

# **Job Fairs and Firm Recruitment in Addis Ababa: Pre-Analysis Plan**

Drafted by: Girum Abebe, Stefano Caria, Marcel Fafchamps,  
Paolo Falco, Simon Franklin and Simon Quinn

**Fieldwork location:** Addis Ababa, Ethiopia  
**Fieldwork dates:** March 2014 – December 2015  
**Date of Pre-Analysis Plan:** 15 August 2016

## **1 Introduction**

This document outlines our pre-analysis plan for a job fairs experiment run in Addis Ababa between 2014 and 2015. It is based on the recommendations of McKenzie (28 October 2012, Development Impact Blog) and it summarises our experiment, our data and our planned empirical analysis. We intend to submit this pre-analysis plan to the AEA RCT Registry.

The job fairs experiment described here involved both randomly invited firms and randomly invited job-seekers. The sample of job-seekers and the strategy to estimate the effects of the program on their labour market outcomes is described in the pre-analysis plan of a closely related project (available at

[www.socialscienceregistry.org/trials/911](http://www.socialscienceregistry.org/trials/911)).

This pre-analysis plan concentrates on the firms involved in the experiment and on the impact the job fair had on them. The document is structured as follows. First, we begin by describing the sample of participating firms. Second, we describe the data collected at various stages of the experiment. Finally, we explain how we intend to estimate the effects of interest, the family of outcomes we will focus on and the dimensions along which we will study treatment heterogeneity.

## **2 Firm Sample**

The sample of participating firms comprises 498 of the largest employers in the city of Addis Ababa, drawn from the main sectors of the economy (including construction, manufacturing, banking and financial services, hotels and hospitality, and other professional services). The average number of pay-roll employees in the sample was 171.5, a large figure by Ethiopian standards.

To select participating firms, we first put together the first comprehensive list of the largest firms in Addis Ababa across all sectors, to our knowledge. Since no firm census exists in Ethiopia, we created the list using a variety of different sources, including government records of formal firms held across different ministries. For the manufacturing sector, we used a representative sample of the largest firms from the *Large and Medium Enterprise* surveys, conducted by the Central Statistics Agency (CSA). When the exact size of the firm was known, we imposed a minimum cut-off of 40 workers. In other cases, we requested specific administrative authorities (who manage the firm registries) to provide lists of the largest firms in a given sector. The final list contained 2,178 firms of the largest employers in Addis Ababa.

Table 1: Main Industry Classifications

Industry	Firms in Sample	Total Firms (Weighted)
Tours-Hospitality	92	435
Finance, Services, Retail	102	239
Education, Health, Aid	104	610
Manufacturing	126	301
Construction, Mining, Farming	69	593
Total	493	2,178

A sample of 493 firms was surveyed for the experiment. The proportion of firms drawn from each sector was equal to the share of *employers* in that sector (obtained from representative labour force data). This ensured that our study sample is representative of employment opportunities in Addis Ababa. Table 4 shows the breakdown of selected firms into 5 main categories. Column (2) provides weighted figures obtained by applying the inverse of the weights used to sample firms. For instance, we surveyed NGOs (“Education, Health, Aid”) relatively infrequently because of the large number of NGOs in the data, and their relatively small contribution to total employment in the city.

Selected firms are, on average, very large by Ethiopian and African standards. The mean number of employees per firm is 171.5 workers, but this masks considerable heterogeneity, particularly in the Tours & Hospitality sector, which is dominated by relatively small hotels and restaurants. Average firm size, when this sector is excluded, is 326 workers per firm. These figures are provided in Table 2. The reader should note that these numbers exclude casual daily labourers (firms report having 34 casual labourers in their workforce, on an average day).

The firms in our sample are growing and looking to hire new workers. On average, at the time of the baseline survey, they expected to expand their workforce in the following 12 months by an average of 12%. The types of workers they are looking for are most commonly white collar workers, usually with university degrees.

