, the causal effect of employment risk on performance cannot be estimated because of challenges in measurement and identification. First, measuring employment risk is difficult. For example, research examining the relationship between risk and savings uses such proxies as variability in household income, variability in expenditures, or, in more recent work, the probability of a job loss (Caroll 1994; Dynan 1993; Lusardi 1998). An appropriate proxy when measuring employment risk might be the probability of a job gain rather than a job loss. Still, none of these proxies provides a direct measure of employment risk. An alternate approach is to measure decision making in response to experimentally induced risk in a laboratory setting. Such experiments provide useful insights about potential mechanisms; however, it is often unclear whether evidence from lab-based experiments will translate into real-world behavior. Second, even if one could directly measure employment risk, it is usually endogenous to the outcome of interest. For example, in the case of job-seeker performance, individuals of higher ability are likely to face lower employment risk, yet also perform better on average, making it hard to establish causality. Finally, although effort and performance are key mechanisms through which risk affects employment, these concepts are difficult to measure due to self-reporting biases and lack of good-quality data.

This paper overcomes these challenges by explicitly varying employment risk using a field experiment to examine the impact of employment risk on job-seeker performance. Job seekers' outside options were randomized during a real recruitment process, in collaboration with a real recruiter offering short-term jobs. Each of 268 job seekers was randomly assigned a probabilistic chance (0, 1, 5, 50, 75, or 100 percent) of an alternative job, thus reducing the downside risk of performing poorly during the recruitment process. For those with a guaranteed outside option, employment risk was zero. Both objective and subjective performance assessments from administrative data were used to examine the relationship between employment risk and job-seeker performance, whereas indicators from both administrative and self-reported data sources were used to measure effort.

The findings show that improving a job seeker's outside option leads to improved performance and a decline in effort. Job seekers assigned a guaranteed outside option performed approximately 0.45 standard deviations better on recruiter-administered tests of knowledge taught in training than those who received no outside option. Moreover, the relationship between risk and performance is highly nonlinear.

[1] The World Development Jobs Report (2013) provided an extensive overview of the individual and social consequences of employment. One key strand of literature focuses on the impact of gaining or losing work. Often this empirical work uses an exogenous shock that results in job loss, such as plant closures and retrenchments, in order to examine both short-term and long-term effects on future employment and earnings (Stevens 1997; Chan and Stevens 2001; Ruhm 1991, 1994; Topel 1990; Gregg and Tominey 2005; Couch 2001). A second strand of relevant literature examines risk-coping mechanisms and their impact on the labor market. This literature has examined the roles of unions (Magruder 2012), UI (Gruber 1997; Green and Riddell 1993), and informal networks (Burns, Godlonton, and Keswell 2010; Beaman and Magruder 2012), as well as how individuals use these support structures to mitigate risk of unemployment.

[2] As discussed in-depth in Fafchamps (2010), *shocks* and *risks* are often used interchangeably despite being distinct. Fafchamps highlighted the lack of research on the impact of any type of risk in the empirical development literature, which has instead focused on the effects of shocks, ignoring the anticipatory nature of the shocks. This is in contrast to older theoretical work that explicitly addressed this and showed that risk aversion should lead to underinvestment and underproduction (Sandmo 1971).

These findings are confirmed using the quality of active participation in job training as a measure of performance. The results indicate higher quality average engagement in training by those assigned high outside options compared with those assigned no outside option. For effort indicators, the reverse is shown—that is, job seekers assigned the guaranteed outside option put forth the lowest effort, whereas those assigned no outside option put forth the most effort.

In terms of punctuality, job seekers assigned a guaranteed outside option were 9.3 percentage points more likely to arrive late during the three-day training conducted during recruitment, as compared with those assigned no outside option. However, the difference is not statistically significant. I did find large, robust, and statistically significant differences in self-reported effort: Individuals assigned a guaranteed outside option spent 25 fewer minutes per day studying training materials compared with job trainees assigned no outside option.

In sum, performance is highest and effort is lowest among those assigned the lowest employment risk, and performance is lowest and effort highest among those facing the highest employment risk. These results are robust to a number of different robustness checks, such as weighted regressions and Lee bounds, thus addressing concerns about differential survey nonresponse.

Although this is the first study (to my knowledge) to examine this question in labor markets, the results are consistent with Ariely et al. (2009),[3] who conducted laboratory experiments among 76 participants in rural India, offering either a high, medium, or low incentive for meeting a performance target on six different games that tested concentration, creativity, or motor skills. These performance incentives were, in some sense, the inverse of the variation in my experiment: whereas I decreased risk, high-powered incentives increase it. Ariely et al. (2009) consistently found that performance in the group assigned the high incentive (400 Indian rupees, or equivalent to a month's salary) was lowest.

My findings are consistent with those of Ariely et al. (2009) in that performance is highest in the group assigned a guaranteed outside option and lowest in the group assigned no outside option. My contributions go beyond affirming this finding, however, as the relationship between risk and performance in this context is highly nonlinear. I extended the experiment from the risk associated with wage incentives to study employment risk—a distinct, though clearly related, construct with potentially larger welfare consequences. I also extended the literature from the lab to the field. The variation in risk in laboratory studies is artificial and over windfall income; in my setting, the variation is over risk in securing real, meaningful employment equivalent to that for which subjects have chosen to apply through a competitive process. No known evidence in a real-world setting has illustrated the link among risk, performance, and effort.

This reduced-form relationship between risk and performance is itself inherently interesting; thus, this study also explores the potential mechanism driving this result. One such mechanism is *choking under pressure*, which draws on economic and psychological insights. Improving a job seeker's outside options reduces the incentive to exert effort during recruitment. Therefore, as outside options increase, effort in the recruitment process should decline and, thus, so too should performance. At the same time, however, improving a job seeker's outside options reduces the stress experienced during the job-seeking process—a premise that is supported by psychology and public health literature, which find that uncertain employment prospects are stressful (Feather 1990; de Witte 1999, 2005; Burgard, Brand, and House 2009). This reduction in stress likely has performance implications, as Yerkes and Dodson (1908) showed that performance has an inverse, U-shaped correlation with arousal (stress). Therefore, as stress decreases due to improved outside options, performance could increase or decrease. The resulting predictions suggest that as risk declines, so too should effort; however, it is ambiguous whether performance would increase or decrease. The finding in this paper that when risk is eliminated effort is reduced and at the same time individuals perform better is consistent with this framework.

[3] Psychologists have extensively studied conditions under which increased pressure to perform has resulted in *choking*. Seminal work is presented in Baumeister (1984) and Baumeister and Showers (1986). More recently, Beilock (2010) provided a comprehensive review of this literature, covering performance in sports, academic environments, and professional settings.

The impact of stress on performance is not well studied within economics. The research that does exist focuses on how stress affects performance in professional activities, sports performance, and academic settings.[4,5,6] In fact, Kamenica (2012) stated: “Overall, to date there is no compelling empirical evidence that *choking* plays an important role in any real-world labor market.” My results attempt to fill this gap.

This study also considers other behavioral channels that are consistent with the key result that individuals facing a lower incentive to perform (better outside options) exhibit higher performance. One possibility is gift exchange, which requires the performance results to be driven by increased effort. However, the results show that individuals assigned a high probability of an outside option exert less effort in studying for tests during recruitment—thus, ruling out gift exchange as the mechanism. Second, the nutritional efficiency wage hypothesis might be an alternative mechanism. However, no observed differences exist in food expenditures by treatment group during training; so it is unlikely that this is the driving mechanism. The third consideration is stereotype threat. The findings show that job trainees' perceptions about their own likelihood of being hired by the recruiter do not significantly differ across treatment groups, suggesting that stereotype threat is an unlikely mechanism. Although the results are consistent with a stress response, some other psychological consideration that operates in a similar way to stress should not be ruled out.

The finding that performance is highest among individuals with guaranteed outside options has important policy implications. In this study, differences in employment rates by treatment status show that individuals assigned a 75 or 100 percent chance of alternative jobs were twice as likely to be employed by the recruiter compared with those in the other treatment groups. Randomization was stratified on quantile ability, but the small sample size limits the ability to rigorously examine the heterogeneity of the impacts. However, suggestive evidence indicates that the reduction in employment risk has the greatest impact on individuals in the middle of the ability distribution.

More broadly, these results suggest that individuals with greater income support through employment guarantees, cash transfer programs, family support, or employment income are likely to perform better, which may have positive feedback effects. Poor initial employment probabilities can induce stress-related performance reductions, resulting in poverty persistence across individuals, communities, or countries. Lastly, the results yield insights into what types of people are more likely to be hired with different recruitment strategies. For example, individuals exposed to higher employment risk have a greater chance of employment in hiring processes that place greater emphasis on effort than on performance.

The remainder of the paper proceeds as follows: Section 2 provides contextual information about labor markets and recruitment in Malawi and presents the experimental design. Section 3 outlines the different data sources used. Section 4 presents the estimation strategy, and Section 5 presents the underlying *choking under pressure* framework and prediction. Section 6 presents and discusses the results, and Section 7 concludes.

---

[4] The public health and industrial psychology literatures have shown that stress is correlated with performance among nurses, medical doctors, police officers, and teachers (Jamal 1984; Motowidlo, Packard, and Manning 1986; Sullivan and Bhagat 1992; Band and Manuele 1987).

[5] The literature on sports performance presents relatively mixed results. Primarily, this literature has looked at the probability of scoring penalty kicks in professional soccer. Dohmen (2008) found that when the importance of scoring is greatest, individuals tend to score. Apesteguia and Palacios-Huerta (2010) found that players who shoot second in a penalty shoot-out lose the game 60.5 percent of the time. They argued that this result is driven by increased pressure to perform, and identification is achieved because the order of the shoot-out is determined randomly from a coin flip. However, Kocher, Lenz and Sutter (2012) failed to replicate these findings using an extended dataset. Paserman (2010) examined performance in tennis and set up a structural model, finding that individuals would be substantially more likely to win if they could score when it mattered most.

[6] The literature examining high-stakes academic testing also finds mixed evidence. Örs, Palomino, and Peyrache (2013) found that women perform worse than men on a high-stakes entrance exam for an elite university in France, despite higher performance on other low-stakes exams. In the education literature more broadly, testing anxiety has been widely observed and studied. Evidence shows that test anxiety can both increase and decrease performance. (Sarason and Mandler [1952] and Tryon [1980] provided extensive reviews.)

2. BACKGROUND AND EXPERIMENTAL DESIGN

To examine the relationship between employment risk and job-trainee performance, this study varies individuals' outside options during a real recruitment process. In the absence of this intervention, the distribution of job seekers' outside options is correlated with their own ability, prior work experience, and social networks. Job trainees are offered a randomly assigned probability of an alternative job with the same wage and duration as the job for which they have applied. During this study, I worked in collaboration with a real recruiter and embedded the experimental component into an already-existing recruitment process.

Setting

Urban labor markets in developing countries are characterized by high unemployment and underemployment, as well as high job instability (WDR 2013). High rates of in-migration to urban areas in developing countries suggest that these problems are likely to increase and that rural labor markets are worse. Malawi, the fourth-fastest urbanizing country in Africa (HDR 2009), is no exception. Data from the nationally representative Malawi 2010/2011 Integrated Household Survey (IHS3) show that among urban Malawian men aged 18 to 49, only 29.6 percent reported engaging in economic activity in the preceding seven days. Incidence of job turnover and the pre valence of short-term contracting are not well measured. However, sectors that are characterized by high turnover, fixed-term contracts, and seasonality are the most common among urban residents.[7]

Due to the recruiter's eligibility restrictions, the sample in this paper is restricted to men aged 18 and older who have completed secondary schooling. Approximately 39 percent of urban men aged 18 to 49 have completed secondary schooling in Malawi. However, they too face high rates of unemployment, with only 52.5 percent of them having worked in the preceding year.[8] Due to their relatively higher social status, these men also bear considerable financial responsibility, not only for their immediate families but also often for extended family members. On average, these men report sending 10 percent of their wage income to other households (Malawi, NSO 2011).

Recruitment Process and Timeline

The sample of respondents was drawn from a recruitment process hiring interviewers for a health survey.[9] Contract work on survey projects for government or international organizations, research projects, or nongovernmental organizations is quite common in Lilongwe, Malawi's capital city. Data collected by Chinkhumba, Godlonton, and Thornton (2012), which sampled approximately 1,200 men aged 18–40 in Lilongwe, found that one in ten individuals had ever worked as an interviewer; of those who had completed secondary schooling, the number was one in four.[10] This dataset also provides some descriptive data on hiring practices. A total of 23 percent reported having taken a test for their most recently held job.

[7] For example, 7.9 percent reported working in construction and 46.8 percent in community, social, and personnel services (Malawi, NSO 2011). In addition, it is worth noting that the community, social, and personnel services sector also includes teachers, whom I excluded when calculating the fraction working in this sector because teaching, though low paying, is a stable profession in this context.

[8] When examining responses regarding activities in the past seven days among men who completed secondary school and resided in urban areas in the IHS3 2010/2011 data, 1 percent reported working in household agricultural activities; 6.2 percent had run or were helping to run small household businesses; 1.95 percent were engaged in *ganyu*/day labor; and 21.7 percent had been employed for a wage, salary, or commission.

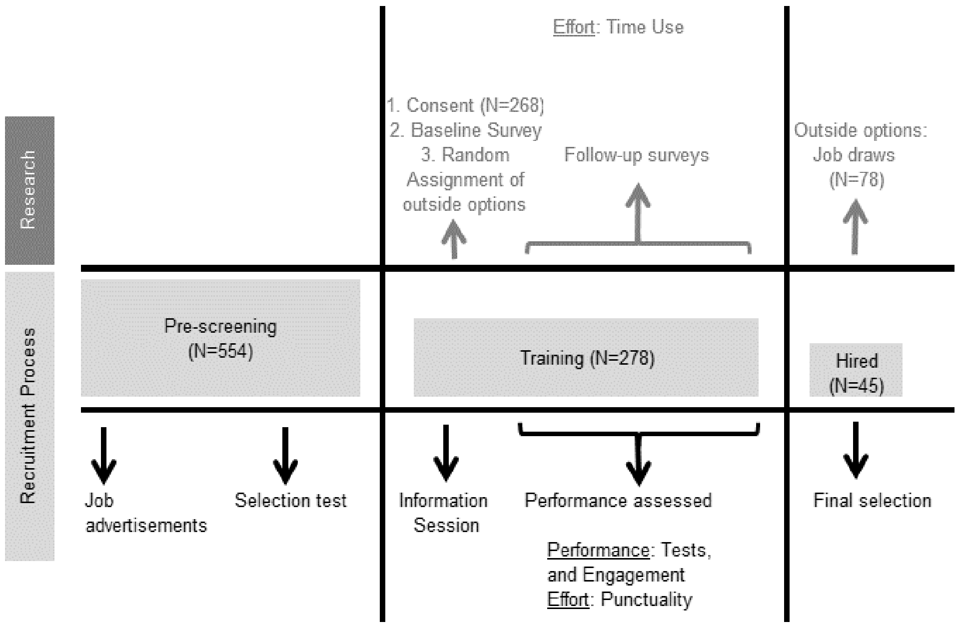
[9] The recruiter conducted independent consulting within Malawi and had, for several years, implemented various randomized controlled trials and other data-collection efforts in Malawi for universities and other international NGOs.

[10] These numbers are high and warrant further explanation. First, the Malawi census took place in 2010. Many individuals are likely to have worked for the census, as the National Statistics Office hired extensively. Second, 65 percent of individuals only reported working once as an interviewer. Third, the concept of *interviewer* is likely broadly interpreted, including work individuals may have conducted as market research or other nonresearch that involved interviewing others.

Approximately half (51.5 percent) reported being interviewed for their most recently held job, and one-third reported attending job-specific training for their most recently held job.[11]

The jobs offered by the recruiter were relatively high paying, offering approximately three times the average wage for men who had completed secondary school.[12] However, the wages offered were comparable for this type of work.[13] The recruitment process timeline is presented in Figure 2.1.

Figure 2.1 Timeline of recruitment and research activities



Source: Author.

Note: Items in the top third of the figure indicate research activities conducted for the purposes of this study. Items in black indicate standard recruitment activities performed by the recruiter.

Prescreening

To advertise positions, advertisements were placed in multiple public places. The placement was determined and conducted by the recruiter and followed their standard protocol. The advertisements included eligibility requirements and the application procedure. To apply, each individual was required to take a prescreening assessment test and submit a copy of his résumé.[14] A total of 554 applicants took this written assessment, which tested numeracy and literacy ability. The recruiter used the assessment to select the top 278 applicants to advance to the job-training phase of the recruitment process. The distributions of these scores are presented in Figure 2.2. Given this selection criterion, the sample of interest is a nonrepresentative sample of applicants. However, it is representative of the individuals who were selected for training by the recruiter and therefore captures the population of interest relevant to the research question.

