Hiring via Internship and Motivating via Salary:

Measuring the Selection and Incentive Effects of Career and Financial Incentives[*]

Bryant Hyuncheol Kim[†], Seonghoon Kim[‡], and Thomas T. Kim[§]

30 May 2016

Preliminary and incomplete (not for citation or distribution)

Abstract

We design and implement a two-stage randomized field experiment to disentangle the selection and incentive effects of career and financial incentives where our collaborating non-governmental organization (NGO) recruits and trains enumerators for a population census of a rural catchment district in Malawi. Career incentives in our setting consist of a future job prospect and a recommendation letter, which are typical components of an internship. The financial incentive we study consists of a fixed wage, which is not dependent on job performance. We find that those selected through the career incentive of an internship perform significantly better than those hired through the financial incentive channel. In addition, we find that an additional financial incentive increases labor productivity of workers recruited via the internship channel.

Keywords: Career Incentive, Financial Incentive, Internship, Worker Selection, Incentive Effect, Labor Productivity,

JEL Classification: J33, O15, M52

1. Introduction

Work incentives are essential to recruit productive workers and motivate them to become more productive. Even though understanding the nature of incentives is crucial in human resource management, it is rare to find an empirical study that estimates the causal effect of work incentives at the recruitment stage on labor productivity (Oyer and Schaefer, 2011). The majority of empirical evidence focuses on the role of incentives in improving the productivity of existing workers following the seminal work of Akerlof (1982), the so-called gift exchange theory.[1] However, there is a growing literature on the role of incentives on worker selection at the recruitment stage. The studies in the literature mostly show that types and levels of work incentives matter in the sense that they are an effective means to hire the right kind of workers, as potential job seekers sort into a suitable job in which they can maximize their utility (Ashraf et al., 2015; Dal Bo et al., 2103; Deserranno, 2014; Gagliarducci and Nannicini, 2013; Goldberg, 2013).

In this paper, we study the role of an internship as a career incentive and a wage as a financial incentive in determining labor productivity.[2] The career incentive we study is a future job prospect and a recommendation letter, which are typical components of an internship benefit. The financial incentive in our setting is a fixed salary, which is not dependent on job performance. Specifically, we first test whether financial incentives and career incentives affect self-selection into the job for individuals in the early stage of their careers (*selection effect*). In addition, we test whether financial and career incentives motivate existing workers to become more productive (*incentive effect*). We collaborate with Africa Future Foundation (AFF), a non-governmental organization (NGO), which recruits and trains a large number of enumerators for a population census in Chimutu, a rural district in Malawi. The census aims to collect demographic and socio-economic information of the household in the survey area.

The main challenge of understanding the impacts of work incentives on labor productivity is to isolate the productivity-enhancing effect of work incentives (*incentive effect*)

---

[1] In gift exchange theory, the provision of incentives leads to an increase in labor productivity as workers exert more effort in return for a "gift" (work incentive) provided by employers.

[2] An internship is a temporary position that can be paid or unpaid and is distinguished from a short-term job in that it emphasizes on-the-job training for students or entry-level workers. According to a 2011 survey of the US-based National Association of Colleges and Employers, more than 50% of graduating college students had internship experiences (Nunley et al., 2016).

from the change in productivity due to endogenous worker sorting (*selection effect*).[3] To address this challenge directly, we design and implement a two-stage field experiment to disentangle the selection and incentive effects of work incentives.[4] As shown in Figure 1, study subjects are randomly assigned to one of three groups: (i) those who received job offers with a financial incentive (hereafter *wage group*), (ii) those who received job offers with a career incentive (hereafter *internship group*), and (iii) those who did not receive any job offers (control group). A one-time temporary work opportunity that provides a wage is offered to those assigned to the wage group. Those assigned to the internship group do not receive any wage, but rather (a) a potential long-term employment opportunity at the collaborating NGO as a regular employee and (b) a recommendation letter specifying their relative job performance, which essentially makes the offer an unpaid internship opportunity.[5] The control group receives no job offer.[6]

Individuals who accept the job offer in the first stage proceed to second-stage randomization. In the second stage, we randomly select half of the subjects in the *internship group* and additionally provide them the same financial incentive as that of the *wage group*. In the same manner, half of the subjects in the *wage group* are randomly selected to additionally receive the same career incentives as that of the *internship group*. This two-stage experimental design allows us to obtain two sub-groups with identical incentives (i.e., both financial and career incentives) during the work period. However, the channels through which these participants were attracted to accept the job offer are different. As a result, we can isolate the selection effect on labor productivity by comparing the two sub-groups (G2 and G3 in Figure 1).[7] In addition, we respectively measure the incentive effects of financial and career incentives on

---

[3] The incentive effect refers to differences in labor productivity when incentives affect work performance holding employee composition constant, while the selection effect refers to the difference in labor productivity driven by workers' self-selection into the job due to different work incentives and worker characteristics, such as ability, motivation, and personality.

[4] Our experimental design follows a two-stage randomization approach similar to that of Ashraf et al. (2010) and Beaman et al. (2014), but it is the first application for estimating the effect of a work incentive on labor productivity, to the best of our best knowledge.

[5] An entry-level regular worker position (enumerator or data entry clerk) at the NGO has career advancement prospects that sequentially lead to more advanced positions, such as a head enumerator, junior project assistant, senior project assistant, and project manager.

[6] The use of career incentives is related to the career concerns model suggested by Fama (1980), Gibbons and Murphy (1992), and Holmstrom (1999). The model explains that firms do not need to provide additional incentives to workers who could be motivated by future promotion and opportunities. It is also related to tournament theory suggested by Lazear and Robin (1981), who treat a promotion as a relative game since relative performance decides those who will be selected as regular workers.

[7] Due to the nature of our experimental design, the selection effect of either a career incentive or financial incentive is evaluated against the other incentive.

labor productivity. By comparing job performance between G3 and G4 in Figure 1, we isolate the incentive effect of career incentives on labor productivity. In the same way, by comparing job performance between G1 and G2, we isolate the incentive effect of financial incentives on labor productivity.

We find that even though those attracted to career incentives perform worse during the training period, they outperform those attracted to the financial incentive during the actual fieldwork period. In particular, the selection effect of a career incentive reduces survey errors by 29%. Moreover, a financial incentive leads those recruited by internship to increase their productivity. Non-cognitive skills including the big five personality traits, self-esteem, and intrinsic and extrinsic motivation explain 28.6% of the observed difference in the survey error rate between those recruited by the financial incentive channel and those recruited by the career incentive channel.

Our study contributes to the literature in the following ways. First, our study is related to literature which examines the impacts of work incentives on job performance.[8] For example, existing research shows that both financial incentives (Dal Bó et al., 2013) and career incentives (Ashraf et al., 2014) could attract more qualified workers and thereby increase job performance. In addition, Ashraf et al. (2014) points that career incentives do not hamper pro-social behavior, whereas Deserranno (2014) suggests that financial incentives could discourage those with pro-social preferences from applying for a job. However, to the best of our knowledge, no study has investigated how financial and career incentives influence labor productivity in the same setting, as our study.

Second, to the best of our knowledge, ours is the first study to examine the role of internships on worker selection and job performance. Most existing studies on internships are mainly descriptive and do not conduct causal analysis (Brooks et al., 1995; D'Abate et al., 2009; Friedman and Roodin, 2013; Liu et al., 2014). One exception is a résumé audit study on the effect of internship experience on employer callback (Nunley et al., 2016). This study randomly changes characteristics of résumés sent out to employers and finds that a résumé with internship experience receives 14% more callbacks from employers. However, the study does not analyze what types of job seekers are drawn to internships and how this affects labor productivity.

---

[8] Oyer and Schaefer (2005) provides an excellent survey of the literature.

Third, our paper contributes to the literature by disentangling the selection and incentive effects from the observed correlation between work incentives and labor productivity. Most studies in the literature estimate either the *selection effect* or *incentive effect* alone. One of the few studies which attempts to separate the selection effect from the incentive effect is Guiteras and Jack (2014).[9] They use the Becker–DeGroot–Marschak (BDM) mechanism (Becker et al., 1964) to infer the reservation wage and randomly vary the actual piece rate for bean-sorting work.[10] By comparing the work performance of workers with an identical reservation wage but different actual wages, they isolate the incentive effect from worker selection. However, there is little consensus that the BDM mechanism works properly in the field experiment setting.[11] Our research design jointly isolates the selection and incentive effects without relying on the indirect inference of an unobserved worker characteristic, such as through the BDM mechanism.

Fourth, our study adds new evidence to the growing importance of non-cognitive skills on labor market outcomes (Deming, 2015; Heckman and Rubinstein, 2001; Heckman et al., 2006; Kautz et al., 2014; Osborne-Groves, 2004; Park, 2015). For example, Deming (2015) shows that the US labor market has increasingly rewarded social skills since 1980 because a job that requires social skills is hard to automate. We show that workers with comparative advantages in a non-cognitive skill, such as extroversion, are more likely to select a job offer with career incentives than a job offer with financial incentives. In addition, we show that non-cognitive skills account for a significant portion of the observed productivity differences.

The remainder of the paper is structured as follows. Section 2 describes the background and context of our research. Section 3 outlines the research design and experimental stages. Section 4 describes the data and reports sample statistics. Section 5 presents the regression results. Section 6 concludes.

---

[9] Other studies that attempt to distinguish selection and incentive effects are based on evidence from the manufacturing sector in the US (Lazear, 2000), a controlled lab experiment (Dohmen and Falk, 2011), and a unique policy design in Italy (Gagliarducci and Nannicini, 2013). Other than the present study, Guiteras and Jack (2014) provide the only evidence from a field experiment thus far.

[10] In theory, the BDM mechanism truthfully reveals the reservation price of an individual, similar to a second-price auction.

[11] Berry et al. (2015) argue that the BDM mechanism could measure reservation prices (willingness to pay) successfully while Bohm et al. (1997) and Horowitz (2006) discuss that BDM is not incentive compatible in practice and could be biased in measuring reservation prices. Moreover, since revealing a reservation wage is not part of the ordinary employment process, job applicants might not be comfortable revealing their true reservation wage.

2. Background

2.1. Labor Market in Malawi

According to the World Bank (2015), Malawi's per capita gross domestic product in 2014 was USD 255, making it one of the poorest countries in the world, and life expectancy at birth was 55 years. The 2010 Malawi Demographic and Health Survey indicates that 19.6% of males aged between 20 and 29 years completed secondary school education. The official statistics show that about 80% of the working population is employed, but only 11% belong to the formal sector and their median income is MWK 13,400 (=USD 28.80) a month (NSO, 2014).[12] According to the 2010–2011 Integrated Household Survey, only about 30% of men residing in urban areas who had completed secondary education and were aged between 18 and 49 years participated in economic activities during the previous 7 days (Goldonton, 2014). A 2014 survey conducted by our collaborating NGO on a representative sample of males who just graduated from a secondary school in rural Lilongwe reveals that only 10% of them worked for pay and more than 60% of them were actively searching for jobs.

2.2. Recruiting Enumerators for the Population Census

We collaborated with Africa Future Foundation (AFF), an international NGO that provides health and education programs in Malawi. In 2015, AFF conducted a district-wide population census to collect demographic and socio-economic information on households in their catchment area, Chimutu district located outside of Lilongwe, the capital city of Malawi. The total population of Chimutu is about 90,000 individuals in 23,000 households, and it consists of 52 smaller catchment areas. Figure 2 shows a map of the project area.

AFF recruited enumerators from a sample pool of 536 male secondary school graduates who participated in the 2011 secondary school survey in Chimutu and nearby district. Specifically, the three qualifications for an enumerator position were that an enumerator 1) must be male owing to security concerns about the fieldwork, 2) must have graduated from a secondary school because field survey work requires some level of cognitive skills, and 3) must live near Chimutu because he should be familiar with the assigned local area to survey. In addition, AFF considered this recruitment drive an opportunity to construct a pool of potential

---

[12] USD = US dollar and MWK = Malawi kwacha. As of January 1, 2015, 1 USD was equivalent to 466 MWK. Throughout the paper, we use this as the currency exchange rate.

enumerators for subsequent surveys because the NGO was planning to hire some of the enumerators whose performance proved excellent during the 2015 population census.

Approaching potential job candidates who just graduated from the secondary school regardless of their employment and schooling status has the following advantages. First, contrary to most existing studies that observe only job applicants, we can observe the population of a young cohort who are potentially interested in a job opportunity in the local labor market. This unique feature of the sampling allows us to have strong external validity in that the sample of applicants would not necessarily be limited to job applicants. Second, one of the treatments in this study is to provide an internship offer, and most internship programs mainly target young and entry-level workers; therefore, it is most relevant that we conduct our study on those who have just graduated from a secondary school. Third, since the data cover individuals who do not participate in the survey, we could test whether and how the study participants differ from those who were invited but did not participate in this study using the 2011 secondary school survey information.

3. Experimental Design and Project Chronology

3.1. Experimental Design

We designed a two-stage randomized controlled trial that allows us to measure selection and incentive effects of financial and career incentive separately. In the first stage, study participants were randomly assigned to the wage group, the internship group, and control group, and those who agreed to work as enumerators were randomized further in the second stage. A randomly selected half of the internship group additionally received a financial incentive (fixed wage) while a randomly selected half of the wage group additionally received a career incentive.

This two-stage randomization leads us to have five study groups: Group 1 (G1) receives the career incentives only while Group 2 (G2) has both career and financial incentives. Group 4 (G4) receives the financial incentive in the form of a fixed wage while Group 3 (G3) has both career and financial incentives. It is important to note that G2 and G3 have identical incentives (both career and financial incentives), but they were attracted to accept the job offer by different work incentives. Therefore, comparing the performance of workers between G2 and G3 isolates

7

the selection effect of the career incentive evaluated against the financial incentive.[13] Any difference in job performance would be attributable to worker selection. In addition, comparing G1 and G2 isolates the effect of the financial incentive and comparing G3 and G4 isolates the effect of the career incentive.