Table 2: Firm Size by Sector

Industry	Worker Type				All workers
	Client services	Production	Support staff	White collar	
Construction, Mining, Farming	2.7	92.7	21.7	21.8	143.2
Tours-Hospitality	15.8	7.4	13.2	7.4	46.4
Finanace, Services, Retail	146.6	33.7	96.6	183.3	473.3
Education, Health, Aid	12.6	5.2	31.2	73.6	131.0
Manufacturing	24.4	149.0	37.4	33.7	250.2
All Industries	26.9	52.4	33.1	52.8	171.5

Table 3: Median rate of expected number of new hires in the coming 12 months, as a percentage of current workforce

Industry	Worker Type				All workers
	Client services	Production	Support staff	White collar	
Construction, Mining, Farming	0.0%	14.3%	9.2%	15.4%	20.0%
Tours-Hospitality	16.7%	10.8%	10.2%	10.6%	14.8%
Finanace, Services, Retail	10.5%	6.3%	10.1%	16.0%	16.1%
Education, Health, Aid	4.5%	5.7%	5.0%	14.3%	13.0%
Manufacturing	0.0%	8.0%	1.6%	3.4%	8.8%
All Industries	7.4%	9.3%	7.4%	11.1%	12.6%

**Attrition:** Rates of attrition are remarkably low in this sample. Among the 493 firms that were interviewed in the baseline survey, we were able to contact 478 in the follow-up survey. Only 3% of firms (15 in total) could not be re-interviewed: 8 were treated, 7 were in the control group. There are no concerns of differential attrition in the sample.

**Phone surveys:** We conducted mid-line phone surveys with firms before each of the job fairs, and after the second job fairs. Between 4 and 2 weeks before the job fairs took place we phoned firms to ask them about their current set of available vacancies. For those firms attending the fairs, this information was used to match firms with workers, based on observable worker characteristics and the corresponding available vacancies.

In addition, in order to measure short run differences in hiring and recruitment between the treated and the untreated, we phoned all firms one month after the second job fair to ask them about their recent hiring experiences. We asked questions about a roster of their recent vacancies, and about the firms' experiences with filling those vacancies. This data is used to measure the outcomes in Family 1a and Family 1b below.

### **3 Description of Intervention**

Two job fairs were held for this project. The first fair took place on October 25 and 26, 2014. The second fair took place on February 14 and 15, 2015. The same workers and firms were invited to both fairs in each round, but job-seekers were invited on a different day (with different firms) the second time around. In this way both sides of the market had the chance to attend two fairs, learn from the first fair, and meet new candidates on the second date.

A total of 1007 workers were invited to the fairs. These were young people who were unemployed at the time of a baseline survey conducted 4 months before the first fair. While some of these respondents had found work by the time of the fair, only about 8% had permanent jobs. Therefore, most of them were still searching for work. Participating job-seekers were between the ages of 18 and 30, and had at least a high school qualification.

About half of the firms in our sample (245 of them) were randomly selected to be invited to the job fairs. The group of invited firms was split into two, one group coming together on the Saturday, and the other group coming on the Sunday of both fairs. Invited firms were given information about the fairs, their purpose and about how to prepare for them. The firms in our sample were typically hiring around the time when the job fairs were held: 89% of them hired at least one worker in the year between baseline

and endline surveys. The job fairs were held in the middle of that year (on average, they hired 52 workers over the course of that year). The mean rate of hiring (number of hires as a fraction of total workforce at baseline) in the year prior to the experiment was 18%.

The job fairs were held at the Addis Ababa University campus, where firms came to set up their stalls in advance of the job seekers arriving. The fairs were split over two days, with a randomly selected half of firms and workers attending on each day. The job-seekers invited on the second day had previously been invited to take series of cognitive and non-cognitive personnel tests, and they had been provided with certificates reporting their performance.

## **4 Randomization & Balance**

We assigned firms to either the treatment group (invited to the fairs) or the control group (not invited) using block level randomization techniques suggested by [Bruhn and McKenzie \(2009\)](#). In all, we assigned roughly half of the firms to attend the job fairs, half were assigned to the control group and were not invited to the fairs. In all, 245 firms were assigned to the treatment group, 248 to the control group.

The following method was used to group firms together: first, firms were partitioned by five main groups of industries, defined in table 4. Then, then firms were partitioned into nearest neighbour groups of 4 firms on the basis of Mahalanobis distance defined over the set of "blocking" variables listed in Table 5.

