What is Informality?

By Matt Juden, Former CGD Research Assistant

Analysis of the informal sector in developing countries began before the existence of either term. Mainstream economists and Marxist writers in the 1950s, 60s and 70s shared an interest in the 'traditional activities' pursued as survival strategies on the margins of modernising economies (see e.g. LeBrun and Gerry 1975; Lewis 1954). As survival strategies, these activities were not understood to contain any potential for growth. Their characterisation as 'traditional', in opposition to the activities of the 'modern' sector is indicative of the sense both sets of writers had that these activities would inevitably cease as processes of modernisation spread to the whole economy.

However, the informal sector has not disappeared; far from it. While such estimates are highly uncertain, Schneider (2013), for example, has estimated that the informal sector was responsible for something like 38 percent of GDP in SSA in the 2000s, though this share may now be slowly diminishing with time. Already, in the 1970s and 80s, the informal sector was recognised as an important part of the economy, especially in developing countries. Hart (1973), and the ILO (1972) identified the 'informal sector' and coined the term. Contrary to those that went before them, they suggested a degree of entrepreneurial dynamism and potential for growth in this sector that could potentially be harnessed to generate economic growth.

The ultimate expression of this optimistic view of the informal sector is the work of De Soto (most famously 2000). De Soto popularised the idea of the informal sector as constituted of a mass of small scale and micro-entrepreneurs, potentially highly capable but stuck in a low-productivity equilibrium as a result of lack of access to credit and the high costs of formalisation. This somewhat simplistic view and the optimism to which it gives rise has since been tempered by detailed empirical analysis of the nature of informal economic activities in a variety of contexts. Many of these studies share as a theoretical framework, the 'informalization' approach championed by Portes (Castells and Portes 1989) and adapted from its largely OECD and Latin-American roots to a Sub-Saharan African context by Meagher (1995). The key insight of this approach is the enormous heterogeneity of economic activities that can be characterised as 'informal', and the need to differentiate between these activities. In the mainstream of economic thought this insight has informed a tendency to differentiate 'survivalist' individuals and firms whose informality is in some sense imposed on them, from intentionally informal, entrepreneurial firms that remain informal as the result of an implicit or explicit cost/benefit decision (World Bank 2007).

If there are some firms which choose to remain informal, then two questions suggest themselves: 1) Can informal firms become more productive by formalising? 2) How can firms be encouraged to formalise? These two questions have given rise to two corresponding lines of enquiry. One investigates the productivity of firms of different sizes and on different sides of the formal/informal divide and the determinants of that level of productivity (e.g. Amin and Islam 2015; Bigsten et al 2004; Gatti and Honoratti 2008). The other investigates the determinants of firm informality so as to inform policy seeking to encourage formalisation (e.g. Djankov et al. 2002; Loayza et al. 2005).

Different answers to the two questions above imply different views on the role of informal firms in the process of development. La Porta and Schleifer (2008) helpfully characterise three different positions taken by authors in the literature. The "romantic view" they associate with the work of De Soto. As was noted above, it must be accepted, *contra* De Soto, that there are a diversity of class relations within the informal sector, meaning that many who work within it are exploited employees or survivalist owner-operators rather than dynamic entrepreneurs (Cramer, Oya, and Sender 2008; Rizzo 2011). Even so, De Soto could be correct that there are a large number of informal firms that are actually or potentially extremely productive, but held back by their lack of access to finance and by the high costs of formalization. If this view is correct, then informal firms should not differ very much from formal firms in terms of their potential for productivity, for example the educational endowment of their managers. The policy solution suggested is to reduce barriers to formalization and increase the availability of capital so as to unleash the unrealised potential of informal firms.

In the second and third positions outlined by La Porta and Schleifer, informal firms are viewed more pessimistically. Their informality is viewed as being determined mainly by the lower level of human capital of their managers, leading to a lower level of productivity than formal enterprises of the same size.[1] Therefore, these firms do not have much potential to become more productive and drive economic growth. These positions differ in their view of the effect of the presence of these firms on the formal sector. The "parasite view" worries that cost savings due to not being subject to tax or regulation mean informal firms will out-compete more productive formal firms. The "dual economy view" sees informal firms as so unproductive that this effect will not materialise and in any case cautions that in reality informal firms overwhelmingly operate in different industries and have different customers to formal firms.

