Review paper 1

Key debates in competition, capabilities development and related policies:
Drawing the link between barriers to entry and inclusive growth

19 February 2015

Prepared by:

Fatsani Banda

Genna Robb

Simon Roberts

Thando Vilakazi

Contents

1. Introduction: the South African policy context

With over 20 years in a democratic South Africa, the country still faces serious economic challenges. The unemployment rate is extremely high, with 24.3% of the population classified as jobless and actively looking for work (Statistics South Africa, 2015). The expanded unemployment rate, which includes people who have given up looking for a job, stands at around 34.6%, or 8.1 million people (Statistics South Africa, 2015). Economic growth has not been as high in recent years as in other emerging market economies and, critically, has not led to large amounts of job creation. Instead the unemployment rate remains stubbornly high and large portions of the population, particularly the young, are essentially excluded from the formal labour market. Inequality also remains high and actually increased between the early 90s and late 2000s (OECD, 2011). In 2011, South Africa's gini coefficient – a measure of inequality – was the highest of all the BRICS countries, and twice the OECD average (OECD, 2011).

In response to these challenges, the need for a higher level of economic growth, combined with a more inclusive approach to sharing the benefits of such growth are universally acknowledged and this has been the tone of recent key policy documents. The New Growth Path (NGP) set out a framework for industrial and economic development, in particular targeted investment into industrial activities, which have the propensity to be labour intensive. The National Development Plan (NDP) goes beyond this, by providing an extensive overview of the current challenges facing South Africa, highlighting the key constraints to achieving inclusive growth and putting forward a broad set of solutions.

One of the key challenges identified by the NDP is "*the legacy of highly concentrated industries that have limited competition and efficiency gains*". It notes that South Africa suffers from uncompetitive goods and services markets as a result of apartheid growth patterns and sanctions-induced isolation. Whilst this results in high profit margins, there is little investment and innovation, new firms are not entering the market and employment creation is low.

South Africa's distance from markets, relatively small market size and high cost structure make it difficult to break up its numerous oligopolistic markets. Similarly, the Industrial Policy Action Plan (IPAP) has identified microeconomic policies such as competition that work towards unlocking markets to allow for greater economic participation by private actors and economic growth. The NGP also speaks of the need to correct for market failures where there is a history of anti-competitive behaviour and rent-seeking, with the idea that this would help fulfil the national goals associated with industrial development. This is based on the notion that high levels of market concentration in key industries hinder the creation of growth and employment linkages from these industries.

Effective rivalry between firms to win over customers encourages firms to produce better quality goods and offer lower prices. This rivalry requires firms to be more prudent in their use of the resources available to them by eliminating inefficient use of resources, cutting down wastage and thus reducing their costs (Evans and Joekes, 2008). In this context, firms can compete on the basis of improved product offerings and investments in improving their capabilities, in which case efficiency, effort and ingenuity is rewarded. On the other hand, firms can compete (unfairly) by leveraging their incumbency and engaging in practices that seek to raise rivals' costs and ultimately diminish the significance of rivals as effective competitors. This can also be described as the difference between 'performance competition' and 'handicap competition' (seeking to handicap rivals) (Gerber, 2010).

The South African experience has been one of high levels of concentration in several key sectors resulting in an economy which largely excludes the majority of the population from ownership of important economic assets. This has deprived the economy of the benefits from increased participation and access to markets, and dynamic rivalry in the form of competition based on innovation, competitive pricing, investment in improved production processes and capabilities, and learning-by-doing.

In practice, the capabilities in the South African economy were narrowly developed to cater for the strategic interests of the apartheid government. This includes state-owned interests and intervention in sectors such as telecommunications, mining, agriculture, and energy. This led to the creation of dominant firms in important sectors whose incumbency has subsequently allowed them to extract rents and shape even the current regulatory regime in their interests. Although the South African economy was largely opened to the global market in the past twenty years, any benefits that could have resulted from liberalisation have been muted by the fact that the actual structure of markets remains concentrated. This appears to be at least partly due to the high barriers to entry in key industries, which, by creating and reinforcing the market power of large firms, tend to lead to higher prices, lower levels of innovation and a less competitive economy.