After that, we randomly split the four firms in each group into treatment and control: two of them (treated) were invited to the job fairs, one on each (randomly chosen) day. The other two were assigned to the control group and were not invited to the fairs.

Table 4: Main Industry Classifications

Main Industry	Freq.	Percent
Tours-Hospitality	92	18.66
Finanace, Services, Retail	102	20.69
Education, Health, Aid	104	21.1
Manufacturing	126	25.56
Construction, Mining, Farming	69	14
Total	493	100

Table 5: **Blocking variables**

VARIABLE	DEFINITION	SOURCE (QUESTION NUMBER)
plc	Firm is a private limited company	$g3 = 3$
total_n_all	Total number of pay-roll employees at the firm	$l1\_1\_n$
prop_p	Proportion of workers who are professionals	$l1\_5\_n/l1\_1\_n$
ed_deg	Number of workers at the firm with a degree	$rowtotal(l1\_19\_ * \_1)/rowtotal(ed * \_total)$
to_all	Rate of turnover in the last year	$rowtotal(l2\_1\_*)/total\_n\_all$
formal_adv	Firms advertise when recruiting for jobs	$l4\_2\_1=1$ or $l4\_2\_2=1$
fairs	Firms expressed interest in attending a job fair	$l4\_31$
hire_all	Rate of new hiring in the last year	$rowtotal(l3\_2\_*)/total\_n\_all$

Additionally, we assigned treatment using a re-randomisation method. We iterated over possible randomization assignments until our randomized a set rule in terms of balance across a set of variables. Following the recommendations of Bruhn and McKenzie (2009), we will control in our estimations for the baseline covariates used for the re-randomisation (that is, the set of variables described in Table 6) and for the baseline covariates used to construct the randomisation blocks.<sup>1</sup>

<sup>1</sup> Most of these are included in the list of Table 6. The two variables used for blocking that are not included in that list are: distance from the city centre and total number of individuals surveyed in the cluster.

Table 6: **Re-randomization Variables**

plc	Firm is a private limited company	$g3 = 3$
ngo	Firm is a registered NGO	$g3 == 6$
Industry	Various categories (see Table 4)	$s22$
stad_dist	Distance of HQ from the centre of city	$GPScoordinates$
total_n_all	Total number of pay-roll employees at the firm	$l1\_1\_n$
prop_p	Proportion of workers who are professionals	$l1\_5\_n/l1\_1\_n$
prop_s	Proportion of workers who are customer service workers	$l1\_5\_s/l1\_1\_n$
prop_w	Proportion of workers who are production workers	$l1\_5\_w/l1\_1\_n$
prop_c	Proportion of workers who are customer service workers	$l1\_5\_n/l1\_1\_n$
ed_deg	Number of workers at the firm with a degree	$rowtotal(l1\_19\_ * \_1)/rowtotal(ed * \_total)$
ed_dip	Number of workers at the firm with a diploma	$rowtotal(l1\_19\_ * \_2)/rowtotal(ed * \_total)$
to_all	Rate of turnover in the last year	$rowtotal(l2\_1\_*)/total\_n\_all$
hire_all_total	Total number of hires in the last year	$rowtotal(l3\_2\_*)$
hire_all	Rate of new hiring in the last year	$rowtotal(l3\_2\_*)/total\_n\_all$
formal_adv	Firms advertise when recruiting for jobs	$l4\_2\_1 = 1 \text{ or } l4\_2\_2 = 1$
fairs	Firms expressed interest in attending a job fair	$l4\_31$
total_sales_n	Total firm sales (revenue) for the year	$p1\_1\_1$
av_sal	Average salary per worker	Total wage bill/total employees - $l1\_15\_ * l1\_17\_ * l1\_18\_*$
hire_exp	Rate of new hiring expected in the next year	$rowtotal(l3\_12\_*)/total\_n\_all$

Table 7: **Summary for Blocking/Re-Randomisation Variables**

N	Mean	S.Dev.	1st Q.	Median	3rd Q.	Min.	Max.	P_val (F)
---	------	--------	--------	--------	--------	------	------	-----------