To minimize the unexpected peer effects between workers with different incentives, the baseline survey and training were conducted separately for the first-stage internship group (G1 and G2) and the first-stage wage group (G3 and G4). In addition, we did not allow enumerators with different incentives to work in the same village.[14]

3.2. Project Chronology

Table 1 shows the number of study participants over the project chronological stages. We describe the chronology of the project as follows.

3.2.1. Recruitment and baseline survey

All 7,971subjects who participated in the 2011 secondary school student survey were invited for the follow-up survey in December 2014, but only to 536 males who met eligibility criteria of male high school graduates in 2014.[15] The follow-up survey in December 2014 serves as a baseline survey of this study. During the follow-up survey, the study participants were not aware of the possibility of receiving a job offer.

Approaching potential job candidates who just graduated from the secondary school regardless of their employment and schooling status has the following advantages. First, contrary to most existing studies that observe only job applicants, we can observe the population of a young cohort whose members are potentially interested in a job opportunity in the local labor market. This unique feature of the sampling allows us to have strong external validity of our findings. Second, one of the treatments is to provide an internship offer (i.e., a job offer with career incentives), and most internship programs mainly target young and entry-level workers; therefore, it is most relevant that we conduct our study on those who have just graduated from a

---

[13] Similarly, comparison of G2 and G3 can be interpreted as the selection effect of the financial incentive evaluated against the career incentive, but for the sake of convenience, we mainly focus on the career incentive.
[14] Note the presence of study participants in the control group who receive no job offer at all. We intentionally created the control group in the first-stage randomization so that we could study the long-term effects of short-term work experience when a follow-up survey becomes available.
[15] The 2011 survey covered secondary school students in Chimutu and nearby districts (Chitukula, Tsbango, and Kalumba). Therefore, study participants had stayed nearby the catchment area for at least 3 years.

secondary school. Third, since the data cover individuals who do not participate in the survey, we could test whether and how the study participants differ from those who were invited but did not participate in this study using the 2011 secondary school student survey information. Of the 536 male secondary school graduates we contacted, 443 successfully completed the baseline survey.[16]

3.2.2. First-stage randomization

On the completion of the baseline survey, those eligible for the census enumerator position randomly received job opportunities with different work incentives. 176 participants were given short-term paid job offers each with a fixed salary of MWK 10,000 for 20 days (equivalent to MWK 500 a day), and 186 participants were given job offers with career incentives, each comprising a recommendation letter and the prospect of a job opportunity at AFF as a regular staff member. A daily wage of MWK 500 is a competitive wage for young workers who have just graduated from secondary school.[17] In addition, the prospect of a regular staff position can be attractive: an entry-level enumerator position is competitive in terms of remuneration and offers a career advancement opportunity to a job seeker. We notified each participant who received a job offer with career incentives that there would be a significant chance of a long-term contract depending on job performance during the contract period and the AFF's job vacancies. However, we did not specify the precise probability of a long-term contract at the time of the announcement. Regarding the recommendation letter, we informed each survey participant who received a career incentive that the recommendation letter would be provided based on his relative performance.[18] The control group (81 individuals) participated in the survey but did not receive any offer.

3.2.3. Training

Those who accepted the job offer participated in a mandatory training program lasting 1 week to empower participants with the necessary skills and knowledge for the census. A job offer was valid conditional on the successful completion of the training. Training performance

---

[16] Those who did not participate in the survey (93 individuals) were unreachable (45%), or refused to participate because they were unwilling (13%), busy with schooling (32%), or busy working (9.7%).

[17] The median monthly salary of secondary school graduates in 2013 was 12,000 MWK, according to the Malawi Labor Force Survey (NSO, 2014).

[18] The recommendation letter was signed jointly by the director of AFF and the head of the Chimutu catchment district.

was measured by a quiz score and accuracy in a mock survey. The quiz tested knowledge on instructions for the procedures of the census enumeration work, provided during the training session. The full text of the quiz is in Appendix B. In addition, enumerator trainees conducted a mock census survey with other fellow trainees using the actual census survey questionnaire.

The acceptance rates were as follows: 74 out of 186 participants in the internship group (39.8%) and 74 out of 176 participants in the wage group (42%) chose to accept the offer. The acceptance rates between the internship group and the wage group were statistically similar. However, 11 trainees from the internship group who decided to work could not meet the minimum score requirement. As a result, 137 enumerators, 63 from the internship group and 74 from the wage group, were finally hired.

3.2.4. Second-stage randomization

Second-stage randomization was implemented immediately after the training. Before being dispatched to the assigned enumeration area, a randomly selected half of enumerators who were hired were given additional work incentives. Specifically, we announced an additional financial incentive in the form of a fixed salary of MWK 10,000 to a randomly selected half of the internship group (i.e., the same financial incentive that the wage group received in the first stage). Similarly, an additional career incentive in the form of a recommendation letter and prospect for long-term employment were given to a randomly selected half of the wage group (i.e., the same career incentive that the internship group received in the first stage). No one refused to accept additional incentives. Note that G1, G2, and G3 had the prospect of a future job opportunity at the end of the contract period while G4 did not.

Enumerators signed the employment contract documents according to their incentives determined by the two-stage randomization. The contract documents for each of G1, G2, G3, and G4 are included in Appendix C. The contract document specifies, regardless of the research group, that their performance will be evaluated by three measurements; 1) error rate, 2) speed, and 3) work attitude. Different work incentives for different incentive groups are described clearly in the contract. For example, a contract for G1 explicitly states that they will not be given any financial enumeration and that they will be provided with a recommendation letter and a future job opportunity based on their relative performance.

3.2.5. Census survey

After signing the employment contract, enumerators were dispatched to 52 catchment areas in late January 2015. Enumerators were randomly assigned to a catchment area stratified by population and land size. Enumerators in the same catchment area have the same incentive in order to prevent unexpected peer effects. In addition, enumerators were not assigned to areas from which they originally came, as locality would also affect their performance.

Once enumerators were deployed to the assigned catchment areas, each one worked independently on his own. To monitor and supervise the enumeration work and to collect the completed survey questionnaire, supervisors who have long experience of conducting field surveys in the past few years randomly visited enumerators at least once during the census.[19] Enumerators were aware that supervisors would visit them but did not know the dates. We report and discuss the effects of the supervisor visit on job performance in Appendix H.

4. Data

4.1. 2011 Secondary School Student Survey and 2014 Baseline Survey

We use data from the 2011 secondary school student survey to examine whether our baseline survey participants and non-participants are systematically different. About 9,000 students from all 33 secondary schools in four catchment districts located in rural Lilongwe were involved in the survey. The questionnaire consists of a variety of areas covering demographics, socio-economic status, health, and cognitive ability.

In addition, the 2014 baseline survey collects information on demographics, post-school training, employment history, health status, income and household assets, cognitive ability, and non-cognitive traits. Cognitive ability is measured by a cognitive ability index, defined as an average z-scores of the Raven's matrices test score, the math and English scores of 2014 Malawi School Certificate of Education (MSCE) test, and the verbal and clerical ability test scores of the O*NET ability test, following the approach of Kling et al. (2007).[20] Non-cognitive traits include

---

[19] Since a relatively small number of supervisors had to cover large areas, 37% of enumerators conducted only two visits during the work period.

[20] Raven's progressive matrices test is a non-verbal test of thinking and observation skills. The MSCE is a test that all Malawian students must take to graduate from a secondary school. O*NET test is a tool for career exploration. We use verbal and clerical perception ability test scores of O*NET, which are directly related to enumerator job characteristics.

likert-type psychometric scales on self-esteem, intrinsic motivation, extrinsic motivation, and five factors in the Big Five personality test (extraversion, openness, conscientiousness, agreeableness, and neuroticism). Appendix D provides the details of definitions of cognitive ability measures and non-cognitive traits used in this study.

4.2. Training and Job Performance Measures

Training performance is measured by 1) a test score of the quiz, which consists of 12 questions on the census survey procedures, and 2) accuracy in a mock survey.

Labor productivity during the census survey is measured by speed, accuracy, and attitude: survey speed is measured by the number of households he surveyed per day; survey accuracy is measured by the proportion of systematically inconsistent or incorrect entries[21]; work attitude is reported by supervisors.[22]

4.3. Baseline Characteristics and Randomization Balance

Table 2 presents the baseline characteristics of 443 study participants categorized by the first-stage randomization (Columns 2 to 5) and the second-stage randomization (Columns 6 and 7). In each row, we report the number of observations, means, and standard deviations of a corresponding variable by the first-stage randomization groups. In addition, we report the pairwise mean differences between the first-stage and second-stage randomization groups and their p-values for the test of equality.

Panel A in Table 2 shows the descriptive statistics on demographics and socio-economic status. Study participants are about 20 years old, reflecting our sampling criteria (secondary school graduates in 2014). Their height is on average about 165 cm, and the body mass index (BMI) is around 20 indicating normal weight. They have an average of 4.4 siblings. Parental support is a self-reported measure of the involvement and support of parents. The higher the score, the more supportive one's parents are. Asset score is a sum of three kinds of assets a survey participant's household own (improved toilet, refrigerator, and bicycle). On average,

---

[21] For example, if the birth year and age of the respondent are not compatible, it is considered an error made by an enumerator, not by a respondent, because the enumerator should have caught the error during the interview. Appendix E provides the details of how we calculate a survey error rate.

[22] Work attitude is an enumerator-specific measure evaluated by randomly assigned supervisors who visited enumerators without prior notice regarding their professional attitude toward respondents and supervisors.

survey participants own 1.15 out of three asset items. It is notable that only 9% are currently working even though the sample consists of male secondary school graduates reflecting a weak labor demand in Malawi.

Panel B reports statistics on non-cognitive traits and cognitive ability, in which a larger value of a variable means a stronger propensity to possess such a trait. Panel C illustrates the training performance among those hired as enumerators.

Except for number of siblings, all p-values of the test of equality between the first-stage career incentive group (G1 and G2) and the first-stage financial incentive group (G3 and G4) are above .05, as reported in column (6). There are some cases in which the p-values of the tests of joint equality are above .05 when we additionally include the control group, but that is mainly due to the control group.[23] Columns (10) and (11) of Table 2 report the group mean differences and the p-values of the test of equality between the second-stage randomization groups (G1 vs. G2 and G3 vs. G4). Except for the level of parental support, none of the equality tests has a p-value above .05.

Note that the original sample pool of 536 individuals was previously surveyed in 2011, so we could examine whether those who participated in the baseline survey and non-participants are systematically different. Appendix A shows that baseline survey participants and non-participants are not statistically different from each other in most dimensions. However, it is noteworthy that the non-participants are relatively wealthier than the survey participants in terms of an asset index, which reflects their higher opportunity cost of survey participation.

Correlations among cognitive ability and non-cognitive traits are reported in Appendix F. In Table F.1, among all survey participants, intrinsic motivation and extrinsic motivation have opposite relationships with cognitive abilities. Intrinsic motivation, or motivation by inner values, shows a positive relationship with cognitive abilities overall, whereas extrinsic motivation, or motivation by external values, has a negative relationship with cognitive abilities. All cognitive ability factors comprising the cognitive ability index show positive correlation, as expected. It is quite notable that in Table F.2, among enumerators, extroversion has a weakly negative correlation with cognitive abilities overall, but agreeableness has a clearly positive relationship with all cognitive abilities. It may be possible that the more agreeable and stable trait

---

[23] Since the control group is also randomly assigned, there is no particular reason why the control group behaves differently.

caused fewer mistakes during the cognitive ability evaluation process, and the more extroverted trait had the opposite effect.

5. Empirical Results

    5.1 Job-takers characteristics
        Table 3 shows the regression results of the following equation:

$$Accept_i = \alpha + \beta Internship_i + \delta Trait_i + \gamma Career_i * Trait_i + \epsilon_i \qquad (1)$$

where i denotes individual i. $Accept$ is a binary indicator variable whose value is 1 if individual i accepts a job offer, and 0 otherwise. $Internship$ equals 1 if individual i has an internship offer, and 0 otherwise. $Trait$ is a vector of traits individual i possess in which it can be either of demographic and socioeconomic characteristics, cognitive ability, and non-cognitive skills. ϵ_i is an error term. It is worthwhile to note that individuals were not aware of an additional incentive randomly to be given in the second stage.

        Column 1 of Table 3 shows that the take-up rates of the internship and the wage group are not statistically different, which is corresponding to the Raw C in Table 1. This means that the perceived market values of an internship offer with career incentives and a job offer with financial incentive measured by the job take-up rate are similar. Therefore, we compare the two different incentives equally appreciated in the labor market, but the composition of workers would be different due to self-selection made by workers.

        Even though the job take-up rates are similar, the composition of job takers under these two schemes could be different. Columns 2 to 15 of Table 3 test whether the career incentives, as opposed to the financial incentive, attracts people with different characteristics. The career incentive dummy is interacted one by one with socio-demographic factors (Columns 2–6), cognitive ability (Column 7), and non-cognitive traits (Columns 8–15). None of the coefficient estimates of the interaction term are statistically significant at the 5% level probably due to the weak statistical power. Table G.1 at appendix G compares job takers in the internship group and the wage group. It reports the means of individual characteristics (Columns 2 and 3) and pairwise mean differences (Column 4).[24] The results in Table G.1 confirm the results in Table 3. Even though we find job takers in internship group are more likely to be extrovert, which is

---
[24] The structure is similar to Table 1 but the sample is restricted to those who accepted a job offer.

14

corresponding to the results shown in Column 12 of Table 3, we acknowledge that this finding should be interpreted with caution due to the concern on multiple hypothesis testing.[25]

5.2. Training Outcome

Figure 3 shows the kernel density estimates of the training outcome measures. Panel A clearly shows that the financial incentive group performs better than the career incentive group in terms of the quiz score on enumerator tasks. Similarly, panel B indicates that the financial incentive group has a lower mock survey error rate than the career incentive group does.