For this reason, the National Treasury has allocated funding to the Centre for Competition, Regulation and Economic Development (CCRED) at the University of Johannesburg to conduct a programme of research focused on barriers to entry and inclusive growth. It is envisaged that this will involve researching and analysing the barriers to entry across a wide range of sectors in South Africa with the intention of formulating policy recommendations that will help to facilitate greater levels of entry and competition and thus drive higher growth.

The review paper will highlight the factors which are critical to facilitating the entry of efficient rivals into traditionally concentrated markets premised on the understanding that investment and innovation emanates from and rests with firms. In this context, government's role (including its various agencies) should be to facilitate the process by lessening structural and strategic barriers in the economy which restrict the dynamic rivalry introduced by firms that compete on the basis of investments, innovation and developing new processes and capabilities. The paper will relate these issues back to the international growth and development literature in order to provide a strong motivation for the relationship between barriers to entry, competition and dynamic efficiency, increased participation in the economy and inclusive growth. This assessment will provide the theoretical context for understanding the sector-specific experiences of firms in the sectors which the project will analyse. A second review paper will discuss the literature on competition policy, economic regulation and industrial policy.

Section 2 presents a discussion of recent developments in the economic growth and development literature, particularly focussing on the idea of inclusive growth, and relates it to the South African context. Section 3 considers the benefits of competition for the economy before presenting a discussion of barriers to entry in theory and practice. Section 4 defines the concept of inclusive growth and illustrates its importance for South Africa. Section 5 then brings the discussion together, making the link between barriers to entry and inclusive growth and motivates for a research agenda around this link.

2. Growth, development and capabilities

2.1 Literature and international experience

The literature on economic growth and development offers divergent views about the critical levers towards unleashing economic development in developing countries. The global focus in recent decades on economic growth as a necessary condition for poverty reduction has yielded some benefits to the poor and marginalized in terms of decreases in the level of poverty. However, growth itself is not a sufficient condition for poverty reduction because it does not guarantee that all persons will benefit in an equal way from growth, or that inequality will be reduced (which has implications for political stability) (Ali and Son, 2007; Ianchovichina, 2009). In this context, the concept of "inclusive growth" has come to the fore, underpinned by an attempt to more explicitly account for market failures and the inability of traditional models of growth to account for distributional factors and the importance of ensuring that the economy is *actually* and not just notionally accessible to all economic actors.

The use of the term 'inclusive growth' has evolved from the earlier conceptions of 'pro-poor' or 'broad-based' growth which were popular in the 1990s and 2000s as levels of inequality began to rise between and within countries. Inclusive growth is related to pro-poor growth, particularly because both necessitate that poverty and inequality be reduced in order for growth to be meaningful (Ranieri and Ramos, 2013). The development of the concept of inclusive growth over the past 10 years has been part of an attempt to broaden the concept of economic growth to include the well-being of all the citizens of a country – most notably the poor – where poverty is understood in both absolute and relative terms. Ramos, Ranieri and Lammens (2013:4) put it this way:

> *"Inclusive growth is both an outcome and a process. On the one hand, it ensures that everyone can participate in the growth process, both in terms of decision-making, for organising growth progression as well as in participating in the growth itself (and earning income). On the other hand, it goes some way towards ensuring that everyone equitably shares the benefits of growth. Inclusive growth implies participation and benefit sharing. Participation without benefit sharing will make growth unjust and sharing benefits without participation will make it a welfare outcome".*

Inclusive growth defined this way is about not just ensuring that the poor benefit from growth, but also that there is increased participation of the poor and disenfranchised individuals in in the process of growth. This would occur via an increase in employment and entrepreneurship, as well through entry and participation at the firm level (access to markets), whereas the benefits thereof are derived from rising incomes and increased social expenditure. Anand, Tulin and Kumar (2014) state that inclusive growth refers to both the pace and distribution of economic growth. The sustainability of growth relies on whether it is inclusive (Berg and Ostry, 2011; Kraay, 2004). Essentially, although economic growth is necessary to improve the level of inclusiveness in an economy, it is not sufficient in and of itself to ensure long-term growth and higher levels of inclusiveness (Berg and Ostry, 2011; Kraay, 2004).