As well as conceptual clarity, La Porta and Schleifer's great contribution is the empirical testing of the three views above. They use World Bank Enterprise Surveys data to assess the productivity and characteristics of informal and formal firms of different sizes in different countries and find no support for the "romantic view", judging the characteristics and levels of productivity of informal firms to be so different from formal firms that it is highly implausible that they would benefit sufficiently from formalization. Rather, they endorse the "dual view". However, closer attention to the differences between countries supports a more nuanced conclusion. Gelb et al. (2009) demonstrate that the picture for East African countries is extremely different to that in Southern African countries, with massive differences in productivity and characteristics suggesting that the 'dual view' best characterises the role of informal firms in Southern Africa but that East African countries have similarly productive informal and formal firms at the same firm size with similarly endowed managers. However, Amin and Islam (2015) make use of more recent Enterprise Surveys data to demonstrate the informal firms decrease in productivity as they increase in size, suggesting that they may possibly struggle to capture the benefits of formalization even in the East African case.

[1] It is an assumption shared by all in the literature that, at least for formal firms, larger firms are more productive due to the presence of fixed costs especially of marketing, development, etc. leading to economies of scale.

Definitional approach

Since Hart (1973) first brought them into existence, 'informality' and 'the informal sector' have been contested terms. This contest has most usually taken the form of disagreements over the correct definition of 'informal sector' (see Gerhani 2004 for an overview). Partly on the basis of their observation of chaos in the definitional literature, some authors have gone so far as to suggest that 'adopting the term obscures analysis of central issues and is counterproductive' (Peattie 1987). Still others have observed that the tendency toward dichotomous characterisation of economic activity in general is damaging. They point out that to divide activities into 'formal' and 'informal', or 'traditional' and 'modern', or 'artisanal' and 'industrial' is a reductive exercise that obscures important processes acting at the boundary between one sector and the other (see e.g. Klein 1999).

A review of the literature shows that 'informality' is a fuzzy concept that cannot be given a unitary definition fit for use in all contexts. That is to say, necessary and sufficient conditions cannot be given for an activity to be 'informal'. However, this is also true of 'game' (Wittgenstein 1953). Informal activities, similarly to games, are related by family resemblance, allowing language users to reliably identify examples of informal activities despite the fact that disagreements about marginal cases cannot be settled by reference to an agreed definition. In this context, it is legitimate to use a restrictive technical definition to identify a subset of informal activities to form the object of a particular study in a particular context.

In every country, the World Bank's Enterprise Surveys group defines 'informal firms' as those firms that are not registered with 'the registration office where businesses are required to register' (Enterprise Surveys 2013). 'Unregistered firms' and 'informal firms' are therefore used interchangeably by Enterprise Surveys to refer to this set of firms. Our analysis contrasts this set of firms with firms who are registered with their local authorities. Thus, our analysis categorises firms as 'formal' or 'informal' based on just one of the criteria that have been suggested in the literature for distinguishing between formal and informal firms. Others have suggested that informal firms are characterised by being family owned, small in scale, labour- as opposed to capital-intensive, utilising skills acquired outside the school system, and so on. In following the Enterprise Surveys methodology for determining whether a firm is 'informal', we cannot be so multidimensional in our analysis. Instead, we restrict ourselves to defining informal firms as those firms that are not registered.

This conceptual choice could be criticised as either being in some sense too restrictive or too broad. It is too restrictive in the sense that some formally registered firms may be involved in informal activities. Thus, informal firms as we define them are not responsible for all informal activity emanating from firms. The set of informal firms as we have chosen to define it might also be thought to be too broad by those who think that a lack of registration is a necessary but not sufficient condition for informality (e.g. ILO 2002). That other authors in the literature are operating with more permissive or more restrictive conceptions of 'informality' should be borne in mind when situating empirical results within the wider literature.