Regression results using a variety of controls confirm the graphical evidence presented in Figure 3. Table 4 shows the estimates of the following regression specification:

$$Training_i = \alpha + \beta Career_i + \delta Demog_i + \gamma Cog_i + \theta NonCog_i + \omega_i \qquad (2)$$

where $Training_i$ is the training outcome of interest for individual i. As stated in Section 3, training performance is measured by a quiz score and a mock survey error rate. $Demog$ is a vector of demographic and socioeconomic characteristics. It includes age, number of siblings, an asset score, and the level of emotional support of parents. Cog is a cognitive ability index variable. $NonCog$ is a vector of non-cognitive traits including test scores of self-esteem, intrinsic motivation, extrinsic motivation, and Big Five personality types (extroversion, agreeableness, conscientiousness, emotional stability, and openness to experience). In the remainder of the paper, we repeatedly use $Demog, NonCog$, and Cog, and their definitions are the same across other equations. β captures the effect of career incentive on training performance evaluated against financial incentive holding other things. Note that Table 4 consists of Panels A and B. They are identical except for the sample composition. Panel A contains those who participated in job training. Panel B contains those hired as enumerators excluding those who failed training. Since 11 individuals failed out of 148 training participants, the regression results between the two panels are qualitatively the same. However, the magnitude of the coefficient estimate of *Career* is larger in Panel A because those who failed training only came from the internship group. In the following, we discuss the results in Panel B.

Column (1) of Table 4 shows that those attracted by the career incentive performed worse than those attracted by the financial incentive in terms of the quiz score when there is no control.

---

[25] We find only one significant result (extroversion) out of 15 traits.

On average, the internship group scored 1.4 points lower than the wage group. Our results are robust to the inclusion of different types of control variables in columns (2)–(5).

Consistent with the quiz score, we find that that those attracted by the career incentive performed worse than those attracted by the financial incentive in terms of mock survey errors. The error rate of the internship group is 8.8 percentage points higher that of the wage group in column (6). As we control for individual characteristics in Columns (7)–(10), the difference becomes somewhat smaller but remains significant at the 1% level.

Note that we do not yet distinguish between the selection and the incentive effects because the differences in training performance reflect not only the differences in unobserved ability between the internship and the wage group (selection effect) but also the differences in efforts between the two groups (incentive effect). We suspect that the internship group is likely to exert more effort than the wage group due to the job prospect nature of the career incentive. Thus, the coefficient estimates of β reported in this subsection could be biased toward rejecting the null of no impact of career incentive on training performance.

5.3. Selection effect of career incentive on job performance (G2 vs. G3)

Figure 4 presents the kernel density estimates of job performance measures of G2 and G3. G2 and G3 have both career and financial incentives but the incentive channel by which they are attracted to accept a job offer is different: G2 initially accepted a job offer that comes with career incentive while G3 initially accepted a job offer with financial incentive. Therefore, we argue that any difference in job performance between G2 and G3 is driven by worker sorting, which we call the *selection effect*.

The graphical evidence suggests that those attracted to the career incentive outperform those attracted to the financial incentive. Panel A of Figure 4 shows that the wage group performs better than the internship group in terms of survey accuracy (lower error rate). Similarly, panel B indicates that the internship group surveyed more households per day than the wage group did although the difference is small.

In Table 5, we present regression-based evidence of the selection effect of career incentive evaluated against financial incentive. We estimate the following equations:[26]

$$Performance_{ijkt} = \alpha + \beta Career_j + \delta Demog_j + \gamma Cog_j + \theta NonCog_j + \sigma_t + \emptyset Z_k + \psi_{ijkt} \quad (3)$$

---

[26] Standard errors are clustered at the enumerator level.

where $Performance_{ijkt}$ is an outcome variable for the survey collected from household i by enumerator j in catchment area k, observed on the t-th work day. We have three major outcome variables: $Error_{ijkt}$ is an error rate, $Speed_{jkt}$ is the number of households surveyed per day, and $Attitude_{jk}$ is a subjective measure of work attitude measured by a supervisor. Note that $Speed_{jkt}$ is aggregated by individual-work day level, and $Attitude_{jk}$ does not vary by survey date. We include supervisor fixed effects when the outcome variable is *Attitude*. Z_k is a vector of catchment area characteristics.[27] σ_t is the work-day fixed effect.

An implicit identifying assumption to claim the causal effect of career incentive on job performance is that G2 and G3 have identical incentives, i.e., the difference in the first-stage incentive is the only difference at the recruitment stage. We acknowledge that the sequence of the provision of first- and second-stage incentives was different by the experimental design. G2 received the career incentive offer first and then received the financial incentive offer; G3 received the financial incentive offer first and then received the career incentive offer. Even if the shadow price of the first-stage incentive is comparable between G2 and G3 (as implied by a small and statistically insignificant difference of job acceptance rate between G2 and G3), our estimate of β would be biased to the extent that the different sequence creates differential reference points for unexpected additional work incentives.

Column (1) of Panel A shows that G2 enumerators (the internship group) outperform G3 enumerators (the wage group) in terms of accuracy. This shows that the selection effect of career incentive reduces systematic errors caused by enumerators in the census survey by 29%.

Note that the specification in Column (1) does not control for demographic and socio-economic status, cognitive ability, and non-cognitive traits other than survey day fixed effect and survey area characteristics. To examine which factors drive this observed difference, we run the regression with different sets of control variables. Column (2) adds demographic and socio-economic status, Column (3) adds the cognitive ability index, Column (4) adds a set of non-cognitive abilities, and Column (5) includes them all.

The results reported in Columns (2) and (3) show that the factors of demographic and socio-economic status as well as of cognitive ability could explain the observed difference

---

[27] This includes the total number of households, total population, average asset score (refrigerator, bicycle, and improved toilet), average number of births in the last 3 years in a household, average incidence of malaria among under 3-year-olds, proportion of under 3-year-olds with the assistance of health professionals, and average number of deaths in households in the last 12 months.

between G2 and G3. However, the biggest reduction in the magnitude of the coefficient estimate comes when we add non-cognitive abilities, and the coefficient estimate is no longer statistically different from zero. We find that the selection effect decreases by 28.6% due to the inclusion of non-cognitive traits. This finding is reasonable because individuals with a more suitable non-cognitive trait, such as extroversion, were more responsive to career incentive offers than financial incentive offers.

Interestingly, Column (5) indicates that 41% of the original selection effect reported in Column (1) is due to the unobservables. As discussed in Ashraf et al. (2015), the finding that a large portion of the selection effect is unexplained indicates that screening via the observables might be imperfect and thus, it is important to devise recruitment to attract workers with strong unobservable skills via self-selection.

The fact that the internship group performed better than the wage group in Panel A is surprising because members of the internship group performed worse than those of the wage group during the training stage. A possible explanation for this finding is that those who have a comparative advantage in an enumerator job sort into the internship group. Enumeration fieldwork requires a particular non-cognitive skill, such as extroversion. Working in the field as an enumerator is quite different from training in the office in that the enumerator needs to introduce himself to many strangers and encourage them to speak about the details of their lives within a short period. Therefore, those who have strong interpersonal skills self-selected more into the internship group and performed significantly better.

However, we do not find similar evidence in speed and attitude. One reason is that the coefficients in Columns (6) and (10) are not precisely estimated, and thus, it is difficult to make a conclusive argument about the selection effect on those performance measures.

5.4. Incentive effect of career incentive (G3 vs. G4)

Panel B of Table 5 compares job performance between G3 and G4, which can also be observed in Figure 5. Members of both G3 and G4 accepted a job offer due to the financial incentive, but only G3 members received the additional career incentive. There was no worker selection in this additional provision of the second-stage incentive by design. Therefore, any difference in job performance between G3 and G4 members is due to the incentive effect of additional career incentive holding worker selection constant.

Column (11) of Panel B shows that the additional career incentive positively affects enumerators' work attitude. We find that the career incentive motivates workers to improve their work attitude by 34%. Interestingly, Columns (12)–(15) show that adding different sets of control variables does not change the estimated coefficient much. This implies that the observed improvement in work attitude is driven mostly by unobservable factors to the researchers.

A possible explanation for our findings is that the additional career incentives could act as a pressure to perform well in the survey since G3 participants now have an incentive to work better to be hired at the collaborating NGO and to obtain a better recommendation letter. This could have made enumerators spend more time in each household they survey. We find a decrease of survey speed by 9.7% (=1.08/11.1), but it did not translate into a decrease in the survey error rate, as reported in Columns (1)–(10). In addition, none of the coefficient estimates in Columns (1)–(10) is statistically significant at the 5% level.

It is difficult to rationalize why we do not observe a significant improvement in survey accuracy but we do observe a significant improvement in work attitude. One possible explanation is that it could have been relatively easy for the wage group enumerators to improve work attitude compared to other dimensions of job performance, such as error rate and speed, because enumerators belonging to the wage group (G3 and G4) were originally at the bottom of attitude distribution compared to the internship group (G1 and G2). Another explanation is that enumerators were trying to influence their supervisors' evaluations in their favor although it did not corresponding to the NGO's objective. If the supervisor evaluation was a primary or sole determinant of worker hiring decision in the future, less productive workers could have been selected due to the employee manipulation.

5.5 Incentive effect of financial incentive (G1 vs. G2)

Members of both G1 and G2 accepted a job offer with career incentive, but only G2 receives an additional financial incentive. Hence, we argue that any difference between G1 and G2 is due to the additional financial incentive effect reported in Panel C of Table 5, holding worker selection constant.

Columns (1)–(5) and (6)–(10) measure the impacts of the financial incentive among those recruited by the career incentive on the survey error rate and speed, respectively. The kernel density estimates of the survey error rate and speed are shown in Figure 6. The financial incentive does not affect the error rate, but it significantly increases survey speed. Those who received an additional financial incentive survey on average 2.1 (19.6%) more households per

day without compromising survey accuracy, as reported in Column (6). We find no strong evidence of the financial incentive effect on work attitude.

An increase in survey speed due to the financial incentive is consistent with the gift exchange model of the efficiency wage theory formulated by Akerlof (1984). In the original gift exchange model, a worker boosts his productivity upon receiving a gift from his/her employer that exceeds the minimum level of compensation for the minimum level of effort. Enumerators in G2 already have career incentive and the additional financial incentive is purely given to them as a surprise gift.

Similar to the career incentive effect, it is somewhat difficult to explain why the financial incentive effect is salient only in survey speed. One possible explanation is that it could have been relatively easy for the internship group enumerators to improve survey speed compared to other dimensions of job performance because their speed is relatively lower while the survey accuracy and work attitude are relatively better than the wage group (G3 and G4).

6. Concluding Remarks

The relationship between work incentives and job performance is highly endogenous and thus, it is difficult to establish a causal link with observational data. We overcome this limitation and disentangle the selection and incentive effects of career and financial incentives on job performance through a two-stage randomized controlled trial. We utilize an opportunity in which our collaborating NGO recruits enumerators for a population census of a rural catchment district in Malawi.

We show that those recruited by the internship channel perform better than those recruited by the financial incentive channel, and that non-cognitive skills play an important role in explaining this difference. In addition, a financial incentive motivates those recruited through an internship to work more productively. We also find that the career incentives motivate enumerators to improve their work attitude, but not productivity.

There are limitations to our study. First, due to the short-term nature of our job, we cannot study whether the estimated selection and incentive effects of career and financial incentives last over longer periods of time. Second, non-cognitive traits used in this study are psychometric scales measured based on a paper test. It would be interesting to know whether such paper-based non-cognitive traits are highly correlated with non-cognitive traits measured in the incentivized setting. Third, we acknowledge that our study setting is a census enumerator job

in developing countries. To the extent that there are differences in the nature of the job and the study context, our study would not have strong external validity. In the future, it would be interesting to conduct a similar study in developed countries or for different occupational types.

References

Akerlof, George A. 1984. "Gift Exchange and Efficiency–Wage Theory: Four Views." *American Economic Review: Papers and Proceedings*, 74(2): 79–83.

Ashraf, Nava, Oriana Bandiera, and Scott S. Lee. 2014. "Do-gooders and go-getters: career incentives, selection, and performance in public service delivery." Harvard Business School Working Paper.

Ashraf, Nava, James Berry, and Jesse M. Shapiro. 2010. "Can Higher Prices Stimulate Product Use? Evidence from a Field Experiment in Zambia." *American Economic Review*, 100(5): 2383-2413.

Beaman, Lori, Dean Karlan, Bram Thuysbaert, and Christopher Udry. 2014. "Self-Selection into Credit Markets: Evidence from Agriculture in Mali." NBER Working Paper No. 20387. National Bureau of Economic Research.

Becker, Gary, Morris H. DeGroot, and Jacob Marschak. 1964. "Measuring Utility by a Single-response Sequential Method." *Behavioral Science*, 9: 226–236.

Berry, James, Greg Fischer, and Raymond Guiteras. 2015. "Eliciting and Utilizing Willingness to Pay: Evidence from Field Trials in Northern Ghana." BREAD Working Paper No. 459. Bureau for Research and Economic Analysis of Development.

Bohm, P., J. Lindén, and J. Sonnegård. 1997. "Eliciting Reservation Prices: Becker–Degroot–Marschak Mechanisms vs. Markets." *The Economic Journal*, 107(443): 1079–1089.

Dal Bó, Ernesto, Frederico Finan, and Martin A. Rossi. 2013. "Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service." *Quarterly Journal of Economics*, 128(3): 1169–1218.

Deming, David J. 2015. "The Growing Importance of Social Skills in the Labor Market." NBER Working Paper No. 21473. National Bureau of Economic Research.

Deserranno, Erika. 2014. "Financial Incentives as Signals: Experimental Evidence from the Recruitment of Health Workers." Working Paper.

Dohmen, T. and Falk, A., 2011. "Performance pay and multidimensional sorting: Productivity, preferences, and gender." The *American Economic Review*, pp.556-590.

Fama, E.F., 1980. "Agency Problems and the Theory of the Firm." *The Journal of Political Economy*, pp.288-307.

Gagliarducci, S. and Nannicini, T., 2013. Do better paid politicians perform better? Disentangling incentives from selection. Journal of the European Economic Association, 11(2), pp.369-398.