[11] These numbers come from the authors' own tabulations from unpublished data collected by Chinkhumba, Godlonton, and Thornton (2012).

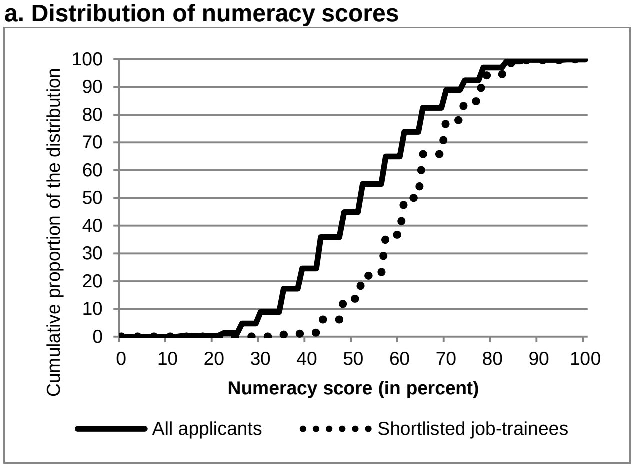
[12] The average wage among 18- to 49-year-old urban men who had completed secondary schooling is approximately US$4.75 per day, and the median is US$2.02 per day (IHS 2010/2011).

[13] Wages at institutions hiring interviewers regularly (such as Innovations for Poverty Action and the National Statistics Office) ranged from US$15 to US$32 per day for urban interviewers. Wages offered in this case are on the low end for this type of work, at US$15 per day.

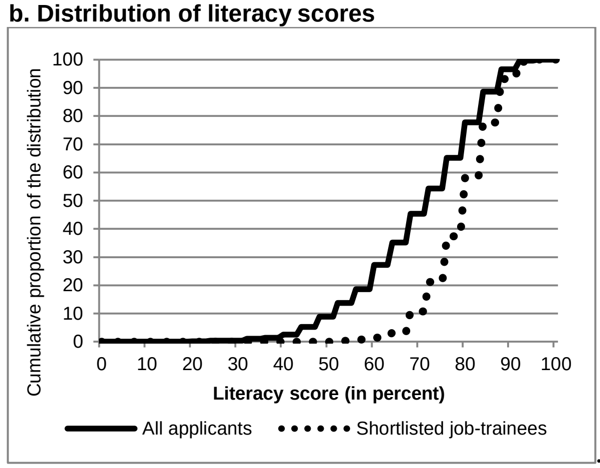
[14] Individuals were encouraged to bring their résumés. Most (95 percent) did bring a résumé. Those who did not were not prevented from taking the prescreening assessment test.

Figure 2.2 Distribution of numeracy, literacy, and ability scores

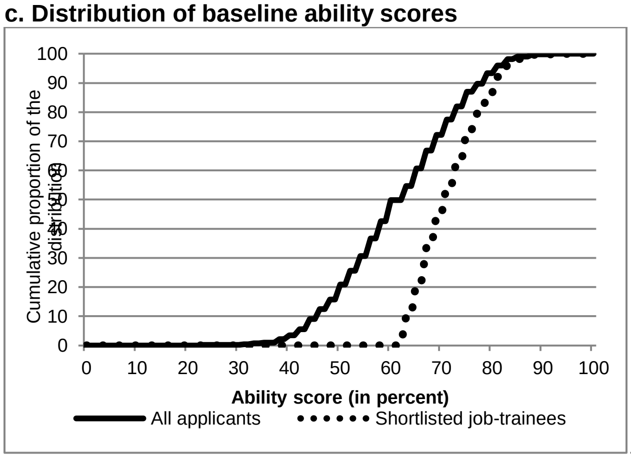
a. Distribution of numeracy scores



b. Distribution of literacy scores



c. Distribution of baseline ability scores



Source: Author's calculations.

Training and Selection by Recruiter

The 278 job seekers who advanced to job training attended a pre-training information session. During this session, job trainees were provided with materials required for training and logistical information related to the training process. They were also informed about the opportunity to participate in this research study. A total of 268 applicants (95 percent) of the applicant pool opted to participate. Consenting participants were asked to self-administer a baseline questionnaire, after which they were issued their randomly assigned probabilistic job guarantee. All 278 job trainees were invited to attend three days of full-time training and further screening.[15] During training, applicants were monitored for their punctuality and the quantity and quality of active participation in the job training. Trainees were also taught and tested on materials specific to the health survey during the training.

In addition, for the purposes of this study, on each day of training, respondents were asked to self-administer a survey questionnaire. The recruitment team did not know who chose to participate in the research, what alternative job probabilities were assigned, or which participants completed the daily questionnaires. Moreover, the recruiter did not get access to the survey questionnaires. This information was carefully explained to respondents, who were monitored to ensure confidentiality regarding participation in the research study.

[15] They were paid a wage equivalent to half of the daily wage of the employment opportunity for each day of training attended.

At the end of the final day of training, the alternative job draws were conducted, and participants learned their alternative job employment realization. The recruiting team was not present at this time, and they were not at any point informed about who had received an outside job offer. Two days after completing the training, the successful applicants for the job advertised by the recruiter were contacted.

Intervention: Probabilistic Outside Employment Options

During the information session and before commencement of the job training, job trainees were randomly assigned some *probability of employment* via a job guarantee for an alternative job. The six different probabilistic guarantees were a 0, 1, 5, 50, 75, or 100 percent chance of an alternative job. Thus, the intervention experimentally altered individuals' outside options.

The alternative jobs were constructed to mimic, as closely as possible, the jobs offered by the recruiter. The alternative jobs were for the same duration and pay as the job being offered by the recruiter. They were real jobs, requiring real effort and paying real wages. Although the recruiter was hiring for interviewer positions, the alternative jobs were other research jobs. In both cases, individuals were working for research projects for the same university, albeit on different projects and performing different types of research tasks. The alternative jobs included data entry, translation, transcription, and archival research.[16]

Individual treatment status was blind to the research and recruitment team but known to the job trainee. Each job trainee was given an envelope with his employment identification number written on it. Inside the sealed envelope was an employment contract stating which probabilistic job guarantee he had received. Job trainees assigned a 0 percent chance of an alternative job also received an envelope. Randomization was conducted at an individual level and stratified on quintiles of baseline ability and an indicator variable for whether they had ever worked for the recruiter.[17] Baseline ability was determined using participants' scores from the numeracy and literacy components of the pre-training assessment. The distribution of the probabilities was pre-assigned to the 278 applicants invited to attend the training stage of the recruitment process. Ten individuals opted not to participate in the research project or in the recruitment process. These participants made their decision before knowing to which treatment group they had been assigned. In the final sample of the 268 male participants, the distribution of the probabilistic job guarantees is similar to the intended assignment (see Table 2.1, Panel A).

Table 2.1 Sample and attrition

Panel A. Sample (pretreatment): Treatment assignment								
		All	**0%**	**1%**	**5%**	**50%**	**75%**	**100%**
Sample frame (intended)	N	278	55	56	56	56	28	27
	%		0.198	0.201	0.201	0.201	0.101	0.097
Main Sample (actual)	N	268	53	56	52	54	28	25
	%		0.198	0.209	0.194	0.201	0.104	0.093

[16] If individuals were selected by the recruiter *and* received an alternative job, they were required to take the recruiter's job and not the alternative job.

[17] "Ever worked for the recruiter" is broadly defined—even individuals who had attended a prior job training session held by the recruiter but had never successfully been employed are included in this category.

Table 2.1 Continued

Panel B. Training participation and survey data completion

	Administrative data	**Survey questionnaires**		
	Attended training	**Pretreatment**	**Post-treatment**	
	Every day	**Baseline**	**At least once**	**Every day**
	(1)	**(2)**	**(3)**	**(4)**
0% job guarantee	0.906	0.981	0.906	0.830
	[0.041]	[0.019]	[0.041]	[0.052]
1% job guarantee	0.964	0.946	0.964	0.946
	[0.025]	[0.030]	[0.025]	[0.030]
5% job guarantee	0.923	0.942	0.981	0.865
	[0.037]	[0.033]	[0.019]	[0.048]
50% job guarantee	0.944	1.000	0.944	0.870
	[0.032]	[0.000]	[0.032]	[0.046]
75% job guarantee	0.964	1.000	0.964	0.893
	[0.035]	[0.000]	[0.035]	[0.059]
100% job guarantee	0.960	1.000	1.000	0.960
	[0.040]	[0.000]	[0.000]	[0.040]
Average of dependent. variable	0.940	0.973	0.955	0.888
N	268	268	268	268
p-value of F-test:				
All (jointly equal)	0.810	0.068	0.031	0.221
0% and 1%	0.220	0.334	0.220	0.055
0% and 100%	0.339	0.319	0.021	0.049
1% and 100%	0.927	0.080	0.156	0.786
50% and 100%	0.759	.	0.079	0.142
75% and 100%	0.936	.	0.315	0.346

Source: Author's calculations.

Notes: The sample frame consists of 278 participants who were short-listed for training by the recruiter. Panel A shows the intended and actual assignment of the job probabilities. These distributions differ due to ten participants who opted out of the research study (prior to learning their treatment status). The main sample used in this paper consists of 268 individuals. Panel B presents average participation rates by treatment group in training and survey data completion rates by treatment group. A partial set of *p*-values from pair-wise comparisons of treatment group means are presented. All those that are not presented have *p*-values greater than 0.10. The full set of results is available on request.

The distribution of treatment allocated approximately 20 percent of the sample to each of the 0, 1, 5, and 50 percent chance groups and approximately 10 percent of the sample to each of the 75 and 100 percent chance groups.[18] Respondents were informed about the distribution of the alternative job probabilities before learning their own treatment assignment, which ensured that all participants had the same beliefs about the distribution. Had respondents not been told the underlying distribution, then individuals would have had variable information, which would have been endogenous to the truthfulness, candor, and size of their social network among other job trainees.

[18] Although equal proportions across groups was desirable, this was not feasible due to budgetary limitations.

Job trainees were also informed that their treatment assignment was confidential. It was consistently emphasized that their probability of an alternative job would have no direct bearing on their probability of being hired by the recruiter. To ensure that individuals were clearly informed about how the probabilities worked and how the draws would be conducted, this information was discussed in detail, and multiple demonstrations were conducted to illustrate the process.

An important concern was whether individuals actually understood the probabilistic nature of the alternative job offers. After the treatment was explained, but before individuals learned their own probability, participants were surveyed to elicit their perceptions related to their understanding of these probabilities. Participants were asked for each treatment arm what they expected the realization of alternative jobs to be. For example, "If 60 participants receive the 50 percent job guarantee, how many of them are likely to receive an *alternative* job?" Modal responses by participants were fairly accurate. Although the modal response for the 1 and 75 percent groups were slight overestimates (at 1.6 and 83 percent, respectively), for 5 and 50 percent, the modal responses were precisely accurate.[19] These data are not differential by treatment status. In general, it seems reasonable to assume that participants understood the assigned outside options.[20]

[19] Given the phrasing of the question for the 1 percent chance treatment group, it was impossible for individuals to select an integer that would map into 1 percent of the distribution getting alternative jobs. The modal response was one person, which maps to 1.6 percent. The second most frequent response recorded was zero.

[20] Open-ended questions on the survey asked respondents to explain how they understood the job probabilities; the responses suggest that respondents generally understood the alternative job probabilities. For example, "The probability criteria are dependent on the chance and not merit of a person in terms of experience and qualification"; "Those that have 75 percent chance have higher chances as compared to those that have 1percent chance"; and "It's a good idea after all if you are guaranteed a 100 percent probability as you don't have to worry about the other job."

3. DATA

Baseline Data

Administrative data from the recruiter was combined with respondent survey data.

Prescreening assessment test (administrative data): The prescreening assessment conducted to select the job trainees provides baseline numeracy and literacy ability scores. For the short-listed candidates—the sample frame for this paper—the average numeracy and literacy scores are 63 and 80 percent, respectively.[21] The ability score used in the paper is a composite measure equal to the normalized measure of the sum of the individual's normalized literacy and numeracy scores. The distributions of these measures are presented in Figure 2.2.

Baseline questionnaire (survey data): The survey instrument was self-administered during the information session to consenting participants before training commenced. It included questions about previous work experience, employment perceptions and attitudes, time use, and a work and health retrospective calendar history.

Training and Post-training Data

I used administrative data collected during the training, as well as the hiring decisions made by the recruiter, to construct the key outcome variables of interest used in the analysis. I supplemented this with daily follow-up survey questionnaires conducted during the training.

Participation in training: Table 2.1, Panel B, presents the participation rates of the 268 consenting participants. Most of the selected job trainees opted to participate in the training, with 94 percent attending training every day. Importantly, there is no statistically significant difference in training participation across treatment groups.

Test scores: On each training day, the recruiter administered a test to job trainees. These tests assessed comprehension of the materials taught during the training sessions and were the most important observable performance indicator used by the recruiter in making employment decisions. (Refer to Appendix A for a detailed discussion on the determinants of hiring decisions.)

Punctuality records: Recruitment staff recorded daily attendance, including job trainee arrival times. Participants were required to sign in, at which time their arrival time was recorded.

Contribution records: Recruitment staff also recorded the verbal contributions made by job trainees. These records enable construction of a performance indicator of engagement. Similar measures of engagement have been used in the education literature, typically in the context of teacher evaluations of student engagement (for example, Dee and West [2011] and Fredricks, Blumenfeld, and Paris [2004] reviewed the education literature pertaining to student engagement). I focus on the subjective assessment of the quality of the contribution made. The quality scale is graded as *good, neutral,* or *bad*. In some cases, multiple members of the recruitment team were documenting these contributions. To eliminate double counting, a contribution is counted only once, assuming that it came within 5 minutes of a second contribution. When a contribution is recorded twice and the two records differ, the lowest quality assessment is used. The double counting allows assessment of the correlation in subjective assessments made. In 61.5 percent of cases, the two separate records are in agreement.[22]

Employment records: I also obtained the employment records of the recruiter providing employment outcomes for the job trainees.

Daily survey questionnaires: Respondents also completed daily self-administered follow-up questionnaires. While respondents completed these surveys in private, all recruitment staff left the training venue. However, research staff were available to address any questions. Participants were asked

[21] Among the 554 job applicants, the average numeracy score was 52.5 percent, and the average literacy score was 70.3 percent.

[22] In only two cases in which the quality assessments differ does one report assess the contribution as *good* and the other as *bad*.

to leave their completed questionnaires in a sealed drop box at the venue. The daily questionnaire asked about time use and mental and physical health, as well as employment attitudes and beliefs.

Table 2.1, Panel B, presents survey data completion rates. There is some evidence of differential nonresponse with the follow-up questionnaires by treatment status.[23] Only 83 percent of the participants who received no chance of an alternative job completed the follow-up survey questionnaire everyday, as compared with 96 percent of those who were assigned a 100 percent chance of an alternative job. This difference of 13 percentage points is significant at the 5 percent level. These follow-up data are primarily used to shed light on the potential mechanism driving the performance results—in particular, to examine the impact of the outside option on short-run effort. To address the differential nonresponse in the survey data, I conducted a number of robustness checks, including weighting, Lee bounds, and Horowitz and Manski bounds.

Sample

The sample in this paper comprises the 268 consenting job trainees. Table 3.1 presents summary statistics about the sample. On average, respondents were 25 years old, and 18 percent were married. Approximately 17.6 percent of the men in the sample had at least one child. Because of the recruiter's secondary schooling requirement, respondents were relatively well educated for Malawi, with an average of 13 years of education. Respondents reported earnings of approximately $220 over the preceding three months.