*Pre-Analysis Plan: Job Fairs and Firm Recruitment in Addis Ababa:*

plc	493	0.51	0.50	0.00	1.00	1.00	0.0	1.0	0.963
ngo	493	0.13	0.34	0.00	0.00	0.00	0.0	1.0	0.958
Tours...	493	0.19	0.39	0.00	0.00	0.00	0.0	1.0	0.949
Finanace...	493	0.21	0.41	0.00	0.00	0.00	0.0	1.0	0.878
Education...	493	0.21	0.41	0.00	0.00	0.00	0.0	1.0	0.944
Manufacturing...	493	0.26	0.44	0.00	0.00	1.00	0.0	1.0	0.937
Construction...	493	0.14	0.35	0.00	0.00	0.00	0.0	1.0	0.940
stad_dist	491	4.93	8.85	1.96	3.42	5.80	0.2	123.6	0.886
total_n_all	493	288.11	972.98	37.00	87.00	225.00	4.0	18524.0	0.598
prop_p	493	0.29	0.23	0.10	0.21	0.45	0.0	0.9	0.921
prop_s	493	0.24	0.15	0.13	0.22	0.32	0.0	0.8	0.401
prop_w	493	0.26	0.29	0.00	0.17	0.50	0.0	1.0	0.863
prop_c	493	0.14	0.16	0.00	0.07	0.22	0.0	0.7	0.873
ed_deg	493	0.23	0.24	0.04	0.13	0.37	0.0	1.0	0.901
ed_dip	493	0.17	0.15	0.05	0.13	0.24	0.0	1.0	0.519
to_all	493	0.21	0.88	0.05	0.10	0.19	0.0	14.3	0.150
hire_all_total	493	54.45	218.42	4.00	11.00	35.00	0.0	3901.0	0.268
hire_all	493	54.45	218.42	4.00	11.00	35.00	0.0	3901.0	0.268
formal_adv	493	0.65	0.48	0.00	1.00	1.00	0.0	1.0	0.703
fairs	493	0.79	0.41	1.00	1.00	1.00	0.0	1.0	0.711
total_sales_n	339	554756.78	3.84e+06	7175.00	23017.00	121831.00	0.0	6.0e+07	0.492
av_sal	493	2885.07	3010.35	1303.03	1990.18	3190.00	0.0	27683.2	0.812
hire_exp	493	0.22	0.85	0.00	0.08	0.19	0.0	14.9	0.571

## 5 Empirical specification

Our empirical analysis will follow a standard approach. Half of the firms were randomly invited to the job fairs. For each outcome of interest we run a regression of that outcome on a dummy variable equal to 1 if the firm was invited, the lagged dependent variable, and a set of covariates used for the randomization. We do not cluster our standard errors since randomization was conducted at the firm level. Firms are spread across the city and might reasonably be thought to operate in the same large labour market.

We estimate:

$$y_i = \beta_0 + \beta_1 \cdot \text{fairs}_i + \alpha \cdot y_{i,pre} + \delta \cdot \mathbf{x}_{i0} + \mu_i \quad (1)$$

```
ivreg2 y fairs
```

```
y_pre balance_*, partial(balance_*) robust
```

In this specification, the “balance” variables included in  $\mathbf{x}_{i0}$  are all the variables listed in Table 7. Variable  $y_{i,pre}$  is simply the dependent variable measured at baseline.

## 6 Families of outcomes

To test the effect of our treatments on employment, we will construct a family of outcome variables. For each of these outcomes, we will run the estimation and hypothesis tests outlined earlier in section ???. For each hypothesis test, we will report two values:

- (i). The usual  $p$ -value from a Wald test; and
- (ii). We will report False Discovery Rate  $q$ -values, taken across the family of outcomes (Benjamini, Krieger, and Yekutieli, 2006). (That is, for each type of test, we will construct a  $q$ -value for that test

across outcomes. We will construct a set of  $q$ -values using all  $p$ -values for the null hypothesis ‘The job fairs had no effect’. )

Following [Olken \(2015\)](#), we separate our outcomes into primary families and secondary families of outcomes, expecting that our interventions are most likely to have impacts on outcomes related to hiring and recruitment methods, with knock-on effects to other aspects of firm outcomes assigned to families of secondary outcomes.