Gibbons, R. and Murphy, K.J., 1992. Does executive compensation affect investment?. NBER Working Paper, (w4135).

Goldonton, Susan. 2014. "Employment Risk and Job-Seeker Performance." IFPRI Discussion Paper No. 1332. The International Food Policy Research Institute.

Guiteras, Raymond P., and B. Kelsey Jack. 2014. "Incentives, selection and productivity in labor markets: Evidence from rural Malawi." NBER Working paper No. 19825. National Bureau of Economic Research.

Heckman, James J., and Yona Rubinstein. 2001. "The Importance of Noncognitive Skills: Lessons from the GED Testing Program." *American Economic Review: Papers and Proceedings*, 91(2): 145–149.

Heckman, James J., Jora Stixrud, and Sergio Urzua. 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics*, 24(3): 411–482.

Holmström, B., 1999. Managerial incentive problems: A dynamic perspective. The Review of Economic Studies, 66(1), pp.169-182.

Horowitz, J.K. 2006. "The Becker–DeGroot–Marschak Mechanism is Not Necessarily Incentive Compatible, Even for Non-random Goods." *Economics Letters*, 93(1): 6–11.

Kautz, T., James Heckman, R. Diris, B. Ter Weel, and L. Borghans. 2014. "Fostering and Measuring Skills: Improving Cognitive and Non-Cognitive Skills to Promote Lifetime Success." NBER Working Paper No. 20749. National Bureau of Economic Research.

Kling, Jeffrey, Jeffrey Liebman and Lawrence Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica*, 75(1): 83–119.

Kim, Booyuel, Hyuncheol Kim, and Cristian Pop-Eleches. 2015. "Complementarities and Externalities of HIV/AIDS Prevention Programs: HIV/AIDS Education, Male Circumcision, and Cash Transfers." Cornell University Working Paper.

Lazear, Edward P. 2000. "Performance Pay and Productivity." *American Economic Review*, 90(5): 1346–1361.

Lazear, Edward P. and Rosen, Sherwin. 1981. "Rank-Order Tournaments as Optimum Labor Contracts." Journal of Political Economy, Vol. 89, No. 5, pp. 841-864.

National Statistical Office (NSO). 2014. Malawi Labour Force Survey 2013. Zomba, Malawi.

National Statistical Office (NSO). 2010. Malawi Demographic and Health Survey.

Malawi National Examination Board (MANEB). 2014. Malawi School Certificate of Education Examination- Grades and Awards for Candidates, Malawi.

Osborne-Groves, Melissa. 2004. Personality and the Intergenerational Transmission of Earnings. In Samuel Bowles, Herbert Gintis, and Melissa Osborne-Groves (Eds) Unequal Chances: Family Background and Economic Success pp. 208–231. Princeton University Press: Princeton.

Oyer, P. and Schaefer, S., 2011. American BigLaw lawyers and the schools that produce them: a profile and rankings.

Park, Sangyoon. 2015. "Socializing at Work: Evidence from a Field Experiment with Manufacturing Workers." Working Paper.

Salop, J. and Salop, S., 1976. Self-selection and turnover in the labor market. The Quarterly Journal of Economics, pp.619-627.

World Bank. 2015. World Development Indicators. http://data.worldbank.org/country/malawi

Figure 1: Experimental Design

Figure 1: Experimental Design

1st stage
Randomization

Career Incentive
(Internship Group)

- Recommendation letter
- An opportunity to become
  a regular employee

Financial Incentive
(Wage Group)

- Wage: 500 kwacha per day

Control

- No offer

2nd stage
Randomization

2nd stage
Randomization

G1. Career
incentive only

G2. Career and
Financial incentives

G3. Financial and
Career incentives

G4. Financial
incentive only

Figure 2: Project Area (Chimutu)

Map data @ 2015 Google

Figure 3: Kernel density of training outcomes

A. Quiz score

B. Error rate in mock survey
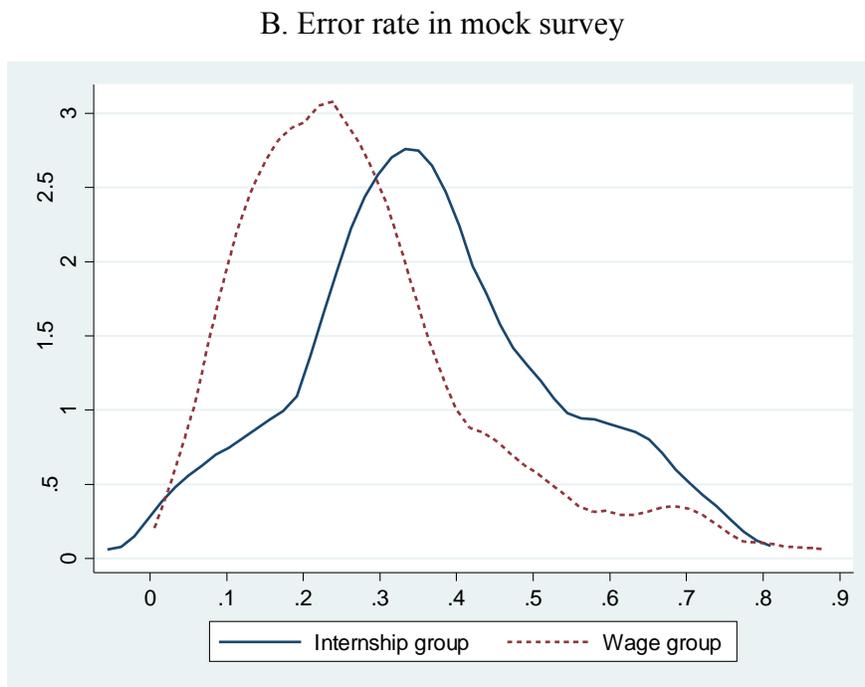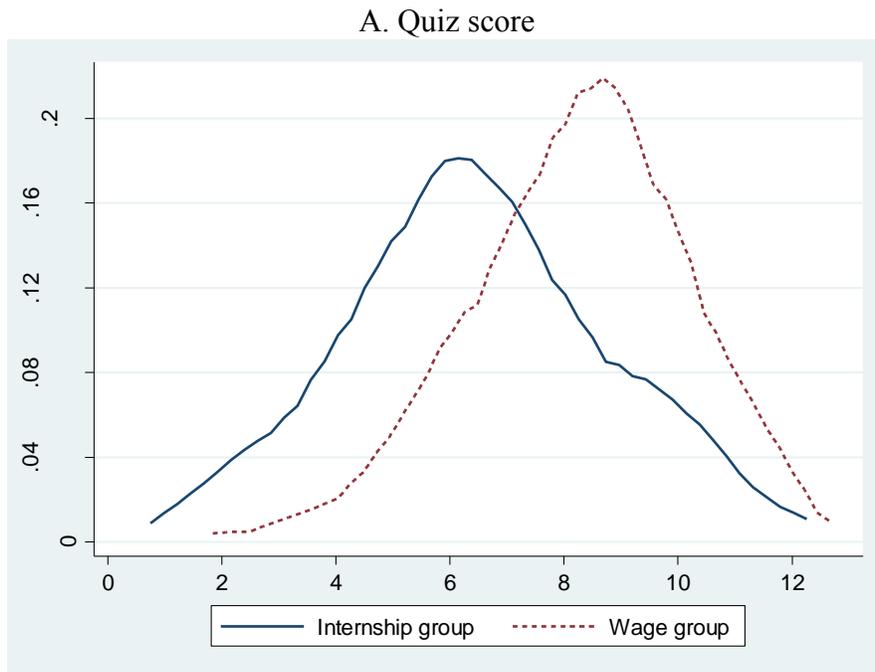
Figure 4 Selection effect (G2 vs. G3)

Panel A. Survey error rate
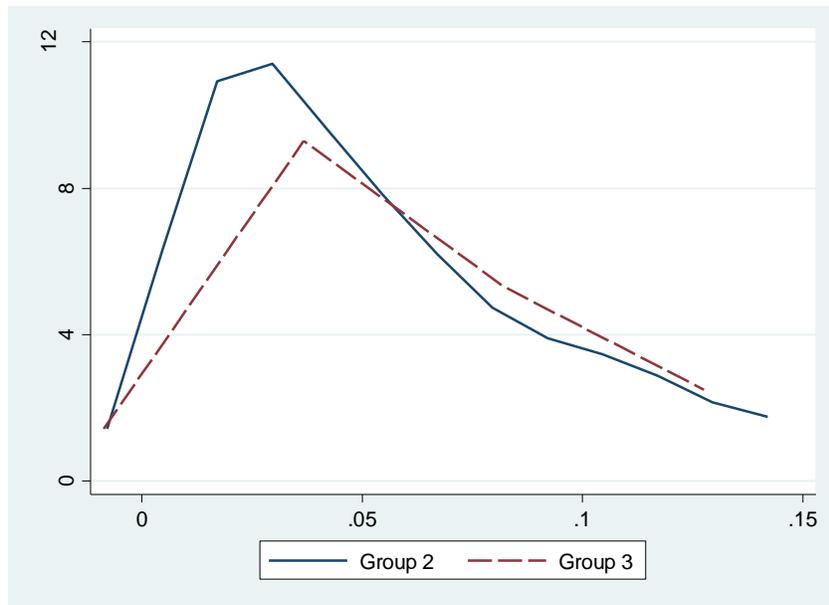
Mean	Group 2	.066
	Group 3	.077

Panel B. Survey speed (number of questionnaires surveyed per day)
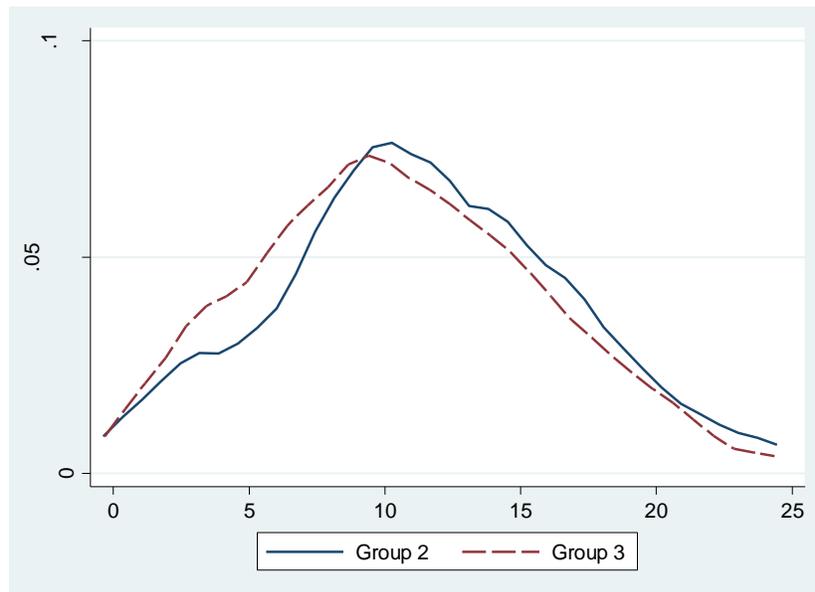
Mean	Group 2	11.6
	Group 3	10.7

Note: Group 2 was offered a career incentive in the first stage and an additional financial incentive in the second stage. Group 3 was offered a financial incentive in the first stage and an additional career incentive in the second stage.

Figure 5: Incentive effect of career incentive (G3 vs. G4)

Panel A. Survey error rate
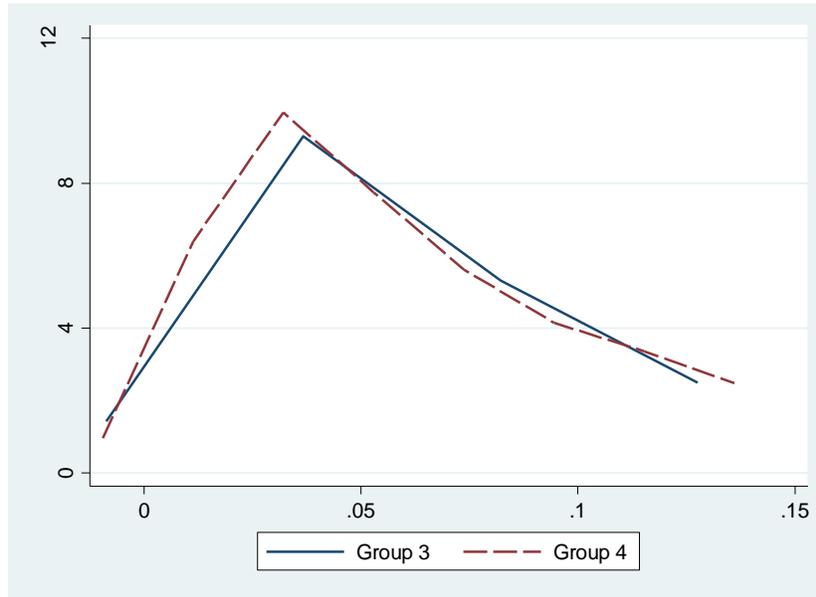
Mean	Group 3[*]	.077
	Group 4[**]	.082

Panel B. Survey speed (number of questionnaires surveyed per day)
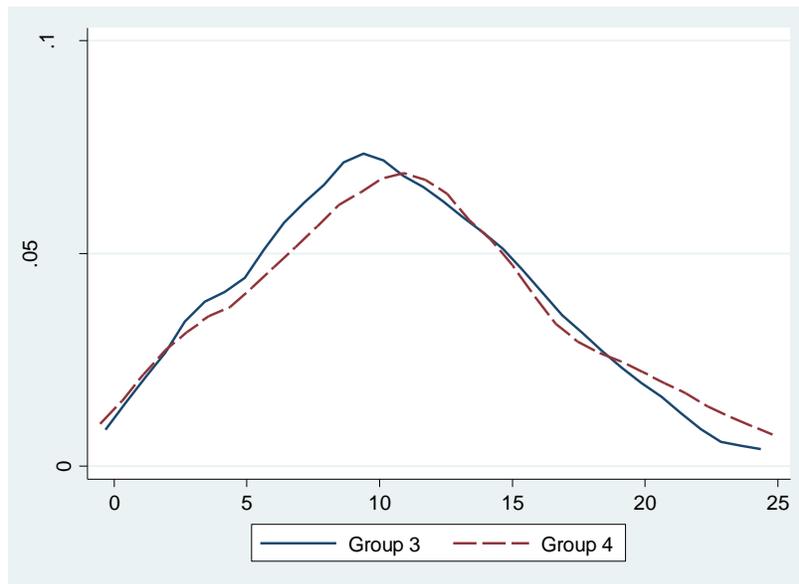
Mean	Group 3	10.7
	Group 4	11.5

Note: Group 3 was offered a financial incentive and provided an additional career incentive while Group 4 was offered a financial incentive only.