Table 3.1 Summary statistics and balancing tests

		Treatment assignment						
Baseline characteristics:	**N (1)**	**0% (2)**	**1% (3)**	**5% (4)**	**50% (5)**	**75% (6)**	**100% (7)**	**F-stat[1] (8)**
Demographics:								
Age	268	25.887	25.893	24.865	25.463	26.464	25.240	0.757
		[5.176]	[4.735]	[4.334]	[3.490]	[5.903]	[4.612]	
Married	268	0.189	0.250	0.135	0.093	0.250	0.120	0.207
		[0.395]	[0.437]	[0.345]	[0.293]	[0.441]	[0.332]	
# of children	250	0.388	0.431	0.277	0.132	0.560	0.200	0.154
		[0.909]	[0.922]	[0.743]	[0.520]	[1.083]	[0.577]	
Income (in USD, 3 months)	225	181.72	247.12	167.54	199.88	294.96	282.83	0.240
		[203.65]	[272.13]	[187.69]	[203.72]	[299.11]	[342.76]	
Education, ability, and experience:								
Years of schooling	268	13.264	13.071	13.115	13.130	13.107	13.600	0.277
		[0.858]	[0.931]	[1.041]	[0.953]	[0.786]	[1.000]	
Ability (standardized)	268	-0.075	-0.006	-0.020	0.034	0.116	0.010	0.978
		[0.960]	[1.021]	[0.989]	[1.063]	[0.992]	[1.013]	
Ever worked	268	0.906	0.857	0.750	0.944	0.929	0.840	0.083
		[0.295]	[0.353]	[0.437]	[0.231]	[0.262]	[0.374]	
Worked last month	252	0.600	0.647	0.638	0.577	0.536	0.792	0.357
		[0.495]	[0.483]	[0.486]	[0.499]	[0.508]	[0.415]	
Any work in past 6 months	252	0.780	0.902	0.894	0.808	0.893	0.958	0.137
		[0.418]	[0.300]	[0.312]	[0.398]	[0.315]	[0.204]	
Months worked (maximum 6)	252	2.820	2.922	2.468	2.538	2.429	3.083	0.759
		[2.371]	[2.226]	[2.155]	[2.313]	[2.116]	[2.225]	

[23] Completing the daily questionnaires was not a condition of receiving the alternative job.

Table 3.1 Continued

Baseline characteristics:	N (1)	0% (2)	1% (3)	5% (4)	50% (5)	75% (6)	100% (7)	F-stat[1] (8)
				Treatment assignment				
p-value from F-tests of joint significance of covariates:[3]								
Compared to all other groups			0.175	0.395	0.400	0.060	0.146	0.223
Compared to 0%				0.006	0.397	0.098	0.210	0.014
Compared to 1%					0.782	0.009	0.559	0.147
Compared to 5%						0.468	0.405	0.772
Compared to 50%							0.078	0.025
Compared to 75%								0.004

Source: Author's calculations.

Notes: The table reports group means or proportions (where applicable—for example, married). Standard deviations are reported in parentheses. The main sample of 268 participants is used. Data used derive from both the baseline self-administered questionnaire and administrative data collected by the recruiter. Income is measured in USD (US dollars) and includes all self-reported income from the past three months, including the following explicit categories: Farming; Ganyu (piecework); Formal employment; Own business; Remittances; Pension; and Other. The ability scores are a composite measure of literacy and numeracy scores and are presented in standardized units. See Figures 2.2a, 2.2b, and 2.2c for the distribution of these scores. [1] These *p*-values correspond to the joint F-test of the means and proportions being equal across all treatment groups. [2] This refers to the number of pair-wise comparisons between treatment groups that are statistically significant at the 5 percent level. A total of 15 comparisons are made for each variable. [3] These *p*-values are associated with the joint F-test for whether all the covariates listed are jointly equal in predicting assignment to the treatment group.

Most of the men (86.9 percent) reported having worked previously. Although most men (86.1 percent) had worked at some point during the previous 6 months, they had only worked, on average, 2.7 months of the preceding 6 months. Individuals who had previously worked were asked a series of questions about their three most recent jobs. For their most recent job, 58 percent reported competing for it, 26.8 percent were required to take a test as part of the hiring process, almost 70 percent were required to attend an interview, and slightly more than half were required to complete some job training prior to employment.[24] In sum, the process was not atypical of hiring processes in this context for this sample.

[24] Averages across the three most recent jobs are similar (results not shown).

4. ESTIMATION STRATEGY

Performance was measured using administrative records from the recruiter, including test scores from the training assessment tests, as well as the quality of engagement of the job trainee in training. I used the number of *good*, *neutral,* and *bad* contributions made during the training and constructed a performance index measure as a summary index of these performance indicators. The index was constructed as the average of the normalized values of each of these measures (Kling, Liebman, and Katz 2007).

To measure effort, both administrative data and survey data were used. From the administrative data, I used the rich arrival data and construct measures of punctuality: "ever late," "always late," and "how early or late." From the survey data, the self-reported time use diaries indicated the average number of hours per day spent studying training materials and the amount of time spent on leisure activities (for example, watching television or listening to the radio). I constructed an effort index as a summary measure of effort using the average of the normalized values of the minutes arrived late and time use variables. Lastly, I examined employment outcomes using data from the recruiter regarding which respondents were hired.

The following regression estimates the differential performance, effort, and employment by treatment group:

$$Y_i = \beta_1 T0_i + \beta_2 T1_i + \beta_3 T5_i + \beta_4 T50_i + \beta_5 T75_i + \beta_6 T100_i + X_i{}'\beta + \varepsilon_i, \tag{1}$$

where Y_i indicates job trainee i's average performance or effort. The average for each indicator is constructed using data from three observations per individual. In the case of missing data, the average is constructed from the observations available. The indicators $T0$, $T1$, $T5$, $T50$, $T75$, and $T100$ are binary variables equal to 1 if the individual received a 0, 1, 5, 50, 75, or 100 percent chance of an alternative job, respectively, and 0 otherwise. Rather than assuming a linear relationship, I specifically allowed a flexible nonlinear relationship between the probabilistic job guarantees and the outcome variables of interest, which allowed for examination of the reduced-form relationship between employment risk and performance and effort.

Lastly, X_i is a vector of covariates, including stratification cell fixed effects, ability score, previous experience with the recruiter, age, and other background characteristics. To facilitate easier interpretation of the coefficients, I demeaned all control variables, so that coefficients are interpretable as group means at the mean of all controls in the regression. The main comparison of interest is between those assigned no outside option ($T0$) and those assigned the alternative employment guarantee ($T100$), removing all risk from the job application process. Although employment risk is decreasing in the magnitude of the outside option, uncertainty of the alternative job offers is highest among those in the 50 percent group. I do, however, present the average performance, effort, and employment results for all treatment groups, yielding insights into the relationship of these outcomes across the distribution of the outside options assigned.

Given the random assignment treatment status, the identification assumption that assignment to treatment group is orthogonal to the error term should hold. One test of this assumption is to compare observable characteristics across the different treatment arms. Table 3.1 shows that the different treatment arms appear to be balanced when examining multiple baseline characteristics. In most cases, I cannot reject the null hypothesis that the means are jointly equal across all the treatments. Similarly, for most pair-wise comparisons, I cannot reject equality of the means. Covariates are included to increase precision, but the results are robust, whether or not controls are included. Moreover, as an additional specification check, I constructed a measure of the extent to which omitted variable bias would have to differ in unobservables relative to observables to explain away the observed differences in performance and effort by treatment group (Altonji, Elder, and Taber 2005; Bellows and Miguel 2009). In all cases, the results suggest that the extent of omitted variable bias introduced in unobservables would have to well exceed that observed in observables to explain the estimated coefficients. (Details are provided in Appendix B.)

Further specification checks were conducted to test the sensitivity of the results to missing data in the administrative records and differential nonresponse in survey data records. I used three strategies. First, I followed Fitzgerald, Gottschalk, and Moffitt (1998) to present weighted results. Second, I presented conservative bounded results in which I implemented min-max bounds (Horowitz and Manski, 1998). Third, I restricted the sample to the 0 and 100 percent treatment groups and estimated Lee (2009) bounds on the average treatment effect of the 100 percent group relative to the 0 percent treatment group. The results are robust to these specification checks. (Details are provided in Appendix C.)

5. INCENTIVES AND STRESS: PREDICTIONS

Essentially, the variation in outside options generated by the experiment changes the incentive to perform. Absent the outside options, performance is rewarded with employment for individuals who reach the hiring threshold. Standard economic theory predicts that reducing the incentive to perform (by offering the outside options) should lead to decreased effort. Economic models typically assume that performance is monotonically increasing in effort; thus, reducing the incentive to perform should also reduce performance. Prendergast (1999) reviewed the provision of incentives in the workplace and largely found that incentives have the intended effect on the incentivized outcome, particularly in the case of simple tasks. This review touched on some cases in which incentives fail to lead to the intended outcome. Kamenica (2012) reviewed the more recent literature, focusing on the empirical evidence in which incentives have had anomalous effects. A number of behavioral theories have been put forth to explain these anomalous incentive effects, with the one most relevant to the context studied here being *choking under pressure.*

In addition to reducing the incentive to perform by increasing the value of the outside option, reduced employment risk is also likely to make the recruitment process less stressful to job seekers. The combination of reduced incentives and reduced stress leads to ambiguous predictions for performance when outside options improve.

Intuitively, an improvement in an individual's outside option reduces the marginal benefit of any particular employment opportunity. Therefore, the optimal level of effort should decline as outside options improve, assuming that the cost of effort is not zero. If performance is increasing in effort, then as outside options improve, performance will decline.

In the recruitment setting, assume that p is the probability of being hired in the current recruitment process, and w is the wage associated with the recruiter's job. The probability of being hired is assumed to be a positive and monotonically increasing function of performance (a realistic assumption for this recruiter's hiring process); thus, it follows that performance and employment are also positive and monotonically increasing functions of effort. In addition, b is the expected value of the individual's outside option (that is, his probability of outside employment multiplied by the expected wage of outside employment). Finally, assume that effort is costly. Therefore, an individual selects effort level $e*$ to maximize

$$Max_e U = p.w + (1-p)b - c(e),$$

subject to

$$p = f(e), f'(e) > 0 \text{ and } f''(e) \leq 0$$

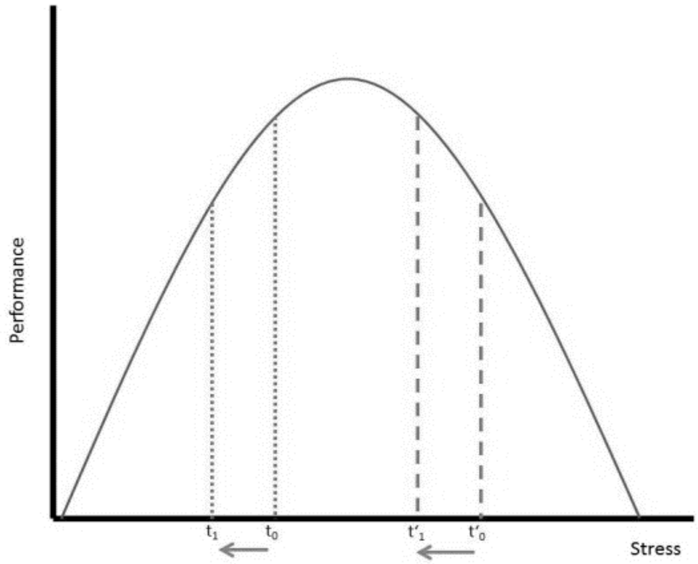
$$c(e) \geq 0, c'(e) > 0 \text{ and } c''(e) \geq 0.$$

In this case, comparative statics imply that performance and employment are both rising in effort, whereas effort is declining in the job seeker's outside option.

However, a second channel through which employment risk may affect performance is through its impact on stress. Extensive literatures in both psychology and public health show that unemployment is stressful, as is perceived job insecurity (Feather 1990; de Witte 1999, 2005; Burgard, Brand, and House 2009). Therefore, it is reasonable to assume that stress is a decreasing function of an individual's outside options—that is, $s = s(b)$, and $s' < 0$. Therefore, reducing the risk of unemployment should reduce stress.

The Yerkes-Dodson law (1908) maps the relationship between stress and performance; performance has been shown to be inverse U-shaped in stress. As stress increases, performance improves up to a bliss point, beyond which performance declines as stress continues to increase (Figure 5.1). Incorporating this prediction means that performance is a function of not only effort but also stress—that is, $p = f(e, s)$. Also, f'_s varies by s.

Figure 5.1 Relationship between stress and performance



Source: Yerkes-Dodson (1908).

As outside options improve (that is, *b* increases), stress (*s*) should decline; however, the impact on performance is ambiguous. Therefore, the stress effect induced by reduced risk should either always be positive or always be negative; or else it should first increase and then decrease across the risk distribution, depending on the underlying values of *s*. In this setup, risk affects performance through both the incentive and stress channels. Therefore, the predicted relationship between risk and performance is ambiguous and depends on the relative size of the incentive and stress effects, as well as on whether the level of risk puts the individual in the increasing or decreasing portion of the Yerkes-Dodson curve. Effort is unambiguously decreasing as the value of the outside option increases. Thus, there are three possibilities when employment risk declines.

- *Performance and effort decline:* If the stress effect leads to a performance decline, then the stress and incentive effects work in the same direction, and effort and performance should decline. Alternatively, if the stress effect leads to a performance improvement but is smaller in magnitude than the negative incentive effect on effort, then effort and performance will both decline.
- *No change to performance, but effort declines:* If the stress effect leads to a performance improvement that exactly offsets the negative incentive effect on effort, then there will be no net effect of employment risk on performance for a lower level of effort.
- *Improved performance, but effort declines:* If the stress effect leads to a performance improvement that exceeds the magnitude of the incentive effect's reduction in effort, then effort will decline and performance will improve.

This framework suggests that ultimately it is an empirical question whether decreasing employment risk will positively or negatively affect job-seeker performance.

6. RESULTS

Performance Indicators

The main performance indicator is job trainees' performance on administered tests and engagement in the training. A secondary outcome of interest is the quality of verbal engagement made by participants during training.

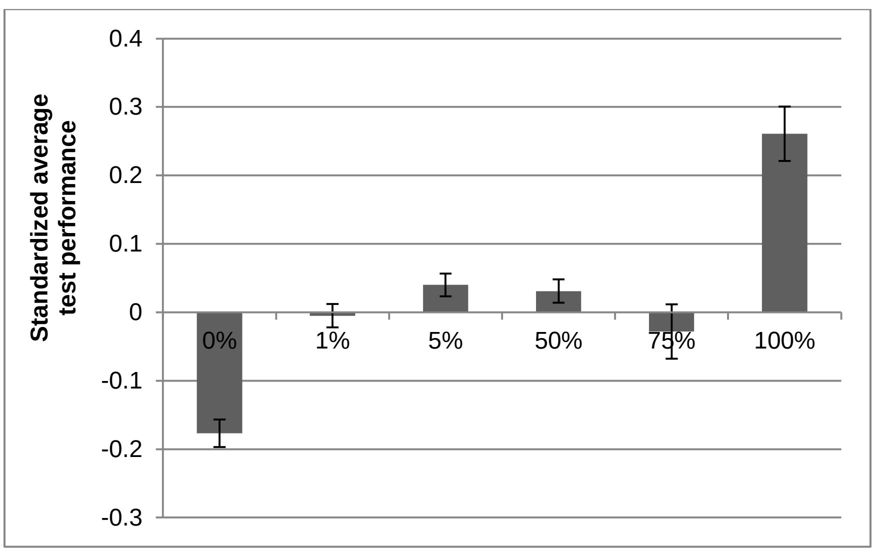
Administrative Training Tests

The recruiter-administered training test scores are the most important assessment tool used by the recruiter for hiring decisions. The R-squared of a univariate regression of employment on the standardized average test score is 0.357 and the coefficient is 0.225 (s.e. = 0.031), implying that for every additional standard deviation in a job trainee's test score, an individual is 22.5 percentage points more likely to be hired. (The determinants of hiring are discussed in Appendix A.)

Figure 6.1 and Table 6.1 present the main test results. The dependent variable is the standardized average performance of the three standardized test scores. Job trainees assigned no outside option performed significantly worse than those assigned a 100 percent outside option. The magnitude of the difference ranges from 0.438 to 0.451 standard deviations, depending on the set of controls used. These effects sizes are large and consistently statistically significant.[25] Individuals with the lowest incentive to perform performed the best during the job-seeking process.

Figure 6.1 and Table 6.1 also show substantial nonlinearities in the performance–risk relationship. However, the discussion here primarily focuses on the difference between individuals assigned no outside option and those assigned a guaranteed outside option, as this comparison provides the cleanest measure of reduced risk.

Figure 6.1 Average standardized test scores by treatment group



Source: Author's calculations.
Note: This figure presents the estimated group means, controlling for covariates and stratification cell fixed effects.

[25] Perhaps the best way of contextualizing the effects is to compare them to education interventions that aim to have an impact on test scores in developing countries. Kremer and Holla (2009) reviewed randomized controlled trials of 26 education interventions conducted in developing countries. Test score effect sizes range from 0 to 0.46 standard deviations, with the exception of a technology-assisted education intervention in Nicaragua, which found an effect size of 1.5 standard deviations (Heyneman et al. 1981). The median effect size in the Kremer and Holla (2009) review is 0.16 standard deviations. Therefore, the effect sizes in this paper are approximately three times larger.