There are therefore clear gaps in the theory and definition of inclusive growth, if policy is to be developed based on attempting to achieve it. For instance, to what extent should policy place an emphasis on incorporating survivalist and informal enterprises within the planning around industrial development? Is inclusion in terms of employment opportunities and income a sufficient marker of successful inclusion, or should the focus be on encouraging dynamism at the firm level with the view that it will create opportunities for labour absorption? Should policy seek to actively encourage inclusion and provide incentives for entry, or is it sufficient to rely on microeconomic tools such as competition policy to deal with strategic barriers to entry?

Increasingly, theories of growth and development acknowledge the reality that markets are imperfect and not always self-correcting and that firms with incumbency advantage and market power have rational incentives to employ strategies (legal or anti-competitive) to protect their position. Recent literature establishes an important link between growth and development and removing these constraints to accessing economic activity. Acemoglu and Robinson (2012) for example talk about “inclusive” versus “extractive” institutions and arrangements where “inclusive” institutions are more likely to lead to economic growth. Inclusive institutions include secure property rights; law and order; markets and state support (public services and regulation) for markets; relatively free entry for new businesses; and access to education and opportunity for the great majority of citizens. In other words, a state which creates incentives for investment and innovation and a level playing for firms to compete is more likely to generate economic growth, although of course it is important to consider the relative importance of these factors to development on a case by case basis.

The ideas of inclusivity, equality of opportunity and ease of entry for new businesses are central to Acemoglu and Robinson’s thesis. In a related vein, North, Wallis and Weingast (2012) introduce the idea of “open access orders” and “limited access orders” to explain the political and economic differences between countries, where open access orders provide citizens with open access to public and private organizations whereas limited access orders are ruled by a dominant coalition and people outside the coalition have only limited access to organizations, privileges, and valuable resources and activities. The former are associated with higher rates of economic growth. The authors argue that “Over the long-term, open access politics cannot be sustained without open access economics, and vice versa”, highlighting again the importance of diverse economic participation and inclusivity. For the most part, open access orders are therefore normative or idealistic frameworks.

In several developing country contexts, the reality is that models of development and for understanding social structures which are drawn from industrialised countries do not fit, in that they may not account for the fact that different leadership groups compete vigorously for control of violence and power, and governments are often changed and deposed in the process which creates instability (North et al, 2007). Instead, social structures are shaped in terms of limited access orders which lie on a spectrum of fragile, basic and mature limited access orders. The positions on this spectrum are determined largely through the degree to which relations between elites and those outside of these arrangements are governed through personal versus increasingly (and preferably) impersonal exchanges (i.e. towards mature limited access orders).

In practice, this means that powerful individuals who control these rents, have an incentive to manipulate the economy to create rents and protect these rents, including through the control of force. This speaks, for instance, to how the arrangements between the state and large firms are created and maintained over time. In the policymaking space, this plays out in the form of incumbent firms or groups of firms which are able to lobby and enter into arrangements with political actors to shape regulation and the economic environment to protect rents, effectively setting the rules of the game to suit incumbent interests in the political and economic space.

The argument advanced in this paper is that it is therefore vital in the role of developing country governments and their agencies, including regulators and authorities, to act in a manner which changes the established rules of the game. Specifically, it becomes critical for these actors to make entry and participation possible by regulating *for* competition and inclusivity through addressing barriers such as anti-competitive conduct or prohibitive tariff and pricing structures, using institutions not governed by personal affiliations and arrangements. In this way, access to markets will depend not on who you know, but rather through developing institutions governed by ‘impersonal exchange within elites’ (North et al, 2007). However, this relies on the dominant coalition, finding it in their interest as elites to move towards impersonal exchange and therefore increase access. The example of regulatory changes and competition law cases in the telecommunications sector in South Africa in the context of the incumbent firms’ interests illustrates this, as we discuss below.