## 6.1 Primary Outcomes: Hiring and recruitment methods

Table 8: **Family 1a: Hiring after the job fairs**

VARIABLE	DEFINITION	SOURCE (QUESTION NUMBER)
short_vacancies	Number of vacancies opened (count h1)	count h1! = .
short_interviews	Total interviews conducted	sum h2
short_hires	Number of people hired- all vacancies	sum h3
hiring_shortfall	Number of people shortfall=1 on hiring (as percentage of number wanted to hire)- all vacancies	sum (h3-h4)/h3
remaining_vacancies	Number of positions that the firm still has vacant (in total)	sum h11

Table 9: **Family 1b: Hire quality after job fairs**

VARIABLE	DEFINITION	SOURCE (QUESTION NUMBER)
short_permanent	Hired workers on permanent contract	h5

*Pre-Analysis Plan: Job Fairs and Firm Recruitment in Addis Ababa:*

short_daysrecruit	Days taken to recruit for positions (weighted mean across all vacancies)	mean h6
short_salary	Average starting salary (weighted mean across all vacancies).	mean h9
short_degree	Hired new recruits with University degrees.	h7

**Table 10: Family 2A: Recruitment in the last year: Recruitment**

VARIABLE	DEFINITION	SOURCE (QUESTION NUMBER)
time_fill_pro	Time taken to fill professional vacancies	13_4_p
time_fill_nonpro	Time taken to fill non-professional vacancies	weighted ave 13_4_c & 13_4_w
num_ints_pro	Number of interviews per position. (professional)	14_9new
pay_recruit_pro	How much did the firm pay per recruit ,professional	14_14
pay_recruit_nonpro	How much did the firm pay per recruit , /non-professional)	14_14new
unfilled_vacancies	How many vacancies were you unable to fill as percentage of vacancies opened.	14_7.

**Table 11: Family 2B: Recruitment in the last year: worker characteristics**

VARIABLE	DEFINITION	SOURCE (QUESTION NUMBER)
new_hire_pro	Number of new hires for the year professional	13_2_p
new_hire_nonpro	Number of new hires for the year non-professional	13_2_c 13_2_w
hire_pro_degrees	New recruits by education: did firms mostly hire people with degrees (professional positions)	13_5
new_hire_perm_nonpro	Percentage of new hires hired in permanent positions in last 12 months - non- professional	(13_11_c +13_11_w)/(13_2_c +13_2_w)
new_hire_perm_pro	Percentage of new hires hired in permanent positions in last 12 months - professional	13_11_p/13_2_p

**Table 12: Family 3: Recruitment methods**

VARIABLE	DEFINITION	SOURCE (QUESTION NUMBER)
interviews_pro	Perform formal interviews (professionals)	13_10_p
interviews_nonpro	Perform formal interviews (non-professionals)	13_10_c 13_10_w
advertised	Did any advertising for new hires	14_1
advertised_pro	Did advertising for professional positions	14_4
advertised_board	Did advertising on the job boards?	14_2_1

## **6.2 Secondary Outcomes: Impact on aggregate firm level outcomes**

**Table 13: Family 4: Composition of the workforce**

VARIABLE	DEFINITION	SOURCE (QUESTION NUMBER)
total_n_all	Total number of employees	11_1_n
perm_prop_pro	Proportion of professional workers on permanent contracts?	11_20_p/11_5
perm_prop_nonpro	Proportion of non-professional workers on permanent contracts?	$(11\_20\_c+11\_20\_w)/(11\_7+11\_9)$
start_salary_pro	Average starting salary- professionals	11_16_p
start_salary_nonpro	Average starting salary- non-professionals	11_16_w 11_16_c
ed_deg_pro	Proportion of professional workers with degree	11_19_p_1/ rowtotal(11_19_p_*)
ed_deg_nonpro	Proportion of workers with post-secondary education (non-professionals)	11_19_w_1 11_19_w_2 11_19_c_1 11_19_c_2
match_quality	Average worker is not underqualified in any of the worker categories	11_23_*! = 3