Table 6.1 Average performance on training tests by treatment group

Indicator	Average training test score (1)	(2)	(3)
0% job guarantee	-0.176	-0.19	-0.177
	[0.147]	[0.142]	[0.142]
1% job guarantee	-0.015	-0.009	-0.005
	[0.136]	[0.126]	[0.126]
5% job guarantee	0.041	0.066	0.04
	[0.132]	[0.113]	[0.119]
50% job guarantee	0.041	0.039	0.031
	[0.124]	[0.119]	[0.122]
75% job guarantee	-0.039	-0.037	-0.028
	[0.241]	[0.209]	[0.207]
100% job guarantee	0.259	0.261	0.261
	[0.195]	[0.200]	[0.198]
Observations	258	258	258
R-squared	0.01	0.19	0.2
Stratification cell fixed effects?	No	Yes	Yes
Includes controls?	No	No	Yes
p-value of F-test:			
0% and 100%	0.076	0.069	0.073

Source: Author's calculations.

Notes*:* This table presents mean performance on the recruiter-administered training tests by treatment group. The average standardized test score is constructed by taking the average of the standardized test score from the three tests. Individual tests are standardized by using the sample mean and standard deviation for the relevant test. Treatment status was randomly allocated and stratified by ability quintile and prior work experience with the recruiter. The stratification cell fixed effects include a set of dummies for each stratification cell. The set of additional covariates include a dummy variable for whether the individual has worked before, marital status, age, and the individuals' standardized ability score. For covariates with missing observations, the variable is assigned the mean value of the variable, and an indicator variable is included for whether that particular variable is missing. Robust standard errors are presented.

Verbal Contributions

The recruitment team monitored and assessed verbal contributions made during the training sessions. Appendix A highlights the importance of good-quality engagement during training, as it is a key predictor of employment in the current context.[26] Using these records, I constructed a measure of *quality engagement,* defined as the number of good-quality contributions made across the three-day training period. More than half of the participants (67 percent) made a contribution at least once during the training. Individuals who contributed did so an average of 2.3 times. Approximately 46 percent of the contributions made were classified as *good*, 39 percent as *neutral,* and 15 percent as *bad*. Table 6.1 presents these results.

[26] Classroom behavior in schools has also been shown to be important for labor market success (Segal 2008, 2012, 2013). I have a similar measure of behavior to that used in this literature. However, in my setting, training classroom behavior was not an important predictor for determining employment outcomes (see Appendix A). Recruitment staff recorded disruptions by participants during the training sessions. Almost half of the participants (47.1 percent) were recorded as being disruptive at some point during the training. In almost half of those cases, the disruptive behavior related to making noise, chatting with friends, banging on desks, and so on. Another 42 percent related to unnecessary moving around the room or entering and exiting the training room. The remainder (11 percent) referred to the answering of cell phones during training. Using this data, I constructed measures of whether the job trainee was ever disruptive, the number of times he was disruptive, and the number of each type of disruption. I did not observe statistically significant differences across treatment groups.

Participants receiving the guaranteed outside option made 0.410 additional *good* contributions relative to those in the 0 percent job probability treatment group. This difference is statistically significant at the 10 percent level. In fact, individuals in the 0 percent group were the least likely of all groups to make a *good* contribution (only 0.528 contributions on average).[27] This is consistent with the test performance results, which showed that individuals assigned no outside option performed the worst on average, whereas those guaranteed a job performed the best (Table 6.1, Column 3). If individuals with the guaranteed outside option simply made many more contributions, then this measure may reflect increased participation due to a lower cost of being incorrect.[28] However, Table 6.1, Column 3, also shows that individuals receiving the guaranteed outside option made fewer bad contributions, though this is not statistically significant.[29]

Performance Index

To examine overall shifts in performance accounting for both test scores and quality of verbal engagement, I created a performance index. This index is the mean of the normalized value of the average test score and the number of good- and neutral-quality contributions (Kling, Liebman, and Katz 2007). Table 6.2, Column 4, presents these results. Individuals assigned no outside option performed 0.456 standard deviations worse than those assigned the guaranteed outside option. The difference is statistically significant at the 5 percent level.

Table 6.2 Average performance indicators by treatment group

	Engagement in training (contributions)			Performance
	# good	# neutral	# bad	index
Dependent variable	(1)	(2)	(3)	(4)
0% job guarantee	0.528	0.612	0.363	-0.198
	[0.099]	[0.130]	[0.108]	[0.093]
1% job guarantee	0.795	0.585	0.195	0.021
	[0.134]	[0.126]	[0.090]	[0.095]
5% job guarantee	0.690	0.705	0.209	-0.011
	[0.156]	[0.129]	[0.059]	[0.111]
50% job guarantee	0.767	0.386	0.224	0.022
	[0.135]	[0.095]	[0.064]	[0.098]
75% job guarantee	0.705	0.480	0.072	-0.047
	[0.155]	[0.123]	[0.050]	[0.130]
100% job guarantee	0.938	1.035	0.273	0.249
	[0.193]	[0.244]	[0.112]	[0.161]

[27] This also goes against a story in which the 0 percent group is more risk averse and, as a result, set a higher bar on their ex ante beliefs about contribution quality. This implies fewer, but higher average quality of, contributions from the 0 percent group.

[28] Evidence suggests that overall job trainees receiving the 100 percent chance of the alternative job made more contributions in general (11.2 percentage points relative to those receiving no outside option). Probit results are broadly consistent.

[29] An alternative approach for examining the impact on the quality of contributions made and accounting for all contributions is to construct an index. I constructed an index by weighting good-quality contribution by 1, a neutral contribution by 1/2, and a bad contribution by –1/2. Individuals assigned to the guaranteed outside option made 0.8 additional good-quality contributions relative to those assigned no outside option, representing a 100 percent increase in the number of good-quality contributions. (Results available upon request)

Table 6.3 Continued

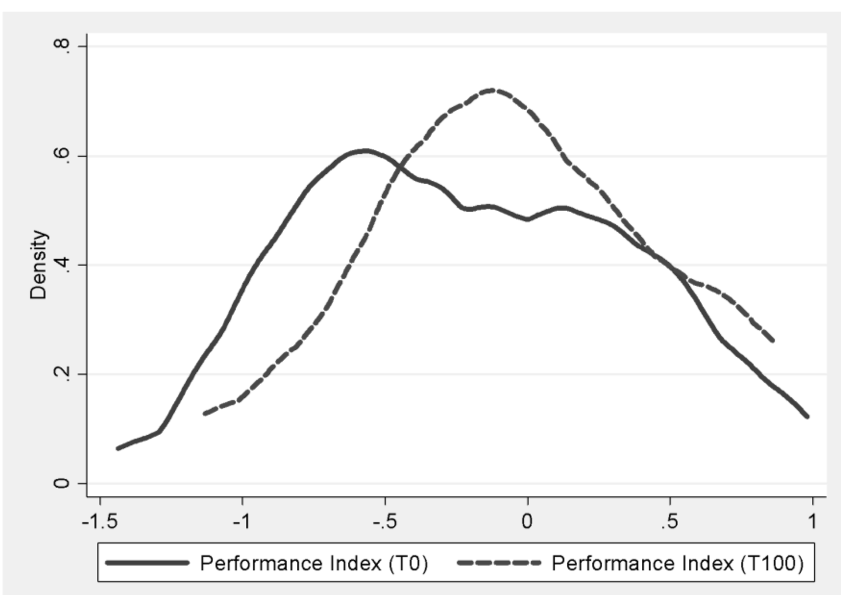
Dependent variable	Engagement in training (contributions) # good (1)	# neutral (2)	# bad (3)	Performance index (4)
Observations	268	268	268	268
R-squared	0.415	0.354	0.170	0.157
Stratification cell fixed effects?	Yes	Yes	Yes	Yes
Includes controls?	Yes	Yes	Yes	Yes
p-value of F-test:				
0% and 100%	0.058	0.127	0.571	0.017

Source: Author's calculations.

Notes: This table presents mean performance as measured by engagement recorded by the recruiter by treatment group. "Any contribution" is a binary indicator if the job trainee ever engaged verbally in training. The "total number of contributions" is the cumulative number of contributions made by the job trainee during the three days of training and then separated by quality as determined by the recruitment staff. The performance index is a summary measure of the performance indicators, constructed by taking the average of the normalized values of "Average test score," and "Number of good contributions." Treatment status was randomly allocated and stratified by quintile ability and prior work experience with the recruiter. The stratification cell fixed effects include a set of dummies for each stratification cell. The set of additional covariates include a dummy variable for whether the individual has worked before, marital status, age, and the individuals' standardized ability score. For covariates with missing observations, the variable is assigned the mean value of the variable, and an indicator variable is included for whether that particular variable is missing. Robust standard errors are reported.

Using this composite performance measure, it is useful to consider the differences among performance index distributions across treatment arms. Figure 6.2 presents the distribution of this index for individuals assigned no outside option (*T*0) and those assigned the employment guarantee (*T*100). This figure shows that the performance distribution for those guaranteed outside employment is shifted quite significantly to the right. The *p*-value associated with a Mann-Whitney distribution test of equality is 0.027 (0.078 for the standardized test scores alone), which indicates a shift in the performance distribution rather than the average performance result being driven by a few outliers.

Figure 6.2 Performance index distribution



Source: Author's calculations.

In sum, the findings show that performance is highest among those assigned a guaranteed outside option and lowest among those assigned the lowest outside option. Differences are large in magnitude and often statistically significant. The results are robust to a number of specification checks. (For further details, refer to Appendixes B and C.) The results suggest that reducing employment risk can, at least in this context, result in overall higher performance. This reduced-form result is interesting in its own right; however, to further understand this relationship, it is important to consider how effort is affected by the outside options.

Effort Indicators

Punctuality

Punctuality is one indicator of effort. On average, job trainees arrived 21 minutes early to training.30 I used three different measures to capture punctuality: binary indicators equal to 1 if the job trainee was either ever or always late and a continuous indicator capturing the average minutes early or late across the three training days. Table 6.3 shows that individuals assigned to the 100 percent treatment group were 9.3 percentage points more likely to ever arrive late and 6.3 percentage points more likely to be always late compared to those assigned no outside option. These differences are large in magnitude but are not statistically significant ($p = 0.34$; $p = 0.271$). Probit results are broadly consistent. Table 6.3, Column 3, presents average minutes arrived early or late. No statistically significant differences in arrival times were observed.

Table 6.3 Mean effort by treatment group

	Administrative data			**Survey data**		
Dependent variable	**Ever late** (1)	**Always late** (2)	**Minutes early or late** (3)	**Studied (hours)** (4)	**Radio/TV (hours)** (5)	**Effort index** (6)
0% job guarantee	0.183	0.017	-24.400	1.179	1.155	0.214
	[0.053]	[0.020]	[2.156]	[0.131]	[0.123]	[0.083]
1% job guarantee	0.185	0.001	-21.405	1.148	1.582	0.000
	[0.052]	[0.003]	[1.856]	[0.110]	[0.132]	[0.079]
5% job guarantee	0.321	0.020	-19.187	0.951	1.356	-0.088
	[0.065]	[0.021]	[2.394]	[0.100]	[0.160]	[0.090]
50% job guarantee	0.175	0.019	-21.747	1.096	1.512	0.017
	[0.056]	[0.020]	[2.146]	[0.099]	[0.133]	[0.069]
75% job guarantee	0.254	0.039	-19.846	1.139	1.408	0.026
	[0.087]	[0.039]	[3.177]	[0.140]	[0.166]	[0.118]
100% job guarantee	0.276	0.080	-19.179	0.750	2.037	-0.373
	[0.091]	[0.055]	[4.153]	[0.079]	[0.247]	[0.144]

[30] Approximately 16 percent arrived late on the first day, 11 percent on the second, and only 5 percent on the final day. Evidently, job trainees realized that their punctuality was being recorded and altered their behavior over time.

Table 6.3 Continued

	Administrative data			Survey data		
Dependent variable	**Ever late** (1)	**Always late** (2)	**Minutes early or late** (3)	**Studied (hours)** (4)	**Radio/TV (hours)** (5)	**Effort index** (6)
Observations	259	259	259	254	254	259
R-squared	0.270	0.070	0.657	0.689	0.707	0.104
Stratification cell fixed effects?	Yes	Yes	Yes	Yes	Yes	Yes
Includes controls?	Yes	Yes	Yes	Yes	Yes	Yes
p-value of F-test:				-		-
0% and 100%	0.340	0.271	0.247	0.005	0.002	0.001

Source: Author's calculations.

Notes: This table presents the average effort by treatment group using both administrative data and survey data. "Always late" is a binary indicator equal to 1 if the job trainee always arrived late for training. "Ever late" is a binary indicator equal to 1 if the job trainee ever arrived late to training. "Minutes early/late" is a continuous variable recording the minutes early (–) or late (+) job trainees arrived at training. Time use in Columns 4 and 5 comes from survey data and is the average hours reported by respondents across the three observations for each activity. The effort index is a summary measure of the effort indicators. It is constructed as the average of the normalized values of "Minutes early/late," "Hours studying training materials," and "Hours watching television/listening to the radio." Treatment status was randomly allocated and stratified by ability quintile and prior work experience with the recruiter. The stratification cell fixed effects include a set of dummies for each stratification cell. The set of additional covariates include a dummy variable for whether the individual has worked before, marital status, age, and the individuals' standardized ability score. For covariates with missing observations, the variable is assigned the mean value of the variable, and an indicator variable is included for whether that particular variable is missing. Robust standard errors are presented.

Time Use

A second dimension of effort is self-reported time use. As part of the daily follow-up questionnaires, individuals reported activities in a time use module. I constructed a measure of the average number of hours spent studying training materials and leisure time spent listening to the radio or watching television per training day.

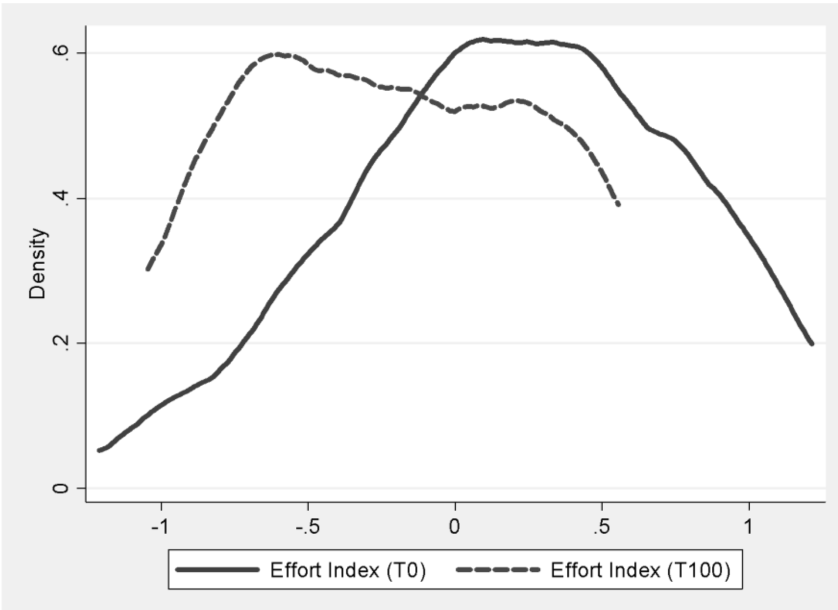
Table 6.3, Column 4, presents the mean number of hours spent studying the training materials for each treatment group. Those with the guaranteed job option reported spending the least amount of time studying the training materials—as much as 25 fewer minutes per day than those who received no chance of alternative employment. Moreover, Table 6.3, Column 5, indicates that individuals in the 100 percent chance treatment group spent 53 more minutes watching television or listening to the radio than those assigned no outside option. These results suggest that individuals substituted time spent studying the training materials for leisure time.

As previously mentioned, the survey data do exhibit differential nonresponse by treatment status. I conducted a host of specification checks to control for the differential nonresponse, including weighting, applying minimum and maximum bounds, and applying Lee bounds. The results are generally robust for all three methods. (See Appendix C for details.)

Effort Index

Due to multiple effort indicators, I created an effort summary measure, analogous to the performance index. The effort index used the mean of the normalized value of the average minutes arrived early or late, the number of hours spent studying the training materials, and the number of hours spent watching television or listening to the radio. The results are presented in Table 6.3, Column 6. Those assigned no outside option exerted 0.587 standard deviations more effort compared with those assigned a guaranteed outside option. Figure 6.3 presents the distribution of this index for the no outside option (*T*0) and the employment guarantee (*T*100) groups. The figure shows that the effort distribution for those guaranteed outside employment is shifted to the left. The *p*-value associated with a Mann-Whitney equality of distributions test is 0.0011.