The success of the late-industrializing nations in East Asia illustrates the point that one size does not fit all in terms of growth strategies and that governments have a role to play in this process. There is certainly evidence that rapid investment-based growth strategies based on countries borrowing innovation (products and processes) from already industrialized countries, and a strict reciprocal partnership between government and the companies that it supports, can succeed in catalyzing backward economies (See Amsden, 1989). The examples of post-war Japan and South Korea, European countries in the nineteenth century, and countries such as Brazil, Mexico, Peru and Turkey in the mid-1970s are illustrative in this regard (Acemoglu et al, 2002). The authors argue that backward economies can grow on the back of investment-based, technology-absorbing growth strategies (directed by the state) however as they approach the ‘world technology frontier’ this growth will need to rely increasingly on their own innovation activities.

In this context, developing countries need to develop strategies that tie in closely with their endowments in resources and capabilities. The importance of building and expanding capabilities has been highlighted as being at the centre of a country’s economic development (Sutton, 2004; Page, 2011; Hausmann et al., 2007). Hausmann et al (2007) argue that ‘what you export matters’ and that it is important to consider the latent level of economic complexity and capabilities with which an economy begins its development process. This suggests that it is important to be strategic in deciding which sectors government should be prioritizing in terms of improving the conditions for new, effective entry to take place given its limited resources. Those sectors need to tie in closely with the *actual* capabilities of the country. More complex economies are those that have a wide range of capabilities which allow them to make decisions to produce an increasingly complex type and range of products.[1] Where a country begins with a relatively low level of capabilities it may be beneficial for government to intervene to 'help coordinate the accumulation of capabilities' for instance through the creation of institutions that help both existing businesses and entrants to develop capabilities (Hidalgo, 2009:18). The latter again implies a role for regulators and government agencies in 'steering' the process of developing capabilities through intervening to ensure wider opportunity and participation by removing barriers to entry.

A further, but no less significant, role for government actors is to ensure that the process of competitive rivalry is instilled in the economy, either ex-post through the competition regime or ex-ante through regulators that 'regulate for competition' by making entry possible or correcting failures in the market.[2] The success of the East Asian countries has been underpinned by state support that still ensures that mechanisms are in place to stimulate competitive rivalry. In the case of South Korea, this pertained to export incentives that firms were required to compete for (Roberts, 2010). Similar evidence has been found in Vietnam in the case of small and medium-sized manufacturing firms (Trung, 2008). These outcomes are directly linked to the idea of ensuring *effective* entry and rivalry.

For the purposes of this paper, inclusive growth however defined needs to account for dynamics at the firm level given that the processes of rivalry are likely to drive the levels of innovation and investment, and the development of capabilities that can in turn lead to the creation of employment, particular as the structure of markets changes in the long-run. Therefore, inclusion is most appropriately defined as a form of growth that increases the economic participation of not only individuals but firms (particularly those owned and operated by historically disadvantaged groups) as well, and has the ability to allow for greater benefit-sharing that results in a more equitable distribution of the benefits of economic growth. It is important to understand rivalry between firms encouraged through pro-competitive regulatory and policy measures and enforcement against abuses of market power, as transmission mechanisms or tools to achieving inclusive growth. In this context, understanding firm behaviour and strategies in the presence of competition (or the lack thereof), and the factors which help or prevent entrants from accessing markets and growing is critical to understanding the processes which can lead to inclusion.

2.2 Inclusive growth in South Africa

The term inclusive growth has not been clearly defined in the South African context either. For example some refer to it as an increase in employment and labour productivity, or the rise in public sector employment and employment schemes, while others see inclusive growth as improved social outcomes in health and education akin to the development of human development indicators, and in some instances social welfare and social protection (Fourie, 2014). Remarkably, the NDP, which frequently mentions the pursuit of an inclusive economy and inclusive growth, does not make an explicit attempt to define the term although it is used as the motivation for particular policy goals (Fourie, 2014).

[1] In this context, 'capabilities' do not simply refer to differing endowments of a small set of factors of production (labour, capital, land), but rather defines products as 'large collections of specific inputs' that are required to be locally available to produce a product. Capabilities can be highly specific inputs such as bridges and road or certain cultures and norms are requisite to producing a particular product. Economies with greater capabilities can produce more varied and sophisticated products.