Figure 6: Incentive effect of financial incentive (G1 vs. G2)

Panel A. Survey error rate

Mean	Group 1[*]	.075
	Group 2[**]	.066

Panel B. Survey speed (number of questionnaires surveyed per day)

Mean	Group 1	9.84
	Group 2	11.6
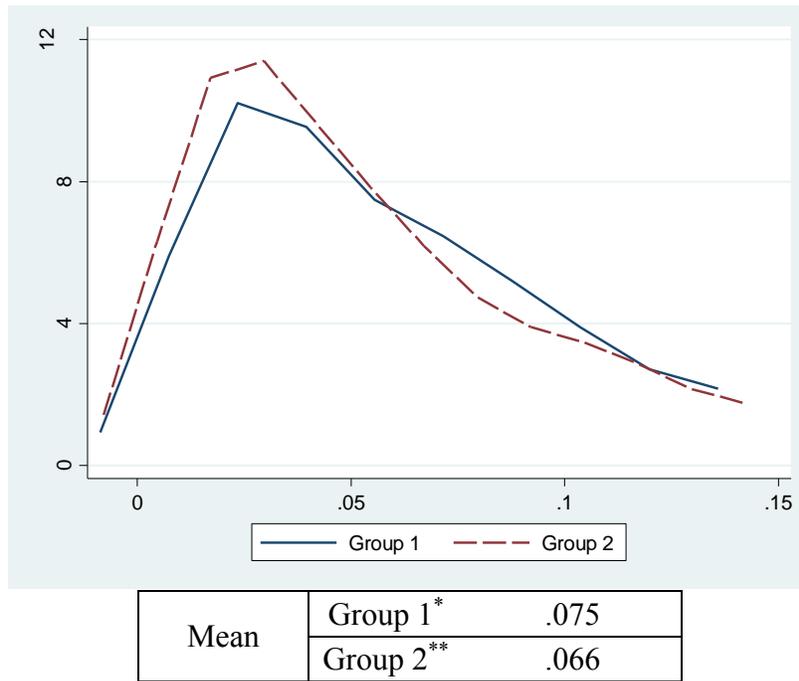
Note: Group 1 was offered a career incentive only while Group 2 was offered a career incentive and provided an additional financial incentive.

Figure 7: Quantile regression

Panel A. Survey error rate
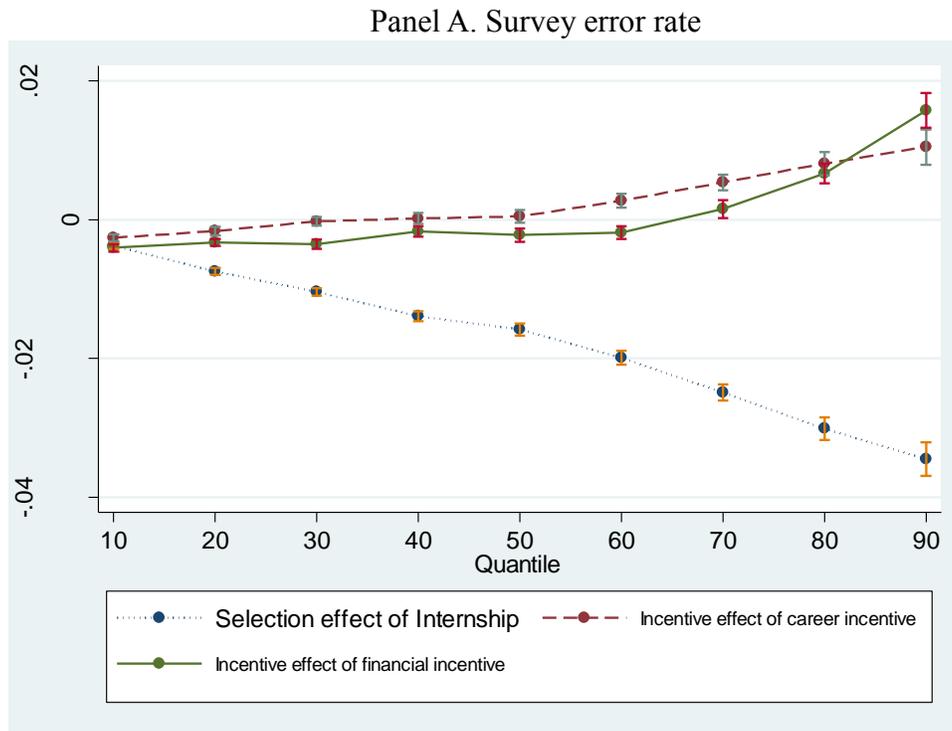
Panel B. Survey speed (number of questionnaires surveyed per day)
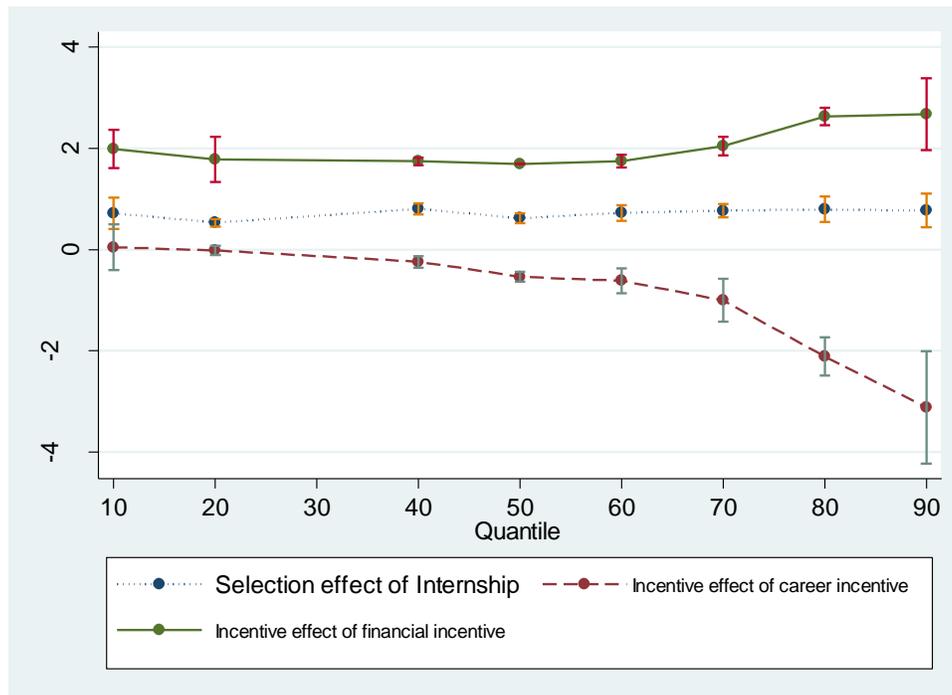
Table 1: Chronological order of the experiment stages

	Stage of experiment	Number of individuals						
		G1 (career incentive only)	G2 (career and additional financial incentives)	G3 (financial and additional career incentives)	G4 (financial incentive only)	Control	P-value	Total
A	Original target subjects	220		220		96		536
B (B/A)	Participated in the baseline survey	186 (84.6%)		176 (80.0%)		81 (84.4%)	.402 (F-stat)	443
C (C/B)	Accepted the conditional job offer	74 (39.8%)		74 (42.0%)		-	.663 (T-stat)	148
D	Failed training	11		0		-	-	11
E (E/B)	Hired as enumerators	63 (33.9%)		74 (42.0%)		-	-	137
		33	30	35	39			

Note: The proportion of individuals remaining at each stage is in parentheses.

32

Table 2: First-stage and second-stage randomization balance check

			First-stage randomization			Second-stage randomization	
Variable	Obs	Full sample	Internship group (G1, G2)	Wage group (G3, G4)	Mean difference (p-value)	Mean difference (p-value) G2 − G1	Mean difference (p-value) G3 − G4
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Demographics and socio-economic status							
Age	443	20.3 (1.61)	20.5 (.120)	20.4 (.126)	.065 (.707)	-.200 (.629)	-.207 (.520)
Height (cm)	441	164.5 (7.70)	164.5 (.625)	164.7 (.556)	-.241 (.774)	1.64 (.343)	1.88 (.256)
BMI (kg/m^2)	441	19.7 (2.10)	19.7 (.165)	19.8 (.151)	-.070 (.756)	-.097 (.868)	.234 (.590)
Number of siblings	443	4.41 (1.83)	4.60 (.132)	4.17 (.134)	.432 (.022)	5.00 (.315)	-.158 (.650)
Parental support	443	15.5 (4.87)	15.3 (.360)	15.5 (.338)	-0.2 (.766)	4.30** (.003)	-.790 (.415)
Asset score	443	1.15 (.901)	1.09 (.066)	1.19 (.067)	-.102 (.282)	.133 (.489)	.048 (.799)
Currently working	442	.088 (.284)	.097 (.022)	.074 (.020)	.023 (.436)	.036 (.514)	-.006 (.913)
Panel B. Non-cognitive traits and cognitive ability							
Self-esteem (Rosenberg scale)	443	19.5 (3.72)	19.4 (3.86)	19.3 (3.51)	.100 (.683)	.441 (.662)	-.768 (.341)
Intrinsic motivation	443	3.10 (.346)	3.10 (.330)	3.09 (.351)	.010 (.644)	.033 (.642)	-.075 (.372)
Extrinsic motivation	442	2.84 (.286)	2.84 (.281)	2.84 (.285)	.00 (.896)	.031 (.646)	.004 (.956)
Extroversion	433	3.53 (1.18)	3.61 (1.12)	3.47 (1.20)	.140 (.237)	.055 (.851)	-.246 (.393)
Agreeableness	438	5.17 (1.38)	5.13 (1.41)	5.10 (1.37)	.030 (.835)	-.165 (.651)	-.268 (.408)
Conscientiousness	442	5.77 (1.34)	5.69 (1.34)	5.68 (1.36)	.010 (.908)	.094 (.778)	-.054 (.850)
Emotional stability	439	5.12 (1.46)	5.08 (1.49)	5.06 (1.42)	.020 (.905)	.064 (.866)	-.190 (.591)
Openness to experiences	439	5.43 (1.34)	5.39 (1.35)	5.32 (1.36)	.070 (.664)	.441 (.187)	-.016 (.958)
Cognitive ability index	443	-.001 (.645)	-.019 (.047)	.049 (.049)	-.068 (.314)	.092 (.556)	.001 (.995)
Panel C. Training performance							
Quiz score	137					.221 (.638)	.101 (.816)
Mock survey error	137					-.036 (.409)	.001 (.965)
Number of observations		443	186	176		63	74

Note: ***, **, and * denote significance level at 1%, 5%, and 10%, respectively.

Table 3: Job-takers characteristics

	Panel A. Demographic and socio-economic characteristics and cognitive ability						
Trait	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Age	BMI	Number of siblings	Asset score	Currently working	Cognitive ability index
Trait		.042	-.030	.039*	-.072*	-.122	-.127**
		(.029)	(.018)	(.020)	(.040)	(.134)	(.054)
Internship	-.023	-.312	-.939	-.011	-.024	-.025	-.032
	(.052)	(.719)	(.482)	(.132)	(.085)	(.055)	(.051)
Trait * Internship		.014	.046	-.006	-.006	.050	-.052
		(.035)	(.024)	(.029)	(.055)	(.179)	(.074)
Constant	.420***	-.436	1.02	.259***	.506***	.429***	.427***
	(.037)	(.586)	(.353)	(.087)	(.063)	(.039)	(.037)
Observations	362	362	360	362	362	362	362
R-squared	.001	.028	.011	.017	.019	.004	.041
Mean (SD) of trait	-	20.4(1.65)	19.8(2.13)	4.39(1.80)	1.14(.896)	.086(.280)	.014(.642)

	Panel B. Non-cognitive traits							
Trait =	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
	Self-esteem (Rosenberg)	Intrinsic motivation	Extrinsic motivation	Extroversion	Agreeableness	Conscientiousness	Emotional stability	Openness to experiences
Trait	-.024**	-.012	-.029	-.059*	.001	.045*	.014	.045*
	(.010)	(.108)	(.132)	(.032)	(.028)	(.026)	(.027)	(.027)
Internship	-.308	.502	.707	-.306*	.025	.252	.173	.258
	(.275)	(.490)	(.502)	(.169)	(.200)	(.216)	(.194)	(.208)
Trait * Internship	.015	-.169	-.256	.079*	-.010	-.049	-.039	-.053
	(.014)	(.157)	(.175)	(.045)	(.038)	(.037)	(.037)	(.038)
Constant	.876***	.458	.501	.634***	.415***	.169	.353**	.182
	(.201)	(.335)	(.378)	(.117)	(.146)	(.147)	(.143)	(.145)
Observations	362	362	361	358	361	361	360	360
R-squared	.017	.008	.014	.012	.001	.008	.004	.008
Mean (SD)	19.3(3.69)	3.09(.340)	2.84(.282)	3.54(1.16)	5.11(1.39)	5.68(1.35)	5.07(1.45)	5.36(1.35)

Note: Robust standard errors are reported in parentheses; Dependent variable is 1 if an individual accepted our job based upon our initial first-stage randomization before knowing their second-stage randomization result, 0 otherwise. ***, **, and * denote significance at 1%, 5%, and 10%, respectively.