Figure 6.3 Effort index distribution



Source: Author's calculations.

In sum, the results show that individuals assigned high outside options exert lower levels of effort, whereas those assigned poor outside options exhibit higher effort. Therefore, the poorer performance among those with poor outside options is not driven by lower effort.

Potential Mechanisms and Discussion

Contrary to the predictions of a standard economic model, in which performance is a decreasing function of the value of the outside option, the findings indicate that performance is highest among those with the best outside option (those facing no employment risk), and it is not driven by effort. However, effort does decrease as expected, and performance improves consistent with the *choking under pressure* framework. The remaining challenge is to determine whether improved performance operates through the stress effect or whether some other mechanism is responsible for the change in performance.

The impact of reduced risk on effort helps, in part, to distinguish between different possible mechanisms, though it does not provide a conclusive test. Although my results are consistent with the theory that a reduction in stress leads to increased performance, a number of alternative behavioral theories could also explain the performance results. I considered gift exchange, stereotype threat, and the nutritional efficiency wage hypothesis.[31]

[31] Ideally, to determine whether the stress effect really is the driver of the observed performance effects, biomarker data collection (for example, cortisol) would have been optimal. Unfortunately, due to budgetary and logistical restrictions, this was not possible. However, in a pilot conducted in a similar setting, I collected four heart rate readings at the same time of day on four different days. Two of these were taken on training days that occurred before job probabilities had been announced, and two occurred after the announcement. I compared the average of the post-announcement heart rates to the average of the preannouncement heart rates for individuals assigned job guarantees compared with those given low chances of outside job options. Individuals assigned a guaranteed outside option experienced a 6.4-point greater decline in their heart rate (s.e. = 3.25)

Gift Exchange

The gift exchange hypothesis, presented in seminal work by Akerlof (1982) and built upon by Akerlof and Yellen (1988, 1990), relies on the key assumption that there is a positive relationship between wages and worker effort. This relationship explains higher-than-market-rate clearing prices, wherein workers reciprocate higher wages with more effort. Substantive lab experimental evidence supports the gift exchange model. Fehr, Kirchsteiger, and Riedl (1993) provided some of the first evidence, and Fehr and Gaechter (2000) provided a survey of the reciprocity literature more generally.

In this setting, job trainees may feel rewarded by the recruiter when allocated a high outside option and may reciprocate by exerting more effort, which, in turn, increases performance. However, although a gift exchange hypothesis yields similar performance predictions, for gift exchange to be the key driving mechanism, effort indicators should increase as outside options increase. But the results are the opposite for effort indicators (Table 6.3). Therefore, the higher observed performance among those assigned high outside options cannot be explained by the gift exchange mechanism.

Nutritional Efficiency Wage Hypothesis

Another framework that would yield similar predictions for the performance results is the extensively researched nutritional efficiency wage hypothesis (Leibenstein 1957, 1958; Stiglitz 1976; Deolalikar 1988). Improved nutritional intake improves both physical and mental well-being, which translates into increased productivity.

In the experimental setting, individuals who were guaranteed an alternative job may have been able to borrow against this guarantee and improve their short-run nutritional intake. A comprehensive caloric intake daily roster was not administered; however, I did collect information on daily expenditures on food, including expenses on food consumed at home and away from home. (See Table 6.4, Columns 1 and 2.) Food expenditures are relatively consistent across the treatment groups, and there were no statistically significant differences in expenditures between the 0 percent and 100 percent treatment groups. Accounting for the differential survey nonresponse using weighted and conservative bounds, I still cannot reject that the 0 percent and 100 percent treatment groups spent equal amounts on food.[32] (Refer to Appendix C for detailed results.)

---

compared with those assigned a 1 percent outside option (in the pilot, the "no outside option" did not exist). This provides further support of a reduction in stress driven by the assigned job guarantee. Although biomarker data collection would yield insight into the presence of a biological stress response, it would not address outstanding questions regarding how stress acts to inhibit performance. Psychological research has identified many factors that contribute to suboptimal performance, including the mere presence of an audience, public speaking, and public announcements about performance (Baumeister and Showers 1986; Beilock 2010). The psychological literature moves beyond identifying factors that affect performance in this way to examine precise mechanisms related to how effects on working memory lead to suboptimal performance. In my setting, it could be that job seekers assigned the low outside option overthink their performance such that paying too much attention actually becomes counterproductive (Beilock et al. 2002). Another possibility is that individuals assigned no outside option experience an increase in distracting thoughts and worries related to their likely continued unemployment, which prevents them from focusing on the important information (Hayes, Hirsch, and Matthews 2008). This study is not designed to determine the precise mechanism through which the stress may operate to impair performance.

[32] There is one exception. However, the exception suggests that those assigned no outside option spend more on food as compared with those assigned the employment guarantee. This exception also works against the possibility that the observed results are driven by increased nutritional intake.

Table 6.4 Alternative explanations

Average	Food expenditures (in MKW) (1)	Eat-out expenditures (in MKW) (2)	Perceived chance of employment with recruiter (3)
0% job guarantee	349.479	124.151	73.058
	[77.118]	[16.339]	[3.557]
1% job guarantee	425.084	165.495	73.538
	[98.487]	[15.067]	[2.996]
5% job guarantee	372.697	154.952	76.109
	[92.836]	[21.179]	[3.170]
50% job guarantee	439.111	147.49	72.706
	[97.689]	[20.097]	[2.343]
75% job guarantee	335.364	183.878	83.596
	[74.342]	[27.507]	[3.376]
100% job guarantee	328.482	123.887	77.596
	[79.742]	[23.159]	[3.553]
Observations	256	256	256
R-squared	0.36	0.6	0.94
p-value of F-test:			
0% and 100%	0.797	0.543	0.363

Source: Author's calculations.

Notes: This table presents the treatment group means for each outcome. Food expenditures (in Malawian kwachas [MKW]) is the average amount spent on food reported by the respondent across the three training days. "Eat-out expenditures (in MKW)" is similar, except it measures food expenditures for food consumed away from the home. "Perceived chance of employment with recruiter" is constructed using the following question: "What percentage chance do you think you have of getting one of the available positions for the RECRUITER'S PROJECT?" with the following options: No chance of getting a job; Less than 25 percent; Between 25 and 50 percent; 50 percent; Between 50 and 75 percent; Between 75 and 99 percent; and Certain about employment with recruiter. To create a measure of the likelihood of employment, I assigned the midpoint to categories that are in brackets and creating a continuous variable. Treatment status was randomly allocated and stratified by ability quintile and prior work experience with the recruiter. The stratification cell fixed effects include a set of dummies for each stratification cell. The set of additional covariates include a dummy variable for whether the individual has worked before, marital status, age, and the individuals' standardized ability score. For covariates with missing observations, the variable is assigned the mean value of the variable, and an indicator variable is included for whether that particular variable is missing. Robust standard errors are presented.

Given these findings, it is unlikely that nutritional intake change is the key driver for the change in performance.

Stereotype Threat

The final potential explanation has its origins in psychology (Steele 1997). The idea is that by highlighting a specific stereotype, performance is impaired in the negatively stereotyped group. A substantive literature addresses stereotype threat and test performance (Spencer, Steele, and Quinn 1999; Maas and Cadinu 2003; Inzlicht and Ben-Zeev 2000; Steele and Aronson 1995).

In this study, job trainees may perceive their outside option as a signal of their ability. Although assignment does not reveal information regarding an individual's ability or performance relative to the other participants, job trainees may still believe that assignment is correlated with their ability.[33] In this

[33] From open-ended questions on the survey, it is evident that respondents understood that the assignment of the outside options was not correlated with ability. For example, "The probability criteria are dependent on the chance and not merit of a

case, performance of individuals could be driven by self-fulfilling perceptions of their own ability. This hypothesis predicts that individuals assigned low outside options are likely to perform worse, consistent with my findings.

In examining the extent to which job trainees updated their beliefs about getting the recruiter's job by treatment status,[34] I did not observe statistically significant differences among most groups. However, those in the 75 percent treatment group did report significantly higher expectations about their chances of getting a recruiter's job compared with all other groups, including the guaranteed outside option group (see Table 6.4, Column 3). However, the starkest key performance differential is again observed between individuals in the 0 percent treatment group and those in the 100 percent treatment group. This difference does not seem to be driven by stereotype threat, as no large differences exist in these two groups' perceptions of their chance of being hired for the recruiter's job.

Neither gift exchange nor efficiency wages nor stereotype threat can explain the full pattern of my results. Instead, it seems that the most plausible mechanism driving the performance and effort results is a framework in which the varied outside options simultaneously reduce effort and stress, enabling a higher return to effort.

Welfare Implications: Employment and Heterogeneity

To assess the welfare implications of the observed performance response to employment risk, I examined employment outcomes.[35] As discussed in Appendix A, though the performance indicators do a relatively good job of predicting performance, there is still a large unobserved component determining employment outcomes.

Figure 6.4 depicts the share of job trainees hired by the recruiter by treatment group. The recruiter offered employment to about 25 percent of trainees in the 75 and 100 percent groups.[36] Of the individuals who received no chance of an alternative job, 13 percent were hired by the recruiter. Those individuals who received a 50 percent chance of an alternative job were the least likely of all treatment groups to be hired by the recruiter—only 11 percent of these participants were hired, making them half as likely to be hired relative to those who knew they had high chances of alternative employment.

---

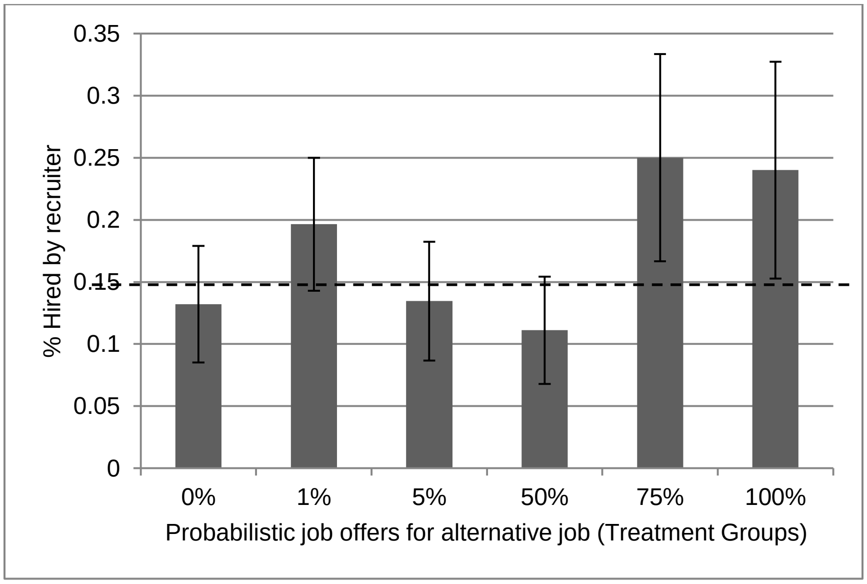
person in terms of experience and qualification"; or "It is about chances"; or "Simply it's about luck."

[34] Respondents were asked "What percentage chance do you think you have of getting one of the available positions for the recruiter's project?" with the following options: No chance of getting a job; Less than 25 percent; Between 25 and 50 percent; 50 percent; Between 50 and 75 percent; Between 75 and 99 percent; and Certain about employment with recruiter.

[35] On-the-job performance was not measured during the alternative jobs; however, the long-term impacts of being assigned an alternative job on future employment and wages are presented in Godlonton (2013). Recall that job trainees were hired as interviewers for a health survey. To measure on-the-job performance for the recruiter jobs, one can use survey data from the health survey. For example, one can measure the number of skip rules incorrectly followed and the number of inconsistencies by interviewer. For these indicators, there is little difference by treatment group; however, the small sample size prevents any substantive analysis.

[36] Note that only one participant who was offered a position by the recruiter opted not to take the job. As such, the offer of a job and the record of who got hired are approximately the same.

Figure 6.4 Fraction employed by recruiter by treatment group



Source: Author's calculations.
Note*:* The dotted line represents the fraction that would have been hired in the absence of the experiment.

Table 6.5 presents both ordinary least squares and probit model results for employment. The marginal effects reported are the partial derivatives evaluated at the mean of the covariates. Given the performance indicator results, it seems reasonable to use the 100 percent treatment group as the omitted category. Individuals in the 0, 5, and 50 percent chance of alternative work treatment groups were less likely to be employed by the recruiter by between 9 and 11 percentage points. These impacts are statistically significant and are large in magnitude, as they translate to a 50 percent lower chance of being hired than those in the 100 percent treatment group.

Table 6.5 Employment (with recruiter) by treatment group

	Ordinary least squares			Probit		
Indicator	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
0% job guarantee	0.132	0.126	0.133	-0.088	-0.095*	-0.090*
	[0.047]	[0.045]	[0.046]	[0.065]	[0.054]	[0.051]
1% job guarantee	0.196	0.195	0.197	-0.034	-0.048	-0.051
	[0.054]	[0.051]	[0.051]	[0.075]	[0.064]	[0.060]
5% job guarantee	0.135	0.139	0.136	-0.085	-0.093*	-0.093*
	[0.048]	[0.043]	[0.044]	[0.065]	[0.052]	[0.049]
50% job guarantee	0.111	0.114	0.108	-0.106*	-0.104**	-0.109**
	[0.043]	[0.043]	[0.044]	[0.061]	[0.051]	[0.046]
75% job guarantee	0.250	0.241	0.238	0.008	-0.024	-0.031
	[0.083]	[0.074]	[0.071]	[0.094]	[0.073]	[0.067]
100% job guarantee	0.240	0.250	0.256			
	[0.086]	[0.091]	[0.089]			

Table 6.5 Continued

Indicator	Ordinary least squares (1)	(2)	(3)	Probit (4)	(5)	(6)
Stratification cell fixed effects?	No	Yes	Yes	No	Yes	Yes
Includes controls?	No	No	Yes	No	No	Yes
Observations	268	268	268	268	268	268
R-squared	0.18	0.28	0.29			
p-value of F-test:						
0% and 100%	0.274	0.224	0.221			

Source: Author's calculations.

Notes: Ordinary least squares results present employment rates (with recruiter) by treatment group. Probit results present the partial derivative at the mean of the covariates of employment of the 0, 1, 5, 50, and 75 percent job probabilities treatment compared to the 100 percent treatment group, where employment risk is zero. Treatment status was randomly allocated and stratified by quintile ability and prior work experience with the recruiter. The stratification cell fixed effects include a set of dummies for each stratification cell. The set of additional covariates include a dummy variable for whether the individual has worked before, marital status, age, and the individuals' standardized ability score. For covariates with missing observations, the variable is assigned the mean value of the variable, and an indicator variable is included for whether that particular variable is missing. Robust standard errors are presented. *** indicates significance at the 1 percent level. ** indicates significance at the 5 percent level. * indicates significance at the 10 percent level.

One threat to the interpretation of the employment results is the potential for a recruiter's strategic behavior in hiring decisions in response to treatment assignments. However, the recruitment team had no knowledge of the specific alternative job probability assigned to each participant. Only participants were in a position to reveal their alternative job probability to recruitment staff. Anecdotally there are no reports of this occurring, and even if it had, one would expect that it would bias the results in favor of higher employment rates for those assigned lower alternative job probabilities. Therefore, if this behavior did occur, then my results are downwardly biased.

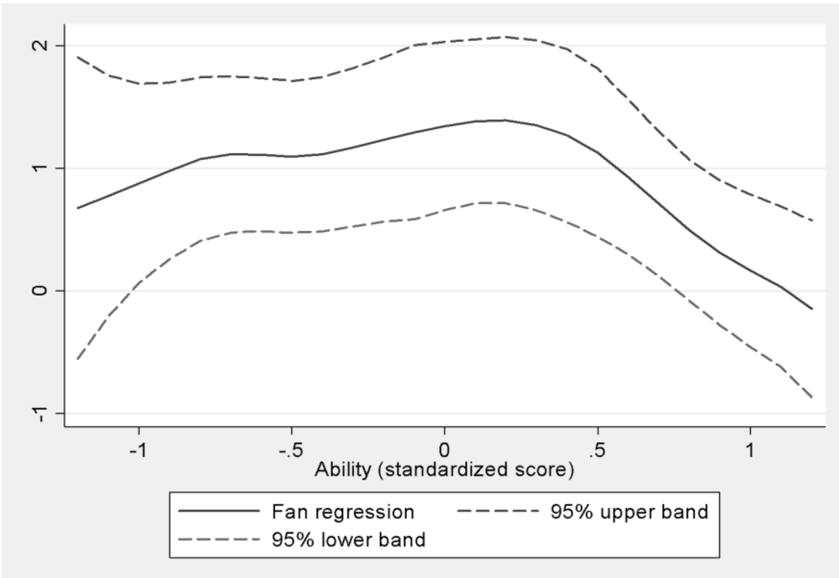
A further concern is that recruitment staff, conditional on learning a trainee's alternative job probability, may have (incorrectly) inferred something about the trainee's ability. However, the recruiter had implemented randomized controlled trials for a number of years within Malawi and was aware that treatment assignment was randomly determined. It is therefore highly unlikely that such inferences were made.