[2] See Das Nair and Roberts (2014); and Das Nair, Mondliwa and Roberts (2012).

It is clear that South Africa's performance in terms of indicators related to inequality and growth has been poor in the past two decades. Measures of inequality could be viewed as one indicator of the extent to which the benefits of growth are shared, and as noted in the introduction, inequality in South Africa is extremely high by comparison to its peers and has increased over the last 15 years. By contrast, Brazil for example has made significant progress in reducing inequality in recent years. Through much of the late twentieth century Brazil was ranked as one of the world's most unequal societies and in 1989 faced a Gini coefficient of 0.634. Between 2001 and 2006, however, Brazil's Gini coefficient fell significantly, to 0.526 in 2009.[3] In Brazil, incomes have risen across all class groupings, but what is most significant about this is how the incomes of the poorest Brazilians have risen much faster than that of the country's wealthy.

Cross-country comparisons of this nature can sometimes help us to understand how different countries have achieved more inclusive economic outcomes. Indeed, a great deal of comparisons, are often drawn between South Africa and its 'peer' nations, particularly those within the BRICS grouping. However, comparisons across countries of this nature are not helpful if they do not consider the comparative structure of industries, and specifically the dynamics of entry and exit that characterise each country. As in the literature developed above, countries need to structure their development path around their own capabilities, based on a clear understanding of the unique circumstances of firms, particularly entrants relative to incumbents, in each domestic market. The possibility for inclusion to take place, based on factors which are broader than just the distribution of incomes, strongly depends on the prevailing social order of a country and the interests of various groups (political and economic actors) in that society which shape the rules of the game. As discussed above in needs to be in the interests of economic and political elites to include other actors, just as a group of firms in a cartel arrangement make strategic decisions as to whether to undercut new entrants which are outside of the arrangement, or to invite them into it. In this way, what might work in one country in terms of increasing participation, may not apply in another, and the same caution should apply in the comparison across different economic sectors without a clear understanding of the competitive dynamics within each.

It is therefore critical to understand the specific experiences of firms and particularly entrants across the South African economy, including their capabilities and the barriers faced in terms of growth and expansion. Without this, policymaking will miss the mark in terms of generating a sustainable growth which is inclusive in the broad conceptualisation considered in this section.

South Africa's history implied that capabilities were narrowly developed to cater for the strategic interests of the apartheid government. This includes state-owned interests and intervention in sectors such as telecommunications, mining, agriculture, and energy. This led to the creation of dominant firms (and groups of firms) in important sectors whose incumbency has subsequently allowed them to extract rents and shape even the current regulatory regime in their interests. Although the South African economy was largely opened to the global market in the past twenty years, any benefits that could have resulted from liberalisation have been muted by the fact that the actual structure of markets has not

[3] Newsweek 13 august 2009. "A crisis fluke: Brazil's shrinking wealth gap". Available online: http://www.newsweek.com/crisis-fluke-brazils-shrinking-wealth-gap-215802

changed significantly (Makhaya and Roberts, 2013:567). The case studies in chemicals, fuels and fertilizer, grain silos, the milling value chain, and telecommunications discussed by the authors illustrate the failure of regulation, competition and industrial policy to stimulate dynamic rivalry and tip 'the balance of power' in favour of new entry and expansion into new dynamic sectors of the economy (Makhaya and Roberts, 2013).

The outcome of the combination of these factors is an economy in which capabilities have been narrowly developed, and clustered around a handful of resource-intensive sectors that have benefited from access to cheap migrant labour and subsidized electricity. Despite very strong capabilities in the extractive industries and industries supporting them (such as financial services), the *range* of capabilities as described in the economic complexity literature is narrow (Hausmann and Klinger, 2008) and incumbent firms have an interest in protecting their position. Arguably, it has not been a focus of South Africa's industrial policy to stimulate entry and encourage linkages into the production of increasingly more complex products. In this regard, the lack of progress in lowering barriers to entry and expanding access and opportunity may be regarded as a failure of industrial policy to-date.