Table 4: Training performance regression

Dependent variable	Quiz score					Mock survey error rate				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Trainees										
Internship	-2.01***	-2.04***	-1.82***	-2.12***	-1.96***	.100***	.098***	.095***	.092***	.085***
	(.333)	(.347)	(.300)	(.329)	(.314)	(.026)	(.026)	(.026)	(.026)	(.028)
Constant	8.43***	11.0***	8.56***	3.00	8.48***	.272***	.315***	.268***	.442**	.364*
	(.212)	(.876)	(.184)	(2.44)	(2.50)	(.017)	(.062)	(.016)	(.181)	(.202)
Demographic and SES	X	O	X	X	O	X	O	X	X	O
Cognitive ability	X	X	O	X	O	X	X	O	X	O
Non-cognitive ability	X	X	X	O	O	X	X	X	O	O
Observations	148	148	148	147	147	148	148	148	147	147
R-squared	.200	.291	.378	.338	.515	.092	.169	.113	.137	.224
Mean (SD)	7.43 (2.26)					.321 (.165)				
Panel B: Enumerators										
Internship	-1.45***	-1.43***	-1.38***	-1.56***	-1.47***	.089***	.087***	.087***	.079***	.076**
	(.313)	(.323)	(.287)	(.306)	(.295)	(.027)	(.028)	(.027)	(.027)	(.029)
Constant	8.43***	11.3***	8.54***	2.48	7.68***	.272***	.316***	.269***	.483**	.448*
	(.212)	(.821)	(.184)	(2.34)	(2.32)	(.017)	(.063)	(.016)	(.203)	(.237)
Demographic and SES	X	O	X	X	O	X	O	X	X	O
Cognitive ability	X	X	O	X	O	X	X	O	X	O
Non-cognitive ability	X	X	X	O	O	X	X	X	O	O
Observations	137	137	137	136	136	137	137	137	136	137
R-squared	.137	.263	.304	.290	.488	.074	.152	.094	.135	.219
Mean (SD)	7.77 (1.96)					.312 (.162)				

Note: Robust standard errors are reported in parentheses. Demographic and SES includes age, number of siblings, parental support, and asset score (improved toilet, refrigerator, and bicycle). Cognitive ability includes ability index (Raven's matrices test, clerical and verbal tests from O*NET Ability Profiler, and MSCE, 2014). Non-cognitive ability controls include tests for self-esteem, intrinsic and extrinsic motivation, and Big Five personality test (extroversion, agreeableness, conscientiousness, emotional stability, and openness to experience). ***, **, and * denote significance at 1%, 5%, and 10%, respectively.

Table 5: Job performance regression

VARIABLES	Error rate					Speed					Attitude				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Panel A: Selection effect of internship															
G2 (vs. G3)	-.021*	-.018*	-.020**	-.015	-.009	.577	.673	.582	.424	.706	-.045	.010	-.042	-.108	-.069
	(.012)	(.011)	(.010)	(.011)	(.008)	(.479)	(.507)	(.488)	(.432)	(.441)	(.101)	(.126)	(.100)	(.101)	(.137)
Observations	11,134	11,134	11,134	11,134	11,134	1,003	1,003	1,003	1,003	1,003	65	65	65	65	65
R-squared	.093	.165	.179	.135	.263	.128	.141	.128	.146	.163	.383	.491	.386	.501	.606
Mean (SD)	.072 (.071)					11.1 (5.50)					.796 (.171)				
Panel B: Incentive effect of career incentives															
G3 (vs. G4)	.006	.006	.007	.007	.006	-1.08	-.905	-1.07	-1.35*	-1.25*	.240***	.241***	.238***	.244***	.238***
	(.013)	(.012)	(.012)	(.013)	(.012)	(.698)	(.619)	(.698)	(.700)	(.666)	(.047)	(.047)	(.049)	(.054)	(.054)
Observations	11,775	11,775	11,775	11,775	11,775	1,063	1,063	1,063	1,063	1,063	74	74	74	74	74
R-squared	.137	.167	.158	.182	.215	.113	.136	.113	.136	.159	.617	.699	.620	.634	.731
Mean (SD)	.080 (.076)					11.1 (5.92)					.709 (.194)				
Panel C: Incentive effect of financial incentive															
G2 (vs. G1)	-.003	-.0004	-.005	-.002	-.002	2.10***	2.26***	2.10***	1.71***	1.81***	.048	.054	.049	.086	.107
	(.010)	(.010)	(.007)	(.010)	(.008)	(.545)	(.598)	(.545)	(.557)	(.635)	(.061)	(.084)	(.063)	(.081)	(.101)
Observations	9,785	9,785	9,785	9,647	9,647	914	914	914	899	899	63	63	63	62	62
R-squared	.160	.260	.253	.187	.348	.169	.182	.169	.191	.208	.366	.441	.367	.482	.576
Mean (SD)	.070 (.064)					10.7 (5.42)					.770 (.164)				
Demographic and SES	NO	YES	NO	NO	YES	NO	YES	NO	NO	YES	NO	YES	NO	NO	YES
Cognitive ability	NO	NO	YES	NO	YES	NO	NO	YES	NO	YES	NO	NO	YES	NO	YES
Non-cognitive ability	NO	NO	NO	YES	YES	NO	NO	NO	YES	YES	NO	NO	NO	YES	YES

Note: Standard errors clustered at the enumerator level are reported in parentheses in Columns (1)–(10). All specifications contain catchment area controls, which include the total number of households, total population, asset score (own refrigerator, bicycle, and improved toilet), birth rate in the last 3 years, incidence of malaria among under 3-year-olds, proportion of under 3-year-olds born with the assistance of health professionals, and deaths in the last 12 months. Columns (1)–(10) include the work-day fixed effects. Columns (11)–(15) include supervisor fixed effects. ***, **, and * denote significance at 1%, 5%, and 10%, respectively.

A. Balance Check between Baseline Survey Participants and Non-participants

Table A1: Balance check between those who participated in the enumerator baseline survey and those who did not (based on information collected from the survey in 2011)

Variable	Balance check based on t-test for all participants			
	Full sample	Survey participants	Survey non-participants	Mean difference between (2) and (3) (p-value)
	(1)	(2)	(3)	(4)
Height	164.5	164.5	164.5	.047
	(7.60)	(.743)	(.367)	(.955)
Weight	53.6	53.9	53.5	-.430
	(7.64)	(.984)	(.342)	(.680)
Age	16.1	16.0	16.1	.065
	(1.56)	(.197)	(.070)	(.758)
Living with a father	.640	.645	.639	-.006
	(.480)	(.050)	(.023)	(.908)
Living with a mother	.733	.667	.747	.081
	(.443)	(.049)	(.021)	(.134)
Asset score	1.21	1.41	1.17	-.240**
	(.910)	(.106)	(.042)	(.037)
High self-health rating	.451	.538	.433	.104*
	(.498)	(.052)	(.024)	(.070)
Raven's matrices test score	19.8	18.7	20.0	1.32*
	(4.96)	(.696)	(.244)	(.077)
2014 Malawi School Certificate Exam score	11.6	11.5	11.6	.075
	(3.93)	(.442)	(.197)	(.877)
Number of observations	536	96	443	

Note: Columns (1)–(3) show group-specific means and standard deviations. As indicated in the table, 536 male secondary school graduates were contacted and invited for the baseline survey but 443 individuals showed up on the survey date. The asset score is constructed as the number of items an individual has access to at home: (1) flush/improved toilet, (2) refrigerator, and (3) bicycle. ***, **, and * denote significance level at 1%, 5%, and 10%, respectively

B. Training Quiz Questionnaire

No.	Question	Answer (Point)
1	An important reason for conducting the census is to achieve an improvement of overall quality of health in TA Chimutu. Describe the other two reasons why we conduct the census.	a. To make possible to reach out to every pregnant women who wanted to participate AFF MCH program. (0.5) b. To enrich the stock of socio-demographic data in T/A Chimutu that is necessary for elaboration of AFF MCH program. (0.5)
2	Regarding the roles of the enumerator, there are two functions you should NOT perform. Please fill them in the blank spaces below. A) Not to _____ B) Not to _____	a. Not to make any influence on answers (0.5) b. Not to change orders or words of questions (0.5)
3	What is the main standard required for households to be enumerated in the "2015 census of TA Chimutu," a modified version of the "population and housing census"?	Enumeration of all people, all housing units, and all other structures in TA Chimutu, who have stayed in TA Chimutu for more than 3 months during the past 12 months (1)
4	What is the name of the document that proves your eligibility to conduct the census?	Endorsement letter (1)
5	As what kind of structure would you categorize the following? *"A structure with sun-dried brick walls and asbestos roof"*	Semi-permanent (1)
6	Choose one that is **not** counted as a collective household. A) Hospitals, including three staff houses sharing food B) Lodge, including staff dwelling and sharing food C) Prison with many inmates' dwelling D) Store with owner's dwelling E) Military barracks with soldiers' dwelling	D (1)
7	What is the name of the document you have to sign before you start enumeration?	Consent form (1)
8	What are the three things you have to check before you leave the household?	Questionnaire, outbuildings, and Household ID number. (1, 0.5 point for partially correct)
9	What number do you put when you cannot meet any respondent from the household?	a. Do not put any number and just note down the household. (0.5) b. Put a latest number on it if you arrange to meet later. (0.5)
10	Your distributed alphabet is "C" and this household is the third household you enumerated in the catchment area. How did you place an ID number on the wall of the household?	0003C (1)
11	True or false questions A) It is okay if the questionnaire gets wet when there is heavy rain. B) You should not come to the completion meeting if you did not finish enumeration of your area. C) If you complete enumeration in your area, you should report to your supervisors immediately. D) You should bring all your housing necessities to the kickoff meeting.	A) False (0.5) B) False (0.5) C) True (0.5) D) True (0.5)

C. Contract Letters
C.1. G1 (career incentive only)

22 January 2015   Mr _____

<div align="center">

INTERNSHIP PROGRAM CONTRACT

</div>

The Management of Project Chimutu has the pleasure to offer you an internship opportunity on the following terms and conditions.

JOB TITLE: **ENUMERATOR**

1.0 TERMS OF CONTRACT

This is a maximum 30-day contract and will be effective from 23 January to 21 February 2015. You will be released from this contract as soon as you receive approval from your supervisor upon the completion of your assignment.

2.0 WORK SCHEDULE

You will be required to work from Monday to Sunday for up to 30 continuous days. If you complete the enumeration of your assigned catchment area before 21 February 2015 (30th day after 23 January 2015), you can report to your supervisor and terminate your contract earlier. Your official working hours are from 07:30am to 04:30pm, but you are strongly required to manage flexible working hours. You may work even before 07:30am or after 04:30pm, whenever it is necessary.

3.0 PERFORMANCE EVALUATION

Your work performance will be evaluated in terms of speed and accuracy of your enumeration. Therefore, quick and accurate enumeration is strongly encouraged. Also, if you complete the assigned enumeration work before the end of the contract, it will be highly appreciated.

Note that, after the census is completed, the supervisors will re-enumerate every catchment area again and evaluate each one's enumeration in terms of accuracy and your attitude toward household members you interviewed.

Hence, your performance will affect your recommendation letter and a future job opportunity at the AFF.

4.0 SALARY

This position is an unpaid internship so you will not be given any financial remuneration throughout the contract period (i.e., no salary). You shall be provided an accommodation in your assigned catchment area during the contract period.

5.0 INTERNSHIP PROVISIONS

Upon the successful completion of the contract, you will be given **an official certificate**, which certifies that you worked as a census enumerator for the Africa Future Foundation (AFF) project, and **a recommendation letter** from the director of Project Chimutu (Mr. Hanyoun So) and the chief of TA Chimutu upon your request for your future job applications.

The recommendation letter will specify your relative performance of the enumerator work compared to your peers. In other words, if you do a good job, the recommendation letter will say so, whereas, if you do a bad job, the recommendation letter will say so.

In addition, upon the successful completion of the contract, you will be considered for the **future hire** of a regular staff position at the AFF office if you show outstanding performance satisfying the standard of the management of AFF.

6.0 TERMINATION OF EMPLOYMENT

In the event of any violation of any of the terms of this contract by you, the Management of Project Chimutu may terminate employment without notice and compensation.

I am looking forward to a cordial and mutual relationship.
Yours faithfully,

Project Director
Hanyoun So

I,…………………………………………………………, have read and understood the above basic terms and conditions of service and hereby accept the offer as stipulated therein.

Signature:………………………………………………………………………………………

C2. G2 (1st stage career incentive + 2nd stage financial incentive) and G3 (1st stage financial + 2nd stage career incentives); G2 and G3 have the same contents in the contract letter.

23 January 2015   Mr _____

<div align="center">**INTERNSHIP PROGRAM CONTRACT**</div>

The Management of Project Chimutu has the pleasure to offer you an internship opportunity on the following terms and conditions.

JOB TITLE: ENUMERATOR

1.0 TERMS OF CONTRACT
This is a maximum 30-day contract and will be effective from 24 January to 22 February 2015. You will be released from this contract as soon as you receive approval from your supervisor upon the completion of your assignment.

2.0 WORKING SCHEDULE
You will be required to work from Monday to Sunday for up to 30 continuous days. If you complete the enumeration of your assigned catchment area before 22 February 2015 (30th day after 27 January 2015), you can report to your supervisor and terminate your contract earlier. Your official working hours are from 07:30am to 04:30pm, but you are strongly required to manage flexible working hours. You may work even before 07:30am or after 04:30pm, whenever it is necessary.

3.0 PERFORMANCE EVALUATION
Your work performance will be evaluated in terms of speed and accuracy of your enumeration. Therefore, quick and accurate enumeration is strongly encouraged. Also, if you complete the assigned enumeration work before the end of the contract, it will be highly appreciated.

Note that, after the census is completed, the supervisors will re-enumerate every catchment area again and evaluate each one's enumeration in terms of accuracy and your attitude toward household members you interviewed.