Heterogeneity

Heterogeneity in the effect of reduced risk on performance and employment may also have important policy considerations or distributional implications. Recall that treatment assignment was stratified by ability and prior work experience with this recruiter. Although other dimensions are certainly interesting to consider, the small sample, in combination with stratification only for these two observable characteristics, limits the interpretation of additional subgroup analyses. My power to detect differences by familiarity with this specific recruiter is further limited, because only 10 percent of the sample (26 individuals) had prior work experience with the recruiter. Due to this, I present and discuss the heterogeneity results by ability only.

Figure 6.5 plots the estimated difference in performance between those assigned the guarantee versus those with no outside option across the ability distribution. It is clear that the largest differences accrue to job seekers in the middle of the ability distribution.

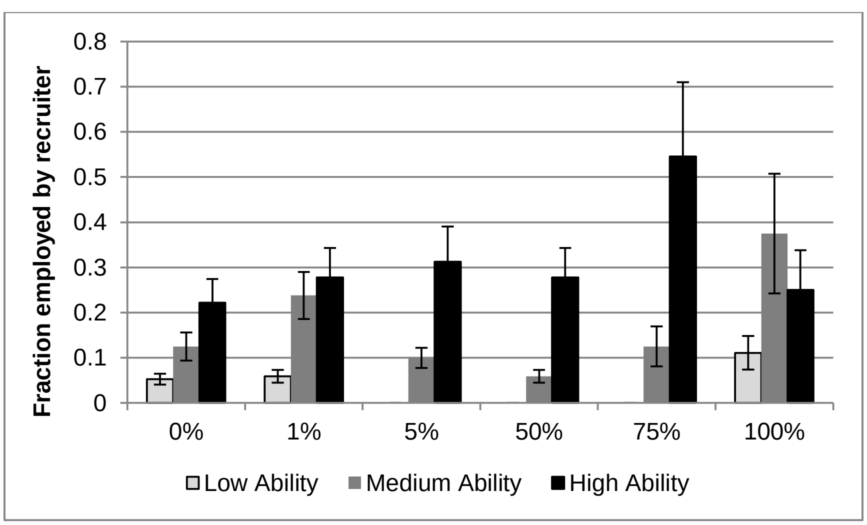
Figure 6.5 Estimated difference between guaranteed outside option and no outside option across baseline ability distribution



Source: Author's calculations.

Figure 6.6 plots the mean performance for high-, medium-, and low-ability trainees and shows large differences in employment across the different treatment groups.37 High-ability types are consistently the most likely to be hired, and their employment rates are least affected by the varied employment risk. Individuals in the middle of the baseline ability distribution benefit the most from the eliminated employment risk (the guaranteed outside option). Although reduced risk also benefits low-ability types, those differences are marginal. Figures 6.5 and 6.6 suggest that individuals in the middle of the ability distribution are the most susceptible to such risk.

Figure 6.6 Average employment by recruiter by treatment group and baseline ability



Source: Author's calculations.

[37] High-, medium-, and low-ability types are defined by cutting the sample into thirds based on the baseline ability test.

7. CONCLUSION

The results of the study indicate that job-seeker performance during recruitment is highest and effort is lowest among those assigned the highest outside employment options, while the converse is true for those assigned the worst outside options. The latter group of job trainees performs better on tests of materials taught during training and participates with higher-quality comments in the recruitment process. However, these performance improvements are not driven by changes in effort and are not linear in the probability of outside employment.

It is important to consider the limitations to the findings. First, this study was conducted using short-term job opportunities; thus, the effects of longer-term job security cannot be assessed in this context. Second, the experiment was conducted among a sample of relatively well-educated men in Lilongwe, the capital city of Malawi. This paper cannot speak to how other groups might respond. Third, in the absence of biomarker indicators, the stress channel cannot be directly tested. Fourth, due to limited power conclusions pertaining to the heterogeneity of the results, the results should be interpreted with caution.

However, these findings are consistent with prior laboratory evidence (Ariely et al. 2009), which observed lower performance under high-stakes incentives. The variation in risk in laboratory studies is artificial and over windfall income, whereas in my setting, the variation is over risk in securing real, meaningful employment equivalent to that for which subjects have chosen to apply through a competitive process. Moreover, my results speak to the growing *choking under pressure* literature that, to date, has focused only on professional sporting activities. To my knowledge, no evidence in real-world settings has illustrated the link among risk, performance, and effort. I provide the first evidence that previous findings do extend beyond the laboratory into real-world labor markets—something noted as an open question as recently as Kamenica (2012). This research has many possible extensions in real-world settings. The results examine performance during recruitment—how performance may be affected on the job is an important and interesting avenue for further research. In addition, the results were obtained in a context in which cognitive performance is important. Whether such results will be observed in manual rural labor markets is also interesting—both theoretically, as it pertains to the mechanism through which uncertainty affects performance, and practically, because of its policy relevance to the large fraction of adults in developing countries who do manual labor.

To some, these results may seem inconsistent with the unemployment insurance (UI) benefit literature, which shows that job search increases when UI benefits are about to expire (for example, Krueger and Mueller 2010). However, the findings here can be reconciled with this literature. For example, I did observe declining effort on behalf of the job trainees that is consistent with this literature. What is typically not observed in the UI literature is the effectiveness of job search in recruitment processes that are cognitively demanding. Thus, the results in this paper suggest new avenues for research in examining the impact of UI on job search effort and performance.

This paper also contributes to conceptual questions about the relationship between risk and performance. The results suggest considerable nonlinearities in this relationship that deserve further attention. Because realized outcomes are binary, studies conducted using secondary data typically do not observe the full distribution of uncertainty between an event occurring with probability zero and it occurring for certain. My results suggest that conclusions about the relationship between risk and performance are sensitive to the range of risk observed.

Moreover, the observed relationship between risk and performance has implications for how to model behavior under uncertainty. Typically, when risk is considered in theoretical models, it is modeled as a parameter of the utility function. Although my results do not reject this approach, they do imply that it is important to consider risk in production functions.

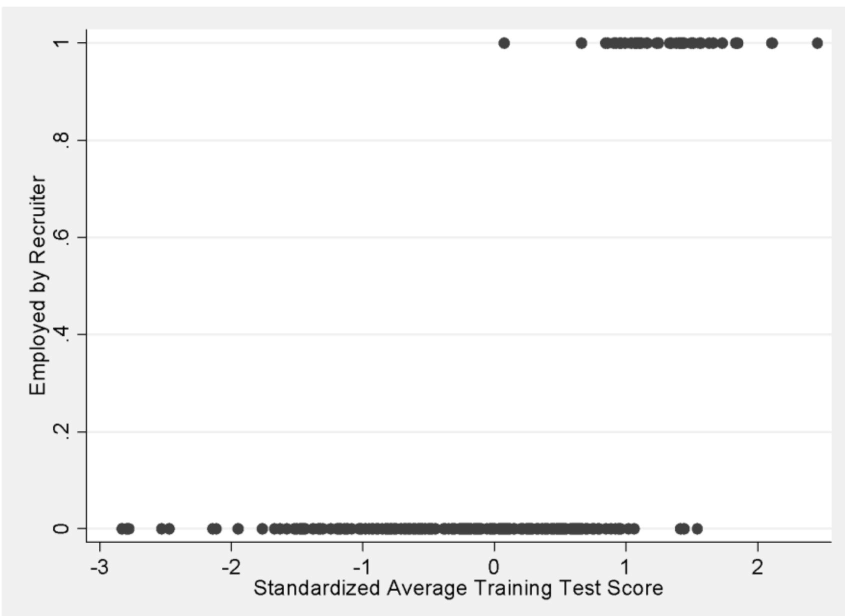
Although the reduced-form effect of risk on performance is interesting in its own right and has real-world welfare implications, I also explored the mechanisms that might drive the key results. Using rich baseline and outcome data, I combined economic and psychological insights to explore potential mechanisms through which risk and uncertainty affect behavior. The findings suggest that the relationship between risk and performance is likely driven by a stress response. However, unlike laboratory evidence, which directly measures stress using biomarkers such as cortisol (for example, Angelucci et al. 2012), I was not able to measure hormonal stress in this manner. That said, my results do not seem to be driven by models of reciprocity, the efficiency wage hypothesis, or stereotype threat. However, it is possible that some other psychological consideration that operates similarly to stress is driving the result. Moreover, I cannot determine the precise psychological mechanism through which stress operates—that is, distraction or overthinking. Future research that more precisely measures stress would be a natural extension of this work.

Finally, although this paper is most closely tied to laboratory experiments about the effect of risk and stress on performance, it also speaks to the growing literature about the effect of high-stakes testing. In this study, performance under high stakes (low probability of an outside job) is worse than performance under low stakes (high probability or guarantee of an outside job). A growing body of literature demonstrates heterogeneous responses to high-stakes versus low-stakes settings. For example, Örs, Palomino, and Peyrache (2013) showed large gender disparities in low-stakes versus high-stakes testing situations. In low-stakes testing environments, females outperform males; however, the same females perform suboptimally and, on average, worse than the males on a high-stakes entrance exam to an elite university. Suggestive evidence shows that differences by ability are important for employment outcomes; in particular, individuals in the middle of the distribution are the most affected by the reduction in risk. Understanding which types of individuals are most susceptible to risk-related performance declines could have substantive policy implications for job training or recruiting processes and deserves further attention in future research.

APPENDIX A: DETERMINANTS OF HIRING DECISIONS USING ADMINISTRATIVE DATA

The recruiter accounted for a number of factors in determining who got hired. The recruiter conducted multiple tests to ensure that trainee participants paid attention and to ensure an objective measure of assessment was available. Figure A.1 shows that the recruiter did partially rely on this in the hiring process. No participants were hired who had a standardized test score less than 0.05. All participants who had a standardized test score greater than 1.3 were hired. However, between standardized test scores of 0.05 and 1.3, the recruitment team relied on additional information about the job trainees. Performing well on the test was a necessary condition to get hired. However, it was not a sufficient condition for those participants with a standardized test score between 0.05 and 1.3.

Figure A.1 Scatter plot employed by recruiter and training test score



Source: Author's calculations.

Table A.1 presents the determinants of the recruiter's hiring decision-making process. This shows that the standardized test score is an important determinant of whether the person gets hired. A 1 standard deviation increase in the composite test score results in a 9.7 percentage point increase in the likelihood that the individual is hired. Other key indicators measured by the recruiter include punctuality, contributions, and disruptions. Figure A.1 suggests that any alternative measures of evaluating performance should be interacted with the test score.

Punctuality appears to have little impact on the hiring decision. Table A.1 also shows that for those performing well (in terms of standardized test scores), making *good* and *neutral* contributions during the training sessions increased their probability of being hired. In such a large hiring process, being noticed in a good way mattered for those participants who performed well but not exceptionally well. Lastly, Column 4 of Table A.1 includes measures for disruptions made by participants during training. This appears not to have any significant impact on the hiring decision-making process, as the magnitude of the coefficients are small and statistically insignificant.

Evidently, the most significant factor taken into account by the recruiter in hiring decisions was participants' performance on the written tests. However, conditional on performing sufficiently well, making good-quality contributions also mattered.

Table A.1 Predicting employment

Dependent Variable:	Offered a job by the recruiter				
	(1)	(2)	(3)	(4)	(5)
Age	0.107***	0.097***	0.069***	0.065***	-0.007
	[0.018]	[0.016]	[0.016]	[0.016]	[0.005]
Married	0.036	-0.008	-0.007	-0.007	0.101
	[0.071]	[0.005]	[0.005]	[0.005]	[0.070]
Ever worked	0.067	0.086	0.104	0.103	0.093
	[0.058]	[0.069]	[0.067]	[0.069]	[0.059]
Ever worked with recruiter	0.150	0.096*	0.087	0.093	0.117
	[0.094]	[0.055]	[0.058]	[0.059]	[0.078]
Ability score (standardized)	0.104***	0.139*	0.122	0.12	0.046**
	[0.024]	[0.082]	[0.078]	[0.078]	[0.023]
Test score		0.097***	0.092***	0.067***	0.063***
		[0.016]	[0.016]	[0.016]	[0.016]
Minutes late			-0.035	-0.035	0.001
			[0.043]	[0.043]	[0.001]
Minutes late x test score			0.114**	0.096*	0.001
			[0.052]	[0.051]	[0.002]
Any good contribution				-0.031	-0.031
				[0.043]	[0.043]
Any good contribution x test score				0.114**	0.098*
				[0.052]	[0.052]
Any neutral contribution				0.023	0.023
				[0.042]	[0.042]
Any neutral contribution x test score				0.078	0.068
				[0.052]	[0.050]
Any bad contribution				-0.012	0.019
				[0.041]	[0.055]
Any bad contribution x test score				0.062	-0.052
				[0.041]	[0.061]
Any disruption					-0.009
					[0.041]
Any disruption x test score					0.059
					[0.042]
Constant	0.281**	0.272**	0.269**	0.250*	0.240*
	[0.137]	[0.129]	[0.130]	[0.137]	[0.145]
Observations	268	268	268	268	268
R-squared	0.11	0.25	0.26	0.31	0.32
Average of dependent variable	0.158				

Source: Author's calculations.

Notes: The dependent variable is a binary indicator equal to 1 if the recruiter offered the job seeker a job and 0 otherwise. For covariates with missing data, the variable is assigned the mean value of the variable, and an indicator variable is included for whether that particular variable is missing. Robust standard errors. *** indicates significance at the 1% level. ** indicates significance at the 5% level. * indicates significance at the 10% level.

APPENDIX B: COVARIATE IMBALANCE SPECIFICATION CHECKS

Treatment status was randomly assigned, and covariates appear to be balanced at baseline. Given the relatively small sample, however, there may be persistent concerns regarding unobservable imbalance. This appendix presents and discusses two additional specification checks to adjust for this potential bias.

First, due to the availability of baseline survey data, it is possible to control for multiple covariates. Doing so does not substantively affect the performance (Table B.1), effort (Table B.2), or employment results (Table 6.4).

Table B.1 Performance indicators (no covariates)

		Contributions		
	Tests	**# good**	**# neutral**	**# bad**
Indicator	(1)	(2)	(3)	(4)
0% job guarantee	-0.176	0.491	0.604	0.358
	[0.147]	[0.088]	[0.122]	[0.108]
1% job guarantee	-0.015	0.804	0.589	0.196
	[0.136]	[0.134]	[0.127]	[0.086]
5% job guarantee	0.041	0.692	0.692	0.212
	[0.132]	[0.152]	[0.128]	[0.057]
50% job guarantee	0.041	0.778	0.389	0.222
	[0.124]	[0.139]	[0.093]	[0.063]
75% job guarantee	-0.039	0.750	0.500	0.071
	[0.241]	[0.150]	[0.120]	[0.049]
100% job guarantee	0.259	0.920	1.040	0.280
	[0.195]	[0.198]	[0.239]	[0.107]
Observations	258	268	268	268
R-squared	0.013	0.380	0.342	0.161
p-value of F-test:				
0% and 100%	0.076	0.048	0.105	0.607

Source: Author's calculations.

Notes*:* This table presents mean performance using an average across training days for each job trainee. I used the average of the standardized test scores, which are standardized by using the sample mean and standard deviation for the relevant test. Treatment status was randomly allocated and stratified by quintile ability and prior work experience with the recruiter. The stratification cell fixed effects include a set of dummies for each stratification cell. The set of additional covariates include a dummy variable for whether the individual has worked before, marital status, age, and the individuals' standardized ability score. For covariates with missing observations, the variable is assigned the mean value of the variable, and an indicator variable is included for whether that particular variable is missing. Robust standard errors are presented.