**Table 14: Family 5: Turnover/higher retention overall**

VARIABLE	DEFINITION	SOURCE (QUESTION NUMBER)
fire_rate_pro	Professional firing rate	12_3_p /total_n_p
fire_rate_nonpro	Non-professional firing rate	$(12\_3\_c + 12\_3\_w) /total\_n\_nonp$
quit_rate_pro	Professional quit rate	12_4_p /total_n_p
quit_rate_nonpro1	Non-professional quit rate	$(12\_4\_c + 12\_4\_w) /total\_n\_nonp$

*Pre-Analysis Plan: Job Fairs and Firm Recruitment in Addis Ababa:*

growth_rate_all	Growth rate in employment	$(\text{hire\_total\_all} - \text{to\_total\_all}) / (\text{total\_n\_all} - (\text{hire\_total\_all} - \text{to\_total\_all}))$
growth_rate_pro	Growth rate in professional employment	$(\text{l3\_2\_p} - \text{l2\_1\_p}) / (\text{total\_n\_p} - (\text{l3\_2\_p} - \text{l2\_1\_p}))$

**Table 15: Family 6: Firm growth and productivity**

VARIABLE	DEFINITION	SOURCE (QUESTION NUMBER)
for_profit	The firm is for profit	p0
sales_revenue2007	Total sales revenue in 2007 EC	p1_1_1
value_added2007	Value added in 2007 EC	$(\text{p1\_1\_1} - \text{p1\_4\_1} - \text{p1\_5\_1})$
reported_profit2007	Profit reported by the firm for 2007 EC	p1_10_1
calculated_profit2007	Profit calculated from the firm's revenue and expense details for 2007 EC	$(\text{p1\_1\_1} - \text{p1\_4\_1} - \text{p1\_5\_1}) - (\text{p1\_3\_1} + \text{p1\_6\_1} + \text{p1\_7\_1} + \text{p1\_8\_1} + \text{p1\_9\_1})$
capital_stock2007	replacement value of machine, equipment, vehicles and land and building	$(\text{p4\_1}) + (\text{p4\_2})$
investment_12_months	investment in machine, equipment, vehicles and land and building	$(\text{p3\_1}) + (\text{p3\_2})$
sales_per_worker2007	Sales revenue per worker	$(\text{p1\_1\_1}) / (\text{l1\_1\_n})$
va_per_worker2007	value added per worker	$(\text{p1\_1\_1} - \text{p1\_4\_1} - \text{p1\_5\_1}) / (\text{l1\_1\_n})$

sales_growth	Sales revenue growth in 12 months	$(p1\_1\_1) - (p1\_1\_2) / (p1\_1\_1)$
--------------	-----------------------------------	--

Table 16: **Family 7: Impacts on HR and attitudes**

VARIABLE	DEFINITION	SOURCE (QUESTION NUMBER)
hr_problem	Firms face some labour management problem	$l4\_21_{new!} = -7$
hr_incentives	Firms use monetary incentives to keep workers around	14_29
fair_wage	Fair wage	14_27
contract_labour	Use of contract professional labour?	14_19
reward_systems_pro	Reward system for non-professionals	m1
reward_systems_nonpro	Reward system for professionals	m2
poor_performers	Action taken on poor-performers	m9=3

## 7 Sub-group analysis for Heterogeneous treatment effects

We will look at heterogeneity of treatment effects along the following baseline outcomes (relevant question number in parentheses):

- Main industry (s22)
- Sector level of competition in product market. (c4) - above/below the median.
- Total employment size (total\_n\_all) - is above/below median.

- Whether the firm used formal advertising for recruitment. (l4\_2\_1=1 or l4\_2\_2=1)
- Proportion of firm's workers who are professionals (l1\_5\_n/l1\_1\_n)- above/below median.
- Recruitment costs ( l4\_14) -above or below the median.

## **References**

- BENJAMINI, Y., A. M. KRIEGER, AND D. YEKUTIELI (2006): "Adaptive Linear Step-up Procedures that Control the False Discovery Rate," *Biometrika*, 93(3), 491–507.
- BRUHN, M., AND D. MCKENZIE (2009): "In Pursuit of Balance: Randomization in Practice in Development Field Experiments," *American Economic Journal: Applied Economics*, 1(4), 200–232.
- OLKEN, B. A. (2015): "Promises and Perils of Pre-Analysis Plans," *The Journal of Economic Perspectives*, 29(3), 61–80.