Hence, your performance will affect your recommendation letter and a future job opportunity at the AFF.

4.0 SALARY
You will receive 10,000 MK as your wage for this contract.
2,000 MK will be provided at the beginning of the project, and the rest will be given upon the completion of the enumeration work. You will be expected to enumerate a minimum of 160 households during the contract period, averaging 8 households per day. When you enumerate more than 160 households, you will be given an additional financial incentive of 500 MK per 8 households. You shall be provided accommodation in your catchment area during the contract period.

5.0 INTERNSHIP PROVISIONS
Upon the successful completion of this contract, you will be given **an official certificate**, which certifies that you worked as a census enumerator for the Africa Future Foundation (AFF) project, and **a recommendation letter** from the director of Project Chimutu (Mr. Hanyoun So) and the chief of TA Chimutu upon your request for your future job applications.

The recommendation letter will specify your relative performance of the enumerator work compared to your peers. In other words, if you do a good job, the recommendation letter will say so, whereas, if you do a bad job, the recommendation letter will say so.

In addition, upon the successful completion of this contract, you will be considered for the **future hire** of a regular staff position at the AFF office if you show outstanding performance satisfying the standard of the management of AFF.

6.0 TERMINATION OF EMPLOYMENT
In the event of any violation by employee of any of the terms of this contract. Employer may terminate employment without notice and with compensation to employee only to the date of such termination.

I am looking forward to a cordial and mutual relationship.
Yours faithfully,

Project Director
Hanyoun So

I,………………………………………………………………, have read and understood the above basic terms and conditions of service of the best of my knowledge and hereby accept the offer as stipulated therein.

Signature:…………………………………………………………………………………………………

C3. G4 (wage only)

AFRICA FUTURE FOUNDATION

PROJECT CHIMUTU

26 January 2015   Mr _____

TEMPORARY EMPLOYMENT CONTRACT

The Management of Project Chimutu has the pleasure to offer you a temporary employment opportunity on the following terms and conditions.

JOB TITLE: ENUMERATOR

1.0 TERMS OF CONTRACT
This is a maximum 30-day contract and will be effective from 27 January to 25 February 2015. You will be released from this contract as soon as you receive approval from your supervisor upon the completion of your assignment.

2.0 WORK SCHEDULE
You will be required to work from Monday to Sunday for up to 30 continuous days. If you complete the enumeration of your assigned catchment area before 25 February 2015 (30th day after 27 January 2015), you can report to your supervisor and terminate your contract earlier. Your official working hours are from 07:30am to 04:30pm, but you are strongly required to manage flexible working hours. You may work even before 07:30am or after 04:30pm, whenever it is necessary.

3.0 PERFORMANCE EVALUATION
Your work performance will be evaluated in terms of speed and accuracy of your enumeration. Therefore, quick and accurate enumeration is strongly encouraged. Also, if you complete the assigned enumeration work before the end of the contract, it will be highly appreciated.

Note that, after the census is completed, the supervisors will re-enumerate every catchment area again and evaluate each one's enumeration in terms of accuracy and your attitude toward household members you interviewed.

Hence, your performance will affect your recommendation letter and a future job opportunity at the AFF.

4.0 SALARY
You will receive 10,000 MK as your wage for this contract.
2,000 MK will be provided at the beginning of the project, and the rest will be given upon the completion of the enumeration work. You will be expected to enumerate a minimum of 160 households during the contract period, averaging 8 households per day. When you enumerate more than 160 households, you will be given an additional financial incentive of 500 MK per 8 households. You shall be provided accommodation in your catchment area during the contract period.

5.0 TERMINATION OF EMPLOYMENT
In the event of any violation of any of the terms of this contract by you, the Management of Project Chimutu may terminate employment without notice and compensation.

I am looking forward to a cordial and mutual relationship.
Yours faithfully,

Project Director
Hanyoun So

I,……………………………………………………………., have read and understood the above basic terms and conditions of service and hereby accept the offer as stipulated therein.

Signature:…………………………………………………………………………………………………

D. Measurement of Cognitive Ability and Non-cognitive Traits

In this appendix, we explain our measures of cognitive ability and non-cognitive personality traits. The measures we report here are all collected during the baseline survey in December, 2014.

D.1. Cognitive ability

D.1.1. Raven's Progressive Matrices test

It is a widely used non-verbal test that evaluates "observation skills and clear-thinking ability" (Raven et al., 1998). Since it is independent of language skills, it is very easy to conduct in any setting including developing countries where the mother tongue is not English. The following is one example of the test questionnaire. In the question, a test subject is required to choose one of eight options that would match a missing pattern in the box. All questions follow the same pattern in that there is a missing component in the visual patterns.
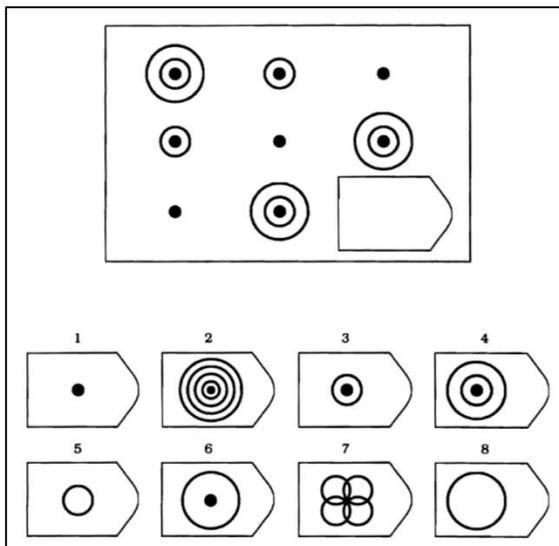
D.1.2. O*NET Ability Profiler

The O*NET Ability Profiler (AP) is originally developed by the US Department of Labor as a "a career exploration tool to help understand job seekers on their work skills" (O*NET Resource Center, 2010, p. 1). We use verbal and clerical tests of the AP relevant for the enumerator job.

a. Verbal ability test measures how a test subject understands the definition of English words and properly uses them in conversation. Basically, it is a vocabulary test. The following is an example of the test questionnaire.

> Choose the two words that are either most closely the same or most closely opposite in meaning

> 1. A. push
>   B. dine
>   C. nap
>   D. eat

b. Clerical perception test: a test to measure an individual's "ability to see details in written materials quickly and correctly. It involves noticing if there are mistakes in the text and numbers, or if there are careless errors in working math problems. Many industrial occupations call for clerical perception even when the job does not require reading or math. This ability is measured by the Name Comparison exercise." (O*NET Resource Center, 2010, p. 2) The following is an example of the test questionnaire.

> On the line in the middle, write <u>S</u> if the two names are exactly the same and write <u>D</u> if they are different.

> | | | | |
> |---|---|---|---|
> | 1. | Paramore & Co. | — | Paramore & Co. |
> | 2. | Bimler | — | Binler |
> | 3. | E-Z Neon | — | E-Z Neon |
> | 4. | Blackstone | — | Blackstone |
> | 5. | Chris Brasch | — | Chris Grasch |

More details can be found at http://informationanthology.net/CareerMentor/Discovery/Ability-Profiler/FAQ-Ability-Profiler.html

D.1.3. Math and English scores of Malawi School Leaving Certificate Exam in 2014

All secondary school students in Malawi are required to take the Malawi School Leaving Certificate Exam in the third semester in Form 4 of secondary school (Grade 12 in the US) in order to achieve an official secondary school graduation status. The Malawi National Examination Board (MANEB) administers the whole process of the exam. Each student chooses

6–8 subjects out of about 20 subjects prepared by MANEB (MANEB, 2014). Mathematics and English are mandatory subjects. The results of each subject are reported in terms of a scale from 1 to 9. We use English and Math test scores because they are mandatory subjects so there are no missing values in the exam transcripts. We obtained the administrative record of the MSCE exam transcripts for all study participants through the Malawi Ministry of Education.

D.2. Non-cognitive traits

D.2.1. Rosenberg self-esteem scale

It is a 10-item scale developed by Rosenberg (1965) and is widely used to measure self-esteem by measuring positive and negative feelings about the self. All items are answered using a 4-point Likert scale format ranging from strongly agree to strongly disagree. It consists of following items.

1. On the whole, I am satisfied with myself.

2. At times I think I am no good at all.

3. I feel that I have a number of good qualities.

4. I am able to do things as well as most other people.

5. I feel I do not have much to be proud of.

6. I certainly feel useless at times.

7. I feel that I'm a person of worth, at least on an equal plane with others.

8. I wish I could have more respect for myself.

9. All in all, I am inclined to feel that I am a failure.

10. I take a positive attitude toward myself.

D.2.2. Intrinsic Motivation

Intrinsic motivation is an individual's trait that captures whether an individual is motivated to do things by intrinsic rewards such as an individual's own desire to pursue goals or challenges. It is the opposite of extrinsic motivation described below. We measure intrinsic motivation using a 15-item scale (Amabile et al., 1994). All items are answered using a 4-point Likert scale format ranging from strongly agree to strongly disagree. It consists of following items:

1. I enjoy trying to solve difficult problems.

2. I enjoy simple, straightforward tasks.

3. I enjoy tackling problems that are completely new to me.

4. What matters most to me is enjoying what I do.

5. It is important for me to be able to do What I most enjoy.

6. The more difficult the problem, the more I enjoy trying to solve it.

7. I want my work to provide me with opportunities for increasing my knowledge and skills.

8. I like to figure things out for myself.

9. No matter what the outcome of a project, I am satisfied if I feel I gained a new experience.

10. Wanting to know more is the driving force behind much of what I do.

11. I prefer work I know I can do well over work that goes beyond what I can manage.

12. I am more comfortable when I can set my own goals.

13. I enjoy doing work that is so involving that I forget about everything else.

14. It is important for me to have space to express myself.

15. I want to find out how good I really can be at my work.

D.2.3. Extrinsic Motivation

Extrinsic motivation is an individual's trait that captures whether an individual is motivated by external rewards such as reputation to do things. We use a 15-item scale to measure a level of motivation triggered by extrinsic values. All items are answered using a 4-point Likert scale format ranging from strongly agree to strongly disagree (Amabile et al., 1994). It consists of following items:

1. I am not that concerned about what other people think of my work.

2. I prefer having someone set clear goals for me in my work.

3. I am very much aware of the income goals I have for myself.

4. To me, success means doing better than other people.

5. I am very much aware of the career promotion goals I have for myself.

6. I am less concerned with what work I do than what I get for it.

7. I am concerned about how other people are going to react to my ideas.

8. I rarely think about salary and promotions.

9. I believe that there is no point in doing a good job if nobody else knows about it.

10. I am strongly motivated by the money I can earn.

11. I prefer working on projects with clearly specified procedures.

12. As long as I can do what I enjoy, I am not that concerned about exactly what I am paid.

13. I am strongly motivated by the recognition I can earn from other people.

14. I have to feel that I am earning something for what I do.

15. I want other people to find out how good I really can be at my work.

D.2.4. Ten-item Big Five personality inventory (TIPI)

We measure an individual's personality types using a 10-item scale that assesses the respondent's characteristics based on traits defined by *Five Factor Theory of Personality*: openness to experience, conscientiousness, extroversion, agreeableness, and emotional stability (Gosling et al., 2003). All items are answered using a 7-point scale format; Disagree strongly, Disagree moderately, Disagree a little, Neither agree nor disagree, Agree a little, Agree moderately, and Agree strongly. It consists of following items:

I see myself as:
1. Extraverted, enthusiastic.
2. Critical, quarrelsome.
3. Dependable, self-disciplined.
4. Anxious, easily upset.
5. Open to new experience, complex.
6. Reserved, quiet
7. Sympathetic, warm.
8. Disorganized, careless.
9. Calm, emotionally stable.
10. Conventional, uncreative.

D.3. Parental Support

We ask how strongly
My parents are interested in the details of my life.
I consider my parents friends.
My parents support decisions I make.
My parents make a point of letting me know that they are available to me whenever I need them.
My parents encourage me to obtain higher education.
My parents view me as an equal.
My parents value my thoughts.
My parents value my opinions.

E. Survey accuracy

We checked each questionnaire one by one and collected systematically inconsistent errors. First, project supervisors listed all possible systematic errors that could come from enuemrators. Second, data-entry clerks go through repeated training to catch those errors. Then, they started counting the number of systematic errors caused by enumerators for each sheet of the census survey.

Error collecting work was carried out in the following steps:

1. Two error-collecting data entry clerks checked one questionnaire separately.

2. Count four types of errors from each page of the questionnaire, as follows.

    1) The total number of all questions that must be answered.

    2) The total number of questions that must be answered but are blank.

    3) The total number of questions that must be answered but are wrongly answered.

    4) The total number of questions that must not be answered but are answered.

3. Add up all the numbers on each page and record the total number of errors

4. Compare the total number of errors independently counted by two clerks.

5. Count again if the difference between the total errors counted by two data entry clerks is larger than 5.

6. Record the mean of the number of errors counted by two data entry clerks.

The following table provides the basic statistics of each number counted.

Index	Measurement	Mean (SD)
A	The total number of all questions that must be answered	221.7 (61.8)
B	The total number of questions that must be answered but are blank	7.59 (10.3)
C	The total number of questions that must be answered but are wrongly answered	3.90 (4.26)
D	The total number of questions that must not be answered but are answered	5.53 (9.28)

Note: A could be different across households due to differences in household-specific characteristics such as family structure.

Finally, the final outcome variable we use for survey accuracy in the analysis is as follows:

$$\text{Survey error}_i = (B_i + C_i + D_i)/A_i$$

where *Survey error$_i$* is the final survey error rate outcome for census household i . That is, i corresponds to one census questionnaire set surveyed by an enumerator. A_i, B_i, C_i, and D_i are the corresponding numbers detected from the i-th census survey questionnaire.