Table B.2 Average effort indicators (no covariates)

Indicator	Ever late (1)	Always late (2)	Minutes early or late (3)	Studied (hours) (4)	Radio/TV (hours) (5)
0% job guarantee	0.180	0.020	-24.230	1.177	1.142
	[0.055]	[0.020]	[2.228]	[0.137]	[0.121]
1% job guarantee	0.182	0.000	-21.467	1.151	1.580
	[0.053]	[0.000]	[1.794]	[0.109]	[0.132]
5% job guarantee	0.314	0.020	-19.209	0.959	1.358
	[0.066]	[0.020]	[2.310]	[0.100]	[0.154]
50% job guarantee	0.176	0.020	-21.843	1.093	1.520
	[0.054]	[0.020]	[2.099]	[0.098]	[0.138]
75% job guarantee	0.259	0.037	-19.914	1.125	1.419
	[0.085]	[0.037]	[3.023]	[0.134]	[0.162]
100% job guarantee	0.280	0.080	-19.320	0.754	2.020
	[0.091]	[0.055]	[4.354]	[0.078]	[0.247]
Observations	259	259	259	254	254
R-squared	0.238	0.043	0.647	0.699	0.676
p-value of F-test:					
0% and 100%	0.347	0.306	0.316	0.008	0.002

Source: Author's calculations.

Notes: This table presents the average effort by treatment group using both administrative data and survey data. "Ever late" is a binary indicator equal to 1 if the job trainee ever arrived late for training. "Always late" is a binary indicator equal to 1 if the job trainee arrived late for training every day. "Minutes early/late" is a continuous variable recording the average minutes early (–) or late (+) job trainees arrived across the training period. Time use in Columns 4 and 5 comes from survey data and is the average number of hours conducting each activity. Treatment status was randomly allocated and stratified by quintile ability and prior work experience with the recruiter. The stratification cell fixed effects include a set of dummies for each stratification cell. The set of additional covariates include a dummy variable for whether the individual has worked before, marital status, age, and the individuals' standardized ability score. For covariates with missing observations, the variable is assigned the mean value of the variable, and an indicator variable is included for whether that particular variable is missing. Robust standard errors are presented.

Second, I constructed a ratio that measures the extent to which selection on unobservables would need to exceed selection on observables to explain away the estimated coefficients for the performance and effort indicators (Altonji, Elder, and Taber 2005; Bellows and Miguel 2009). A larger ratio implies that the relative omitted variable bias from unobservables relative to observables is greater, and therefore estimated effects are less likely to be explained away. Table B.3 presents the ratios for each performance and effort indicator for which significant differences between those assigned no outside option and those guaranteed an outside option exist.

Table B.3 Omitted variable bias ratio

Indicator	Ratio (1)
Performance indicators:	
Tests	67.994
Engagement:	
* # of contributions	-1.504
* # good contributions	-1.672
* # neutral contributions	-1.796
Effort indicators:	
Time use:	
* Hours studied training materials	7.003
* Hours watching TV/listening to radio	9.668

Source: Author's calculations.

Notes: Following Altonji, Elder, and Taber (2005) and Bellows and Miguel (2009), I constructed a ratio that assesses the extent of omitted variable bias that would be required to explain away the results. This table presents the ratios for each performance and effort indicator for the estimated difference between those assigned no outside option and those assigned a guaranteed outside option. The ratio measures the extent to which selection on unobservables would need to exceed selection on observables to explain away the coefficient. Therefore, a larger ratio implies that the relative omitted variable bias from unobservables relative to observables is greater, and therefore estimated effects are less likely to be explained away.

For the main performance result—that is, the estimated impact on the average training test score—the required ratio is 68. This means that the selection on unobservables would have to be 68 times greater than selection based on observables controlled for. For contributions to the discussion, the ratios are negative, which suggests that the omitted variable bias results in an underestimate of the treatment effect, rather than an overestimate. Similarly, the effort indicators suggest that selection on unobservables would have to be much larger than the selection based on observables, by ratios ranging from 7 to 9.7.

Overall, the results are robust to both of these robustness checks.

APPENDIX C: MISSING ADMINISTRATIVE DATA AND DIFFERENTIAL NONRESPONSE IN SURVEY DATA

The administrative data used in most of the analysis exhibits missing values for some of the performance and engagement indicators. For example, a subset of test score data is missing (5.2 percent). These data are missing for various reasons, including misplaced test papers, illegible or incorrect employment identification numbers on submitted tests, and partial training attendance that resulted in some individuals not taking all tests.[38] Given that participation rates in training are not differential across treatment groups, the missing data should not to affect the results.

The survey data exhibit differential completion rates by treatment status. The differential nonresponse by treatment status may have implications for the results based on the survey data—primarily, the effort and mechanism results. Conducting robustness checks in this case is particularly important.

To assess the robustness of the results for outcomes deriving from both survey and administrative data, I used three approaches. First, I presented weighted results (Fitzgerald, Gottschalk, and Moffitt, 1998). To do this, I first predicted the probability of noncompletion. Using these predicted probabilities, I constructed propensity score weights for each individual. I then reran the regressions using these weights. Second, I presented conservative bounded results in which I implemented min-max bounds (Horowitz and Manski, 1998). I imputed the maximum test score for all treatment groups, except for the 100 percent treatment group, for which I imputed the minimum test score. In a second regression, I imputed the minimum test score for all treatment groups, except the 100 percent treatment group, for which I imputed the maximum test score. Lastly, I restricted the sample to the 0 and 100 percent treatment groups and estimated Lee (2009) bounds on the average treatment effect of the 100 percent group relative to the 0 percent treatment group. Overall, the results are fairly robust to the different robustness checks.

Performance indicators: Table C.1 presents the weighted and min-max bounded results. Columns 1 and 4 present the weighted regressions, Columns 2 and 5 present conservative minimum bounds, and Columns 3 and 6 present conservative maximum bounds. Table C.2 presents the Lee bounds.

Overall, the three specification checks have similar findings to the main results for test scores. Point estimates from weighted results for test scores are similar to the main results. Using the conservative min-max bounds, the differential testing performance between individuals receiving a 0 and a 100 percent chance of an alternative job is no longer statistically significant. However, the differential effect remains positive, albeit considerably smaller (Table C.1, Column 2).

[38] One potential behavioral response in this setting is that job trainees assigned poor outside options reduced their participation in training, instead opting to increase external job search effort. Recall that individuals were paid for participation during the training at a wage that is relatively competitive in this environment. Although some evidence supports lower attendance of individuals in the 0 percent treatment group relative to the 100 percent treatment group, the difference is neither large (4 percentage points) nor statistically significant (though the *p*-value is 0.147). Job search among those who attended training would have been difficult. Participants spent approximately 8 hours per day in training and reported another 1.6 hours in transit and 6.8 hours sleeping (on average). Moreover, the job-training period was conducted over a relatively short time, and delaying job search by three days would not be seen to be costly.

Table C.1 Average performance by treatment group: Weighted results and bounds

	Average test score			Good quality contributions		
		Min-max bounds			Min-max bounds	
	Weighted	0–75 = max; 100 = min	0–75 = min; 100 = max	Weighted	0–75 = max; 100 = min	0–75 = min; 100 = max
Treatment group	(1)	(2)	(3)	(4)	(5)	(6)
0% job guarantee	-0.174	-0.067	-0.288*	0.242	0.259	0.24
	[0.141]	[0.147]	[0.154]	[0.051]	[0.056]	[0.051]
1% job guarantee	-0.004	0.045	-0.1	0.371	0.414	0.359
	[0.126]	[0.129]	[0.139]	[0.061]	[0.067]	[0.060]
5% job guarantee	0.038	0.089	-0.052	0.342	0.393	0.332
	[0.120]	[0.120]	[0.139]	[0.069]	[0.079]	[0.068]
50% job guarantee	0.03	0.16	-0.049	0.351	0.38	0.346
	[0.122]	[0.133]	[0.125]	[0.063]	[0.067]	[0.062]
75% job guarantee	-0.032	0.04	-0.156	0.297	0.341	0.287
	[0.208]	[0.218]	[0.232]	[0.067]	[0.077]	[0.065]
100% job guarantee	0.261	0.252	0.271	0.391	0.382	0.393
	[0.198]	[0.202]	[0.194]	[0.089]	[0.091]	[0.088]
Observations	258	268	268	262	268	268
R-squared	0.2	0.18	0.18	0.42	0.41	0.41
Stratification cell fixed effects?	Yes	Yes	Yes	Yes	Yes	Yes
Additional controls?	Yes	Yes	Yes	Yes	Yes	Yes
p-value of F-test:						
0% and 100%	0.075	0.203	0.024	0.148	0.249	0.133

Source: Author's calculations.

Notes: This table presents mean performance using an average across training for each job trainee. I used the average of the standardized test scores, which are standardized by using the sample mean and standard deviation for the relevant test. Treatment status was randomly allocated and stratified by quintile ability and prior work experience with the recruiter. The stratification cell fixed effects include a set of dummies for each stratification cell. The set of additional covariates include a dummy variable for whether the individual has worked before, marital status, age, and the individuals' standardized ability score. For covariates with missing observations, the variable is assigned the mean value of the variable, and an indicator variable is included for whether that particular variable is missing. Robust standard errors are presented.

The Lee bounds for the average test performance results are presented in Table C.2. In this case, the analysis was restricted to only the 0 and 100 percent treatment groups and estimated a lower bound of the performance improvement of the *T*100 group (compared to the *T*0 group) at 0.346 standard deviations and the upper bound at 0.492 standard deviations.

The results for engagement indicators are similarly robust. As for performance, the weighted results for engagement indicators are similar to the main results. Using the conservative bounding approach does not affect the direction of the coefficients, though the magnitude of the differences is muted. In addition, for the number of *good* contributions, the difference between *T*0 and *T*100 is no longer statistically significant at the 10 percent level ($p = 0.249$). The Lee bounds for the key engagement variable—that is, the number of *good* contributions—comparing the 0 and 100 percent treatment groups are 0.413 and 0.509, respectively.

Table C.2 Average performance and effort by treatment group: Lee bounds

Indicator	Lower bound Coefficient (1)	Lower bound *p*-value (2)	Upper bound Coefficient (3)	Upper bound *p*-value (4)	Trimming proportion (5)
Performance indicators:					
Tests	0.346	0.154	0.492	0.054	5.66
Engagement:					
* # good contributions	0.413	0.192	0.509	0.099	1.89
* # neutral contributions	0.497	0.143	0.593	0.075	1.89
* # bad contributions	-0.155	0.429	-0.116	0.547	1.89
Effort indicators:					
Punctuality:					
* Always late	0.005	0.945	0.065	0.288	5.66
* Ever late	0.057	0.615	0.117	0.290	5.66
* Minutes early/late	1.894	0.709	6.490	0.206	5.66
Time use:					
* Hours studied training materials	-0.502	0.001	-0.363	0.021	9.43
* Hours watching TV/listening to radio	0.656	0.032	1.043	0.001	9.43

Source: Author's calculations.

Notes*:* This table presents the Lee bounds for the comparison of those assigned no outside option (*T*0) and those assigned a guaranteed outside option (*T*100). I use the average of the standardized test scores, which are standardized by using the sample mean and standard deviation for the relevant test. "Late" is a binary indicator equal to 1 if the job trainee arrived late for training on that day. "Minutes early/late" is a continuous variable recording the minutes early (–) or late (+) job trainees arrived at training. Time use in Columns 4 and 5 comes from survey data and is the number of hours conducting each activity daily.

Effort indicators: Table C.3 presents the weighted and min-max results for the effort indicators. Columns 1, 4, and, 7 present the weighted regressions; Columns 2, 5, and 8 present minimum bounds; and Columns 3, 6, and 9 present maximum bounds. Lee bounds are presented in Table C.2. In all cases, including the time use indicators, the results discussed are robust to these rigorous specification checks. Even using the most conservative bounds for the time use results, the difference between the amount of time spent by *T*0 and *T*100 remains quantitatively large and statistically significant at the 5 percent level. Those assigned the job guarantee (*T*100) spent 19 fewer minutes studying the training materials and 41 more minutes watching television or listening to the radio relative to those assigned no outside option. Particularly important in the case of the time use data, the Lee bounds show that those assigned a guaranteed outside option studied the training materials less and watched more television. The upper and lower bounds are statistically significant and consistently show large differences between those assigned the job guarantee (*T*100) and those assigned no additional probability of outside employment (*T*0).

In sum, the results are generally robust to a number of alternative specifications and bounding exercises. The findings consistently show that performance is highest among those with guaranteed outside options and lowest among those assigned no outside option. However, effort is highest among those assigned no outside option and lowest among those with the guaranteed outside options.

Table C.3 Average effort indicators: Weighted results and bounds

	Punctuality			Hours studied training materials			Hours watching television or listening to the radio		
		Min-max bounds			Min-max bounds			Min-max bounds	
Indicator	**Weighted (1)**	**0–75 = max; 100 = min (2)**	**0–75 = min; 100 = max (3)**	**Weighted (4)**	**0–75 = max; 100 = min (5)**	**0–75 = min; 100 = max (6)**	**Weighted (7)**	**0–75 = max; 100 = min (8)**	**0–75 = min; 100 = max (9)**
0% job guarantee	0.088	0.139	0.083	1.17	1.41	1.069	1.156	1.334	1.044
	[0.030]	[0.040]	[0.028]	[0.131]	[0.159]	[0.127]	[0.124]	[0.139]	[0.123]
1% job guarantee	0.081	0.091	0.079	1.158	1.268	1.127	1.593	1.681	1.536
	[0.025]	[0.027]	[0.024]	[0.111]	[0.134]	[0.110]	[0.134]	[0.146]	[0.134]
5% job guarantee	0.152	0.173	0.146	0.946	1.091	0.889	1.341	1.557	1.256
	[0.036]	[0.037]	[0.035]	[0.105]	[0.122]	[0.102]	[0.166]	[0.191]	[0.162]
50% job guarantee	0.079	0.125	0.073	1.087	1.222	1.03	1.505	1.658	1.429
	[0.030]	[0.040]	[0.029]	[0.100]	[0.121]	[0.099]	[0.133]	[0.150]	[0.136]
75% job guarantee	0.129	0.157	0.124	1.16	1.212	1.138	1.428	1.487	1.374
	[0.051]	[0.058]	[0.049]	[0.147]	[0.152]	[0.145]	[0.167]	[0.177]	[0.168]
100% job guarantee	0.186	0.182	0.186	0.742	0.73	0.747	2.029	2.014	2.037
	[0.066]	[0.067]	[0.066]	[0.078]	[0.090]	[0.074]	[0.247]	[0.246]	[0.250]
Observations	259	268	268	254	268	268	254	268	268
R-squared	0.24	0.27	0.24	0.69	0.65	0.66	0.71	0.69	0.68
p-values of F-test:									
0% and 100%	0.186	0.578	0.158	0.005	0.000	0.028	0.002	0.017	0.001

Source: Author's calculations.

Notes: This table presents the average daily effort by treatment group using both administrative data and survey data. "Late" is a binary indicator equal to 1 if the job trainee arrived late for training on that day. "Minutes early/late" is a continuous variable recording the minutes early (–) or late (+) job trainees arrived at training. Time use in Columns 4 and 5 comes from survey data and is the number of hours conducting each activity daily. The effort index is a summary measure of the effort indicators. It is constructed as the average of the normalized values of "Minutes early/late," "Hours studying training materials," and "Hours watching television/listening to the radio.". Treatment status was randomly allocated and stratified by quintile ability and prior work experience with the recruiter. The stratification cell fixed effects include a set of dummies for each stratification cell. The set of additional covariates include a dummy variable for whether the individual has worked before, marital status, age, and the individuals' standardized ability score. For covariates with missing observations, the variable is assigned the mean value of the variable, and an indicator variable is included for whether that particular variable is missing. Robust standard errors are presented.

REFERENCES

Akerlof, G. A. 1982. “Labor Contracts as Partial Gift Exchange.” *Quarterly Journal of Economics* 97 (4): 543–569.

Akerlof, G. A., and J. Yellen. 1988. “Fairness and Unemployment.” *American Economic Review, Papers, and Proceedings* 83: 44–49.

———. 1990. “The Fair Wage-Effort Hypothesis and Unemployment.” *Quarterly Journal of Economics* 105: 255–283.

Altonji, J. G., T. E. Elder, and C. R. Taber. 2005. “Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools.” *Journal of Political Economy* 113 (1): 151–184.

Angelucci, M., A. Breman, K. Cordova, and L. Ordonez. 2012. “*Mental Overload and Entrepreneurship. Ann Arbor:”* Unpublished, University of Michigan.