The following section presents a brief discussion of the benefits of competition before expanding on the foregoing discussion of the challenges facing South Africa in terms of stimulating rivalry, dynamism and the expansion of capabilities which can drive growth. It then goes on to discuss in more detail the concept of barriers to entry and their importance for South Africa in the context of these difficulties.

3. Competition and growth

3.1 The benefits of competition

Why competition matters

Competition policy has evolved out of a recognition that monopoly and concentrated markets reduce welfare, whereas competition can deliver efficient markets, low prices, a more dynamic and innovative economy and therefore greater overall welfare (Motta, 2004). There are three kinds of efficiency which competition encourages: productive efficiency, allocative efficiency and dynamic efficiency. Productive efficiency is generally lower in concentrated markets as monopolies have less incentive to reduce costs than firms facing strong competition (Motta, 2004; Evans and Joekes, 2008). By contrast, in a competitive market, firms have an incentive to reduce their costs in order to win greater market share and higher returns (Evans and Joekes, 2008). Allocative inefficiency refers to the welfare loss which results from the firm with substantial market power charging a price which is above the competitive price. This implies a lower level of demand (at the higher price) and consequently a lower level of output. This reduces welfare since it implies that there are customers who could have been profitably supplied with the product but will not buy the product at the monopoly price.[4] Finally, dynamic efficiency refers to the extent to which a firm introduces new products or processes (Motta, 2004). A monopolist will not have an incentive to invest in introducing such innovations if it does not have to fight to retain customers.

[4] At each price between the monopoly price and the competitive price, a producer could supply the product at a profit and there would be consumers who demand it.

Higher economic efficiency thus ultimately means lower prices, higher output, and a more dynamic and innovative economy. The competitive process also contributes to the creation of employment in so far as barriers to new entry are reduced especially where firms can enter markets or expand their operations, although there are of course some trade-offs in terms of the short-term losses that may take place where inefficient firms are pressured out of the market. This is by no means an easy trade-off, and the latter needs to be balanced against overall gains to society, particularly in the medium- to long-term. Conversely, anti-competitive mark-ups have been identified as inhibiting growth, productivity and the creation of employment (Aghion et al, 2008).

This is especially relevant in South Africa. Government has emphasised the creation of jobs in the short-term, which has proven to be difficult and unlikely to change significantly in the near future without addressing the levels of concentration and structural constraints on the economy which prevent markets from growing through entry, innovation and welfare-enhancing rivalry. Another way to consider this, is that an incumbent dominant firm facing limited to no rivalry will generally have no incentive to innovate, expand their operations (particularly in sectors where fixed costs are high and scale economies are critical), reduce costs or price, or draw in additional labour over time - society would therefore not benefit from any growth in employment in the longer term in any case. To the extent that incumbent firms of this nature are producers of key intermediate inputs to industrial sectors with potential for growth, growth is further stifled when the costs of those inputs are kept above a competitive level due to the unilateral exercise of market power by established firms. This is precisely the risk inherent in protecting dominant incumbent firms in small developing country markets in particular.

The importance of competition in South Africa

The benefits which rivalry can deliver are reflected in the preamble and purpose of the South African Competition Act which tasks the competition authorities with promoting and maintaining competition in the South Africa, not for its own sake, but explicitly in order:

> *(a) to promote the efficiency, adaptability and development of the economy;*
>
> *(b) to provide consumers with competitive prices and product choices;*
>
> *(c) to promote employment and advance the social and economic welfare of South Africans;*
>
> *(d) to expand opportunities for South African participation in world markets and recognise the role of foreign competition in the Republic;*
>
> *(e) to ensure that small and medium-sized enterprises have an equitable opportunity to participate in the economy; and*
>
> *(f) to promote a greater spread of ownership, in particular to increase the ownership stakes of historically disadvantage persons."*

These objectives are thus concerned with more than efficiency. There is also a clear distributional aspect to competition policy in South Africa, as reflected in the objectives around employment, SMEs and ownership. In this context the specific provisions of the Act seek to restrain particular trading practices which undermine the attainment of these goals.