F. Cross-correlations between Cognitive and Non-cognitive Traits

Table F.1. Baseline survey participants (n=443)

	Self-esteem	Intrinsic motivation	Extrinsic motivation	Extroversion	Agreeableness	Conscientiousness	Emotional stability	Openness to experience	Raven's matrix test	Verbal ability	Clerical ability	MSCE score	Cognitive ability index
Self-esteem	1.00												
Intrinsic motivation	.221***	1.00											
Extrinsic motivation	.092*	.290***	1.00										
Extroversion	.050	.027	-.028	1.00									
Agreeableness	.126***	.030	-.047	-.140***	1.00								
Conscientiousness	.100**	.035	.029	-.033	.263***	1.00							
Emotional stability	.170***	.099**	-.094**	.015	.339***	.266***	1.00						
Openness to experiences	.063	.140***	-.045	-.069	.336***	.392***	.312***	1.00					
Raven's matrix score	.049	.065	-.079*	.059	.070	.025	.135***	.028	1.00				
Verbal ability	.135***	.098**	-.008	.011	.141***	.085*	.140***	.074	.173***	1.00			
Clerical ability	.052	.052	-.125***	.081*	.044	.078	.129***	.075	.177***	.147***	1.00		
MSCE score	.140***	.161***	-.054	.032	.133***	.094**	.325***	.112**	.320***	.419***	.234***	1.00	
Cognitive ability index	.159***	.157***	-.101**	.074	.153***	.113**	.304***	.125***	.588***	.638***	.547***	.832***	1.00

***, **, and * denote significance at 1%, 5%, and 10%, respectively.

Table F.2. Enumerator (n=137)

	Self-esteem	Intrinsic motivation	Extrinsic motivation	Extroversion	Agree-ableness	Conscien-tiousness	Emotional stability	Openness to experiences	Raven's matrix test	Verbal ability	Clerical ability	MSCE 2014	Cognitive ability index	Error rate	Surveys per day	Attitude scale
Self-esteem	1.00															
Intrinsic motivation	.158*	1.00														
Extrinsic motivation	-.026	.186**	1.00													
Extroversion	.115	.084	.083	1.00												
Agreeableness	.132	-.016	-.091	-.083	1.00											
Conscientiousness	.088	-.019	-.009	-.167*	.221***	1.00										
Emotional stability	.251***	.164*	-.089	-.069	.362***	.278***	1.00									
Openness to experiences	-.063	.098	-.142	-.062	.234***	.465***	.393***	1.00								
Raven score	.099	.110	.139	.056	.094	-.023	.102	.105	1.00							
Verbal ability	.183**	.134	.083	-.019	.198**	.081	.143*	.135	.180**	1.00						
Clerical ability	-.024	-.026	-.176**	-.021	.187**	.135	.112	.032	.187**	.127	1.00					
MSCE 2014	.246***	.196**	.133	-.024	.224***	.139	.415***	.255***	.380***	.435***	.032	1.00				
Cognitive ability index	.221***	.177**	.063	-.006	.280***	.146*	.338***	.218**	.634***	.662***	.477***	.787***	1.00			
Error rate	-.165*	.022	.215**	-.075	-.088	-.003	-.152*	-.180**	-.279***	-.158*	-.137	-.256***	-.326***	1.00		
Surveys per day	-.045	.022	.117	-.030	-.042	.106	-.048	.110	-.042	-.094	.089	.015	-.003	-.010	1.00	
Attitude scale	.025	-.044	-.040	.166*	.040	-.037	.057	-.024	-.008	-.119	.056	-.060	-.051	-.063	-.057	1.00

***, **, and * denote significance at 1%, 5%, and 10%, respectively.

12

G. Mean differences of individual characteristics by first-stage randomization group (after self-selection)

Table G.1. Mean differences by first-stage randomization group after self-selection

Variable	Obs	Internship Job-takers	Wage Job-takers	Mean Difference
		(1)	(2)	(3)	(4) =(2)-(3)
Panel A. Demographics and socio-economic status				
Age	362	20.8	20.7	.162
Height	360	165.0	164.7	.368
BMI	360	19.9	19.5	.413
Asset score	362	.932	1.05	-.122
Number of siblings	362	4.86	4.46	.405
Level of parental support	362	15.7	15.3	.369
Currently working	362	.081	.054	.027
Panel B. Non-cognitive traits and cognitive ability				
Self-esteem (Rosenberg scale)	362	19.1	18.6	.521
Intrinsic motivation	362	3.05	3.08	-.029
Extrinsic motivation	361	2.78	2.83	-.046
Extroversion	358	3.67	3.27	.405**
Agreeableness	361	5.08	5.10	-.019
Conscientiousness	361	5.67	5.87	-.196
Emotional stability	360	4.94	5.12	-.182
Openness to experiences	360	5.35	5.52	-.171
Cognitive Ability Index	362	-.199	-.077	-.122
Number of observations		74	74	148

Note: ***, **, and * denote significance level at 1%, 5%, and 10%, respectively

H. Job performance before and after supervisor visits

As stated in Subsection 3.2.5, we dispatched supervisors to monitor and advise enumerators in the field. During the field visit, a supervisor accompanied an enumerator to about three households on average per visit, pointed out common errors the enumerator was making, and provided overall comments about the enumerator's job performance.

Figure H.1 below illustrates changes in error rate and survey speed over time. A vertical line in Panels A and B represents the first and second supervisor visit dates, respectively. It shows that the error rate decreases over time while the survey speed remains relatively the same during the survey period. There is a dip in speed on the day of the first visit of a supervisor probably due to intense supervision during the first visit, but it rebounds after the departure of the supervisor.

Table H.1 shows the estimates of the following equation:

$$Performance_{ijt} = \alpha + \beta First_i + \beta Second_i + \sum_{a=1}^{3} \rho_a I(G_i = a) + \delta X_i + \sigma_t + \emptyset Z_j + \mu_{ijt} \tag{6}$$

where $First_i$ is a dummy variable for the supervisor's first random visit to enumerator i (1 if surveyed from date of the first visit onward, 0 otherwise). $Second_i$ is a dummy variable for the supervisor's second random visit to enumerator i (1 if surveyed from date of the second visit onward, 0 otherwise). $I(G_i = a)$ denotes a dummy variable for whether enumerator i belongs to Group a. G4 (the wage-only group) is the excluded group. X_i is a vector of individual characteristic variables, which combines $Demog_i$, Cog_i, and $NonCog_i$ in Section 5. μ_{ijt} is an *i.i.d.* error term.

Figure H.1 suggests that the first supervisor visit is followed by a reduction in the error rate, yet, the regression evidence in Table H.1 shows that the error-reducing effect disappears once we control for a variety of confounding factors. Rather, the supervisor visit reduces the survey speed, and thus, it actually reduces labor productivity. The reason there is no supervisor effect on the survey error rate might be that it is relatively difficult to improve by advice from a supervisor.

Table H.1. Supervision visit regression

| Dependent variable	Error rate		Survey speed	
	(1)	(2)	(3)	(4)
First visit	.004	.002	-2.38***	-2.45***
(1 if surveyed from date of the first visit onward, 0 otherwise)	(.012)	(.012)	(.775)	(.793)
Second visit	.002	.012	1.15	.911
(1 if surveyed from date of the second visit onward, 0 otherwise)	(.012)	(.010)	(1.16)	(1.25)
Observations	21,560	21,560	1,977	1,977
R-squared	.057	.101	.098	.111
Enumerator characteristics	X	O	X	O
Mean (SD) of the dependent variable	.076 (.071)		10.9 (5.70)	

Note: Standard errors clustered at the enumerator level are reported in parentheses. Catchment area characteristics include the total number of households, total population, asset score (refrigerator, bicycle, and improved toilet), birth rate in the last 3 years, incidence of malaria among under 3-year-olds, proportion of under 3-year-olds born with the assistance of health professionals, and deaths in the last 12 months. Control variable set of enumerator characteristics includes age, number of siblings, level of parental support, and asset score (improved toilet, refrigerator, and bicycle). ***, **, and * denote significance at 1%, 5%, and 10%, respectively. All specifications include catchment area controls, fixed effects for incentive groups, and work day.

Figure H.1.

Panel A: Error rate before and after the first random visit
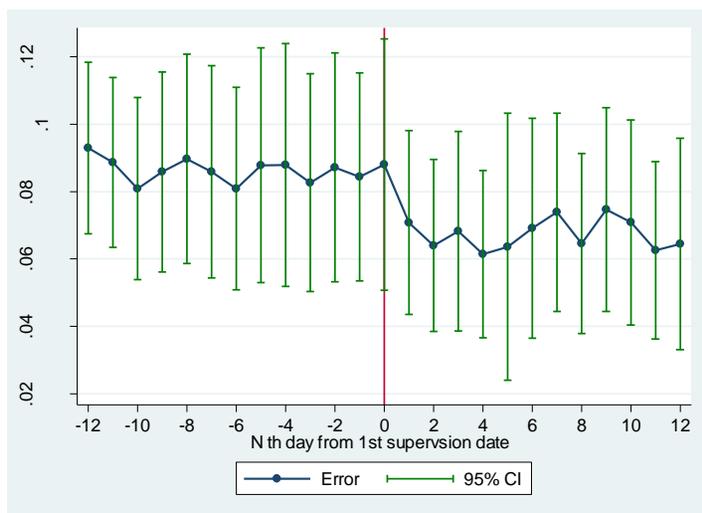

Panel B: Error rate before and after the second random visit
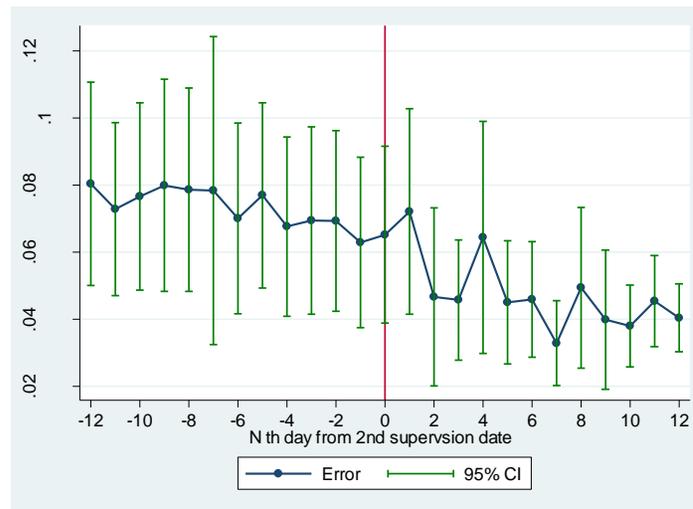
Panel C: Survey speed before and after the first random visit
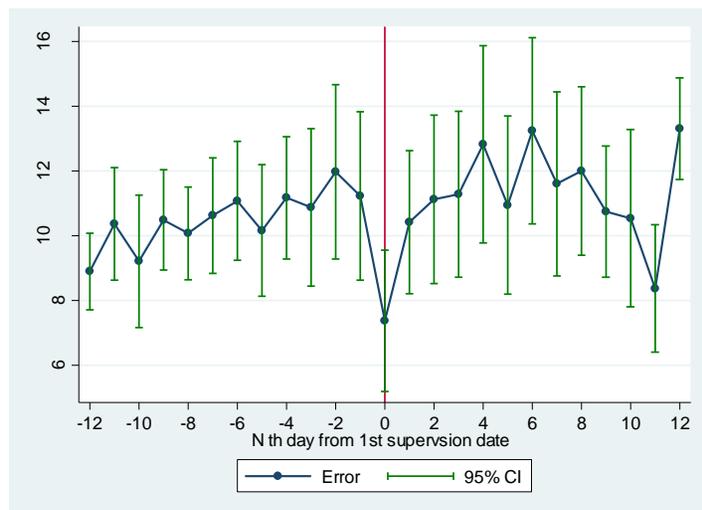
Panel D: Survey speed before and after the second random visit
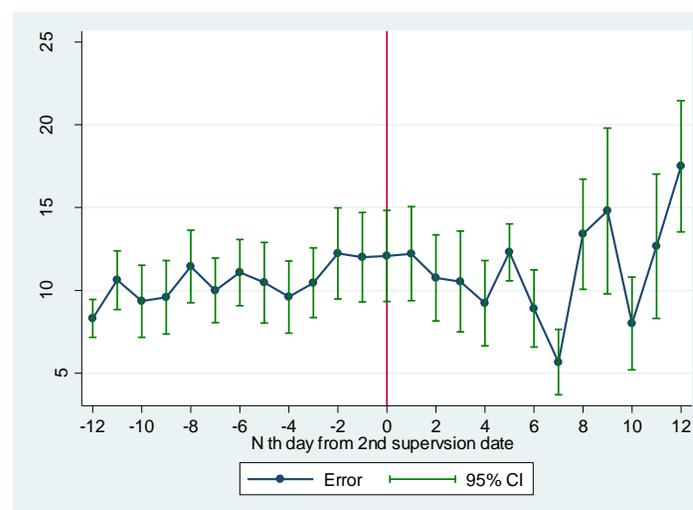
References for the Appendix

Amabile, T.M., K.G. Hill, B.A. Hennessey, and E.M. Tighe. 1994. "The Work Preference Inventory: Assessing Intrinsic and Extrinsic Motivational Orientations." *Journal of Personality and Social Psychology*, 66(5): 950.

Gosling, S.D., P.J. Rentfrow, and W.B. Swann. 2003. "A Very Brief Measure of the Big-Five Personality Domains." *Journal of Research in Personality*, 37(6): 504–528.

Malawi National Examination Board (MANEB). 2014. Malawi School Certificate of Education Examination—Grades and Awards for Candidates, Malawi.

O*NET Resource Center. 2010. *O*NET Ability Profiler^{TM} Score Report*, pp. 1–2

Raven, J., J.C. Raven, and J.H. Court. 1998. Manual for Raven's Progressive Matrices and Vocabulary Scales. The Standard Progressive Matrices, Section 3. Oxford Psychologists Press: Oxford, England/The Psychological Corporation: San Antonio, TX.

Rosenberg, Morris. 1965. Society and the adolescent self-image. Princeton, NJ: Princeton University Press.