Apesteguia, J., and I. Palacios-Huerta. 2010. “Psychological Pressure in Competitive Environment: Evidence from a Randomized Natural Experiment.” *American Economic Review* 100 (5): 2548–2564.

Ariely, D., U. Gneezy, G. Loewenstein, and N. Mazar. 2009. “Large Stakes and Big Mistakes.” *Review of Economic Studies* 76: 451–469.

Band, S.R. and Manuele C.A. 1987. “Stress and Police Officer Performance: An Examination of Effective Coping Behavior.” *Police Studies* Vol. 10 (3): 122–131.

Baumeister, R. F. 1984. “Choking Under Pressure: Self-consciousness and Paradoxical Effects of Incentives on Skillful Performance.” *Journal of Personality and Social Psychology* 46 (3): 610–620.

Baumeister, R. F., and C. J. Showers. 1986. “A Review of Paradoxical Performance Effects: Choking Under Pressure in Sports and Mental Tests.” *European Journal of Social Psychology* XVI: 361–383.

Beaman, L., and J. Magruder. 2012. “Who Gets the Job Referral? Evidence from a Social Networks Experiment.” *American Economic Review* 102 (7): 3574–3593.

Beilock, S. 2010. *Choke: What the Secrets of the Brain Reveal about Getting It Right When You Have To.* New York: Free Press.

Beilock, S. L., T. H. Carr, C. MacMahon, and J. L. Starkes. 2002. “When Paying Attention Becomes Counterproductive: Impact of Divided versus Skill-focused Attention on Novice and Experienced Performance of Sensorimotor Skills.” *Journal of Experimental Psychology: Applied* 8: 6–16.

Bellows, J., and E. Miguel. 2009. “War and Local Collective Action in Sierra Leone.” *Journal of Public Economics* 93: 1144–1157.

Burgard, S., J. E. Brand, and J. S. House. 2009. “Perceived Job Insecurity and Worker Health in the United States.” *Social Science and Medicine* 69: 777–785.

Burns, J., S. Godlonton, and M. Keswell. 2010. “Social Networks, Employment and Worker Discouragement: Evidence from South Africa.” *Labour Economics* 17: 336–344.

Carroll, C. D. 1994. “How Does Future Income Affect Current Consumption?” *Quarterly Journal of Economics* 109: 111–148.

Chan, S., and A. H. Stevens. 2001. “Job Loss and Employment Patterns of Older Workers.” *Journal of Labor Economics* 19 (2): 484–521.

Chinkhumba, J., S. Godlonton, and R. Thornton. 2012. “Demand for Medical Male Circumcision.” Dataset.

Couch, K. A. 2001. “Earnings Losses and Hours Reductions of Displaced Workers in Germany.” *Industrial and Labor Relations Review* 54 (3): 559–572.

Dee, T., and W. R. West. 2011. “The Non-cognitive Returns to Class Size.” *Educational Evaluation and Policy Analysis* 30 (1): 23–46.

Deolalikar, A. B. 1988. "Nutrition and Labor Productivity in Agriculture: Estimates for Rural South India." *Review of Economics and Statistics* 70 (3): 406–413.

de Witte, H. 1999. "Job Insecurity and Psychological Well-being. Review of the Literature and Exploration of Some Unresolved Issues." *European Journal of Work and Organizational Psychology* 8 (2): 155–177.

———. 2005. "Job Insecurity: Review of the International Literature on Definitions, Prevalence, Antecedents, and Consequences." *SA Journal of Industrial Psychology* 31 (4): 1–6.

Dohmen, T. J. 2008. "Do Professionals Choke Under Pressure?" *Journal of Economic Behavior and Organization*. doi:10.1016/j.jebo.2005.12.004.

Dynan, K. E. 1993. "How Prudent Are Consumers?" *Journal of Political Economy* 101: 1104–1113.

Fafchamps, M. 2010. "Vulnerability, Risk Management and Agricultural Development." *African Journal of Agricultural Economics* 5 (1): 243–260.

Feather, N. T. 1990. "The Psychological Impact of Unemployment." *Springer Series in Social Psychology.* New York: Springer-Verlag.

Fehr, E., and S. Gaechter. 2000. "Fairness and Retaliation: The Economics of Reciprocity." *Journal of Economic Perspectives* 14: 159–181.

Fehr, E., G. Kirchsteiger, and A. Riedl. 1993. "Does Fairness Prevent Market Clearing? An Experimental Investigation." *Quarterly Journal of Economics* 108: 437–460.

Fitzgerald, J., P. Gottschalk, and R. Moffitt. 1998. "An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics." *Journal of Human Resources* 33 (2): 251–299.

Fredricks, J.A., Blumenfeld, P.C., and Paris, A.H. 2004. School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74 (1): 59-109.

Godlonton, S. 2013. "Returns to Work Experience: Evidence from Malawi." Unpublished, International Food Policy Research Institute, Washington, DC.

Green, D. A., and W. G. Riddell. 1993. "The Economic Effects of Unemployment Insurance in Canada: An Empirical Analysis of UI Disentitlement. Part 2: U.S. and Canadian Income Maintenance Programs." *Journal of Labor Economics* 11 (1): S96–S147.

Gregg, P., and E. Tominey. 2005. "The Wage Scar from Male Youth Unemployment." *Labour Economics* 12 (4): 487–509.

Gruber, J. 1997. "The Consumption Smoothing Benefits of Unemployment Insurance." *American Economic Review* 87 (1): 192–205.

Hayes, S., C. Hirsch, and A. Matthews. 2008. "Restriction of Working Memory Capacity during Worry." *Journal of Abnormal Psychology* 117 (3): 712–717.

Heyneman, S. P., D. T. Jamison, B. Searle, K. Galda, et al. 1981. "Improving Elementary Mathematics Education in Nicaragua: An Experimental Study of the Impact of Textbooks and Radio on Achievement." *Journal of Education Psychology* 73 (4): 556–567.

Horowitz, J. L., and C. F. Manski. 1998. "Censoring of Outcomes and Regressors Due to Survey Nonresponse: Identification and Estimation Using Weights and Imputations." *Journal of Econometrics* 84: 37–58.

HDR (Human Development Report). 2009. *Overcoming Barriers: Human Mobility and Development.* New York: United Nations Development Programme.

ILO (International Labor Organization). 2012. *Global Employment Trends 2012: Preventing a Deeper Job Crisis.* Geneva: ILO.

Inzlicht, M., and T. Ben-Zeev. 2000. "A Threatening Intellectual Environment: Why Females Are Susceptible to Experiencing Problem-solving Deficits in the Presence of Males." *Psychological Science* 11 (5): 365–371.

Jamal, M. 1984. "Job Stress and Job Performance Controversy: An Empirical Assessment." *Organizational Behavior and Human Performance* 33 (1): 1–21.

Kamenica, E. 2012. “Behavioral Economics and Psychology of Incentives.” *Annual Review of Economics* 4: 427–452.

Kling, J. R., J. B. Liebman, and L. F. Katz. 2007. “Experimental Analysis of Neighborhood Effects.” *Econometrica* 75 (1): 83–119.

Kocher, M., Lenz, V., Sutter, M. 2012. “Psychological Pressure in Competitive Environments: New Evidence from Randomized Natural Experiments. *Management Science* 58 (8): 1585-1591.

Kremer, M., and A. Holla. 2009. “Improving Education in the Developing World: What Have We Learned from Randomized Evaluations?” *Annual Review of Economics*. 1: 513–542.

Krueger, A. B., and A. Mueller. 2010. “Job Search and Unemployment Insurance: New Evidence from Time Use Data.” *Journal of Public Economics* 94 (3): 298–307.

Lee, D. 2009. “Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects.” *Review of Economic Studies* 76 (3): 1071–1102.

Leibenstein, H. 1957. “The Theory of Underemployment in Backward Economies.” *Journal of Political Economy* 65 (2): 91–103.

———. 1958. “Underemployment in Backward Economies: Some Additional Notes.” *Journal of Political Economy* 66 (3): 256–258.

Lusardi, A. 1998. “On the Importance of Precautionary Saving.” *American Economic Review: Papers and Proceedings* 88 (2): 448–453.

Maas, A., and M. Cadinu. 2003. “Stereotype Threat: When Minority Members Underperform.” *European Review of Social Psychology* 14: 243–275.

Magruder, J. 2012. “High Unemployment Yet Few Small Firms: The Role of Centralized Bargaining in South Africa.” *American Economic Journal: Applied Economics* 4 (3): 138–166.

Malawi, NSO (National Statistics Office). 2011. *Third Integrated Household Survey Data (IHS3).* Zomba: NSO; Washington, DC: World Bank.

Motowidlo, S. J., J. S, Packard, and M. R. Manning. 1986. “Occupational Stress: Its Causes and Consequences for Job Performance.” *Journal of Applied Psychology* 71 (4): 618–629. doi: 10.1037/0021-9010.71.4.618.

Örs, E., F. Palomino, and E. Peyrache. 2013. “Performance Gender-Gap: Does Competition Matter?” *Journal of Labor Economics* 31(3): 443–499.

Paserman, D. 2010. “Gender Differences in Performance in Competitive Environments: Evidence from Professional Tennis Players” Unpublished, Boston University.

Prendergast, C. 1999. “The Provision of Incentives in Firms.” *Journal of Economics Literature* 37 (1): 7–63.

Ruhm, C. J. 1991. “Are Workers Permanently Scarred by Job Displacements?” *American Economic Review* 81: 319–324.

———. 1994. “Advance Notice, Job Search, and Postdisplacement Earnings.” *Journal of Labor Economics* 12 (1): 1–28.

Sandmo, A. 1971. “On the Theory of the Competitive Firm under Price Uncertainty.” *American Economic Review* 61 (1): 65–73.

Sarason, G., and S. B. Mandler. 1952. “A Study of Anxiety and Learning.” *Journal of Abnormal and Social Psychology* 47: 166–173.

Segal, C. 2008. “Classroom Behavior.” *Journal of Human Resources* 43 (4): 783–814.

———. 2012. “Working When No One Is Watching: Motivation, Test Scores, and Economic Success.” *Management Science* 58 (8): 1438–1457.

———. 2013. “Misbehavior, Education, and Labor Market Outcomes.” *Journal of the European Economic Association* 11 (4): 743–779.

Spencer, S., C. M. Steele, and D. Quinn. 1999. “Under Stereotype Threat and Women’s Math Performance.” *Journal of Experimental Social Psychology* 35: 4–28.

Steele, C. M. 1997. “A Threat in the Air: How Stereotypes Shape Intellectual Identity and Performance.” *American Psychologist* 6: 613–629.

Steele, C. M., and J. Aronson. 1995. “Stereotype Threat and the Intellectual Test Performance of African Americans.” *Journal of Personality and Social Psychology* 29: 797–811.

Stevens, A. H. 1997. “Persistent Effects of Job Displacement: The Importance of Multiple Job Losses.” *Journal of Labor Economics* 15: 165–188.

Stiglitz, J. E. 1976. “The Efficiency Wage Hypothesis, Surplus Labour and the Distribution of Labour in LDCs.” *Oxford Economic Papers* 28 (2): 185–207.

Sullivan, S. E., and R. S. Bhagat. 1992. “Organizational Stress, Job Satisfaction and Job Performance: Where Do We Go from Here?” *Journal of Management* 18 (2): 353–374.

Topel, R. 1990. “Specific Capital and Unemployment: Measuring the Costs and Consequences of Job Loss.” *Carnegie-Rochester Conference Series on Public Policy* 22: 181–214.

Tryon, G. S. 1980. “The Measurement and Treatment of Test Anxiety.” *Review of Educational Research* 50 (2): 343–372.

WDR (World Development Report). 2013. *Jobs Report*. Washington, DC: World Bank.

Yerkes, R. M., and J. D. Dodson. 1908. “The Relationship of Strength of Stimulus to Rapidity of Habit Formation.” *Journal of Comparative Neurology of Psychology*: 459–482.

RECENT IFPRI DISCUSSION PAPERS

For earlier discussion papers, please go to www.ifpri.org/pubs/pubs.htm#dp.
All discussion papers can be downloaded free of charge.

1331. *Food prices and poverty reduction in the long run*. Derek Headey, 2014.

1330. *Impact of Ghana's Agricultural Mechanization Services Center Program*. Samuel Benin, 2014.

1329. Aid effectiveness: How is the L'Aquila Food Security Initiative doing? Samuel Benin, 2014.

1328. *Improved dairy cows in Uganda: Pathways to poverty alleviation and improved child nutrition*. Nassul Kabunga, 2014.

1327. *A comparative analysis of global cropping systems models and maps*. Weston Anderson, Liangzhi You, Stanley Wood, Ulrike Wood-Sichra, and Wenbin Wu, 2014.

1326. *Population pressure and livelihood dynamics: Panel evidence from Bangladesh*. Shahidur Rashid, Nigussie Tefera, Solomon Lemma, and Mohammed Yunus, 2014.

1325. *Changing sources of growth in Indian Agriculture: Implications for regional priorities for accelerating agricultural growth.* Pratap S. Birthal, Pramod K. Joshi, Digvijay S. Negi, and Shaily Agarwal, 2014.

1324. Risk and Ambiguity Preferences and the Adoption of New Agricultural Technologies: Evidence from Field Experiments in Rural India. Patrick S. Ward and Vartika Singh, 2014.

1323. Land, Assets, and Livelihoods: Gendered Analysis of Evidence from Odisha State in India. Vivien Savath, Diana Fletschner, Amber Peterman, and Florence Santos, 2014.

1322. *Exploring local perceptions of climate change impact and adaptation in rural Bangladesh.* Peter Davis and Snigdha Ali, 2014.

1321. *Closing the gender asset gap: Learning from value chain development in Africa and Asia*. Agnes R. Quisumbing, Deborah Rubin, Cristina Manfre, Elizabeth Waithanji, Mara van den Bold, Deanna Olney, and Ruth Meinzen-Dick, 2014.

1320. *The political economy of Zambia's recovery: Structural change without transformation?* Danielle Resnick and James Thurlow, 2014.

1319. *The impact of Bolsa Família on schooling: Girls' advantage increases and older children gain*. Alan de Brauw, Daniel O. Gilligan, John Hoddinott, and Shalini Roy, 2014.

1318. *Exploring strategic priorities for regional agricultural research and development investments in Southern Africa.* Michael Johnson, Samuel Benin, Liangzhi You, Xinshen Diao, Pius Chilonda, and Adam Kennedy, 2014.

1317. *How do intrahousehold dynamics change when assets are transferred to women?: Evidence from BRAC's Challenging the Frontiers of Poverty Reduction—Targeting the Ultra Poor Program in Bangladesh*. Narayan Das, Rabeya Yasmin, Jinnat Ara, Md. Kamruzzaman, Peter Davis, Julia A. Behrman, Shalini Roy, and Agnes R. Quisumbing, 2013.

1316. *Is there an app for that?: The impact of community knowledge workers in Uganda. Bjorn Van Campenhout, 2013.*

1315. *Can integrated agriculture-nutrition programs change gender norms on land and asset ownership: Evidence from Burkina Faso.* Mara van den Bold, Abdoulaye Pedehombga, Marcellin Ouedraogo, Agnes R. Quisumbing, and Deanna Olney, 2013.

1314. *Moving in the right direction?:Maize productivity and fertilizer use and use intensity in Ghana*. Antony Chapoto and Catherine Ragasa, 2013.

1313. *Women's empowerment in agriculture, production diversity, and nutrition: Evidence from Nepal.* Hazel Jean L. Malapit, Suneetha Kadiyala, Agnes R. Quisumbing, Kenda Cunningham, and Parul Tyagi, 2013.

1312. *Don't tell on me: Experimental evidence of asymmetric information in transnational households*. Kate Ambler, 2013.

1311. *Can dairy value-chain projects change gender norms in rural Bangladesh?: Impacts on assets, gender norms, and time use.* Agnes R. Quisumbing, Shalini D. Roy, Jemimah Njuki, Kakuly Tanvin, and Elizabeth Waithanji, 2013.

1310. *Can government-allocated land contribute to food security?: Intrahousehold analysis of West Bengal's microplot allocation program.* Florence Santos, Diana Fletschner, Vivien Savath, and Amber Peterman, 2013.

1309. *The National Health Insurance Scheme in Ghana: Implementation challenges and proposed solutions.* Gissele Gajate-Garrido and Rebecca Owusua, 2013.

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

www.ifpri.org

IFPRI HEADQUARTERS

2033 K Street, NW
Washington, DC 20006-1002 USA
Tel.: +1-202-862-5600
Fax: +1-202-467-4439
Email: ifpri@cgiar.org