Mergers between firms which could substantially prevent or lessen competition to the detriment of consumers are prohibited and firms which seek to abuse positions of market power to stifle rivalry or exploit consumers can be prosecuted. Furthermore, the explicit provisions regarding the evaluation of public interest concerns that may arise from merger transactions are, by design, a means towards linking competition law with the greater socio-economic development agenda of the country (Evans and Joekes, 2008:32). The objectives of competition law in South Africa, at least on paper, are therefore directly related to the attainment of economic development, transformation of the economy, and economic redress.

There is a clear link between barriers to entry and inclusive growth. In crude terms, if new local businesses cannot access markets or if their costs are raised indirectly by the anti-competitive behaviour of established rivals, these firms will not be profitable. If these entrants are not profitable then they cannot compete on the basis of innovation, efforts to increase efficiency, achieving economies of scale and scope, or building capabilities through learning-by-doing; and they certainly cannot compete with incumbent firms on pricing and quality.[5] Critically these firms also cannot contribute to employment creation as mentioned above, and the objectives of downstream industry development outlined in the country's industrial policy strategy. This breakdown in the processes of competitive rivalry ultimately results in poor competitive outcomes which means downstream firms and consumers pay much more for their inputs and goods, respectively.

These issues are explicitly dealt with in the various policy documents which set out government's economic policy framework for the coming years. The NDP expands notes that South Africa suffers from uncompetitive goods and services markets as a result of apartheid growth patterns and sanctions-induced isolation. Whilst this results in high profit margins, there is little investment and innovation, new firms are not entering the market and employment is low. South Africa's distance from markets, relatively small market size and high cost structure make it difficult to break its numerous oligopolistic markets. Similar points are made by the New Growth Path (NGP) and IPAP 5, which highlight the importance of increasing levels of competition and envisage a strong role for competition policy in lowering input costs and prices to poor consumers (See, EDD, 2010; and DTI, 2013).

The NDP also recognizes that small and expanding firms will generate employment and drive growth in South Africa in the future, including through public and private procurement (NPC, 2011b:117). Importantly, the NDP notes that aside from creating jobs, there are other advantages to broadening the base of new and expanding firms. These include reduced levels of economic concentration and higher levels of competition. The NDP and NGP both suggest that public policy can be supportive of SMEs through lowering barriers to entry, reducing regulatory red tape and providing an entrepreneurial environment for business development.

This view is supported by empirical evidence on the experience of SMEs in South Africa. The SBP SME Growth Index tracks 500 SMEs across three sectors over time in order to understand the dynamics of South Africa's SME sector. In 2013, survey respondents reported that the biggest impediments to their growth were the lack of skilled staff, burdensome regulations, local economic conditions, lack of finance and the cost of labour

[5] See, for example, the case of liquid fuel wholesalers in South Africa, in Paelo et al (2014).

(SBP, 2014). The authors conclude that “the evidence is clear that South African SMEs confront tough, often hostile, operating conditions”. It is not clear, however, that these issues represent constraints for SMEs specifically or whether they are challenges for all firms and in all sectors. This points to the need for sector-specific studies to test the experiences of entry for small firms and to explore the barriers to entry they face in more detail.

There is also a link between policies to facilitate the entry and growth of SMEs and the goal of transformation. The NDP recognizes that transformation policy has suffered from a number of problems including a lack of capital, a preferential procurement regime which is “riddled with corruption and incompetency”, and a lack of policy and regulatory alignment (NPC, 2011b:139). It suggests that in order to ensure transformation going forward, South Africa needs to create an enabling environment for SMEs and entrepreneurs, which includes lowering the cost of doing business and reducing barriers to entry in value chains.

Other areas recognized by the NDP as binding constraints on growth are energy generation and distribution, logistics and telecommunications (NPC, 2011b:119). The plan motivates that regulatory certainty and institutional reforms are required in order to promote competitive outcomes in network industries and to get rapid increases in administered prices under control. The NDP sees reducing costs and increasing levels of competition and productivity in network infrastructure as preconditions for achieving “a diversified dynamic economy” (NPC, 2011b:122-123). IPAP 5 similarly identifies high and escalating administered prices, in particular electricity and port prices, as a major constraint on industrialisation (DTI, 2013:19).

