

Survey design and sampling procedure

IDCEA project (SOAS, University of London)

We adopted a research design based on a sequential approach that was operationalised through a carefully designed comparative framework.

Our research design was organised around a 2-by-2-by-3-by-2 comparative framework:

- two countries (Ethiopia and Angola);
- two sectors (manufacturing and construction) and specific sub-sectors within each of these;
- three origins (national/domestic, Chinese and other foreign);
- whenever possible, two varieties of Chinese capital were considered (private and state-owned), with distributions relevant to each sub-sector. Chinese state-owned enterprises are mainly found in infrastructure construction and private firms mostly in manufacturing.

In order to reduce excessive variation in outcomes and explanatory variables, the surveys focused on the type of workers that represent the vast majority of jobs created in the target sectors. According to evidence collected through interviews with managers and HR departments in selected companies in target sectors, most jobs created for national workers in Ethiopia and Angola are in the low-skilled or semi-skilled categories. Many semi-skilled workers have been upgraded from low-skilled status through on-the-job training and direct work experience. Typically, eight out ten jobs created by firms in these sectors are within these target skill categories. We therefore sampled only low-skilled and semi-skilled workers. The identification of low- and semi-skilled categories was based on a combination of two criteria, namely (a) specific job title and tasks as specified/reported by worker, and (b) qualifications in terms of education level and total number of schooling years. These classifications were also cross-checked against broad salary scales for consistency purposes.

National contexts	Angola			Ethiopia		
Sectors	Road building and dams	Manufacturing of building materials		Road building	Textile and garment, leather products (footwear, etc.)	
Firms	<u>National</u> (Angolan)	Chinese (SOEs, Private)	Other foreign (OF)	<u>National</u> (Ethiopian)	Chinese (SOEs, Private)	Other foreign (OF)
<i>Source: Authors' elaboration.</i>						

Sampling process and outcomes

We collected data via structured quantitative interviews with workers via a two-stage procedure. We first sampled firms in the relevant target sectors and then representative samples of low- and semi-skilled workers within each firm. All of the workers we sampled were nationals of Angola or Ethiopia. While selection within each firm was randomised, the sampling of firms was purposive. Firm selection followed analytical and empirical criteria, starting from the rationale for target sub-sectors:

- ✓ Sectors where job creation for low- and semi-skilled workers had been very significant in the last decade.
- ✓ Sectors where there was a large enough pool of comparable firms of the categories needed for this research: Chinese, domestic and other foreign.
- ✓ Sectors where more low-skilled or semi-skilled labour can be hired, i.e. where barriers to entry are lower, in order to capture some new labour market entrants in such sectors.

In Ethiopia these criteria led us to select three sub-sectors: textiles and garments as well as leather products for manufacturing, and road building in the construction sector, the latter linked to the scaling up of state investment in infrastructure. In Angola, the selected sectors were the building materials sector for manufacturing, and road and hydroelectric dam building in the construction sector.

Once specific sub-sectors were selected, we conducted extensive scoping research to gather the necessary information for company selection. Our sample of companies was selected according to the following criteria in this order of importance:

- Firms had to be important generators of employment, i.e. the largest and more significant job creators within each subsector, according to official data and interviews with sector experts and company managers.
- We included firms that were considered as among the most important in each sector (based on interviews in the scoping phase) and were active at the time of the survey. This was an important constraint for the road construction sector, where activity and employment depend on active projects. Additionally, for logistical reasons we prioritised construction projects that were not in the most remote areas of each country.
- We included both large and medium firms, but not small-scale firms, using the scale standards within each sector.
- We ensured that we selected at least some examples of enterprises that were known for best practice in labour standards, so that the sample had a 'top benchmark' against which other firms could be compared, instead of a sector 'average' for which there was no secondary information.

The final sample included all of the most analytically important firms across the three subsectors, according to these criteria. We compared the leading Chinese firms with the leading other foreign and domestic firms in same sectors, and not to the 'average'. An 'average' sample would have necessitated more complete sample frames and would have unnecessarily added heterogeneity to comparisons, by including firms of very different sizes and capacities.

One aim of the project was to try to obtain *representative samples within each company or site*. In the manufacturing sector the sample was restricted only to workers directly involved in production,

so as to exclude cleaners, security guards and other ancillary staff, as well as clerical and administrative workers. This meant following a number of basic principles for selection protocol:

- First, there should be a large enough absolute sample size for each site/firm: it was decided that sample sizes within each firm/site would range between 20-30 depending on the relative size of total employment in the firm/site. Larger samples sizes within same firm/site would not add much statistical precision and would add to costs unnecessarily. Moreover, the aim was to cover a reasonable number of firms/sites, as variation was expected to happen more between than within them.
- Second, we aimed to work with precise and unbiased sampling frames (i.e. lists of workers) as far as possible. In order to construct suitable local sampling frames, field supervisors were trained in and employed a variety of procedures, including making on-site lists of workers in sections of the factory or site, working with employee lists provided by the company which were then independently checked for completeness by field supervisors, or using systematic random sampling, which obviated the need for precise sampling frames.
- Third, independent of how sampling frames were constructed, interviewed respondents were randomly selected by field supervisors, who used laptops or tablets to generate random numbers for each of the two relevant worker strata, low-skilled and semi-skilled.

Generally sampling protocols were strictly followed in Ethiopia, so samples are comparable according to expectations. In Angola, teams encountered some challenges in a number of Angolan and other foreign firms (i.e. non-Chinese firms), where field teams had to stratify and randomly select workers from relatively restricted sample frames that may not have included temporary workers or recent hires, and represent mainly a *core labour force*. In the construction sector, a crisis in the sector in Angola affected some firms more than others. As a result of a national fiscal squeeze during the time of the survey project execution was hampered and several Angolan and other foreign firms were operating below capacity, with mostly their core permanent employees, whereas most Chinese firms in the sample were operating at higher intensity and initiating projects financed by the new China Credit Line approved in 2015. Therefore their workforces were more mixed and included temporary project workers and new hires in greater proportions than other comparable firms in the same sector. This sample bias is therefore acknowledged as a limitation but was unavoidable given the circumstances of Angola at the time of the survey, especially for the infrastructure construction sector. This experience also shows the methodological challenges in trying to achieve fully comparable samples in research on these sectors, especially given the impact of volatile business cycles. In any case, since the potential bias was captured, we use this information to conduct a more precise statistical analysis and qualify some of the findings for Angola in the Angola country report and the overall synthesis report (please see the link to publications below).

Table 2. Ethiopia sample

Sector	Chinese	Other foreign	Ethiopian	Total
Manufacturing	167	197	170	303
Construction	124	59	120	534
Total	291	256	290	837

Table 3. Angola sample

Sector	Chinese	Other foreign	Angolan	Total
Manufacturing	144	85	68	297
Construction	167	120	98	385
Total	311	205	166	682

The total size of samples for both countries was large enough for an adequate representation of low-skilled and semi-skilled workers in the target sectors, given total employment levels and the number of leading firms active at the time of surveys. The balance between sectors was also designed to capture the relatively greater significance of manufacturing in Ethiopia and infrastructure construction in Angola for research questions on employment outcomes in Chinese and other firms in these two countries. The manufacturing sector in Ethiopia has created many more jobs than factories in Angola, whereas the infrastructure construction boom and associated jobs have been relatively more significant in Angola.

Publications

Initial analysis of the data can be found in the project synthesis report and the country reports for Angola and Ethiopia, all of which are available at:

<https://www.soas.ac.uk/idcea/publications/reports/>.