Evidence of the benefits of competition

This overview of the treatment of competition, increased participation and more dynamic markets in South African policy documents is important in so far as it clearly espouses a role for competition and the reduction of entry barriers in reinvigorating growth in the country. This approach ties in with the growing body of evidence which demonstrates the benefits of rivalry, dynamism and more competitive markets.[6] One way to demonstrate the benefits of competition is through looking at the counterfactual of no competition or where there is/was anti-competitive behaviour, and comparing this world to one where there has been entry, increased participation and competition, and successful pro-competitive intervention. This is effectively the same as estimating the damage caused by (strategic) restraints to competition through, for example, cartel conduct or abuses of dominance by incumbent firms which both raise barriers to rivals. The most recent economic studies in this area are the competition impact assessment studies conducted by the competition authorities and academics in South Africa, following different interventions. This research suggests that there have been significant benefits to consumers from interventions that have sought to enhance competition in various markets.

The most obvious examples of the benefits of effective rivalry, or the harm caused by restraints to rivalry, are overcharge estimates which demonstrate the extent of harm caused by cartel conduct. Specifically, overcharges represent the profitability of cartels or conversely the losses to society from the conduct by estimating the mark-ups charged by cartelists over the duration of the cartel, over some competitive benchmark or counterfactual price that

[6] See for example Mncube (2013), Khumalo et al (2014), Das Nair et al (2014) and Boshoff (2013).

would have prevailed absent the conduct. Cartels, by their nature, face two primary problems being the possibility of cheating by one of the members, and entry by a firm which is outside of the cartel arrangement that could threaten to undercut cartel profits and thus destabilise the arrangement. Cartelists therefore face very strong incentives to deal with new market entrants by undercutting them or preventing entry altogether through signalling effects and/or investing in excess capacity (or threatening to use their excess capacity or any control over key infrastructure or inputs they may have to undermine entry which is typical in cartel cases).

The various methods used to estimate overcharges are not important for this paper, but rather the quantum of the damage as a proxy for the benefits that *would* have accrued to consumers, other things equal, were it not for the cartel. Khumalo et al (2014) studied the cartel in precast concrete products following the end of the cartel in around 2007 (which had started in 1973) and found mark-ups of approximately 16.5% to 28% in Gauteng, and between 51% and 57% in KwaZulu-Natal. This means that users of these products, which included government through infrastructure programmes, paid up to approximately 28% and 57% more in the two regions, respectively, than they would have in a more competitive environment. Interestingly, following the end of the cartel there has been new entry into the industry and some of the firms that were previously party to the cartel have expanded their operations into other geographic areas to the benefit of customers. Mncube (2013) studies mark-ups in the flour cartel which lasted from 1999 to 2007 and estimates that the mark-ups charged to independent bakeries as users of the product were in the range of 7% to 42%. Similar studies have been conducted by Boshoff (2013) on the bitumen cartel, Das Nair et al (2014) on the reinforcing bar cartel, and Govinda et al (2014) on the cement cartel, each estimating mark-ups which are relatively high and comparable to international studies in this area.

The overcharge findings illustrate the substantial costs to consumers from anti-competitive conduct and speak to the importance of enforcing against arrangements which unfairly privilege incumbent firms, and ‘handicap’ any potential rivals that could compete away excess margins. Most notably, cartels have been in key industrial inputs and products, which affect the trajectory of economic growth by affecting key infrastructural and productive inputs. These include bitumen, scrap metal, reinforcing bar, construction, cement, concrete pipes, plastic pipes, industrial gases and steel products. Perhaps the most notable of these is the recent cartel in the construction sector involving fifteen of the largest construction firms operating in South Africa. Several of the tenders affected by this conduct included key public infrastructure projects such as roads and hospitals. This should raise concerns from the perspective of policymakers tasked with managing the processes of economic development and investing in infrastructure as part of that. Similarly, where cartels affect consumer goods, like bread, they directly harm consumers, including the very poor. When inputs or consumer goods are expensive to end-users due to unfair practices by incumbent firms that have locked out entrants and prevented rivalry, growth cannot be inclusive.

As noted in the reference to the NDP above, telecommunications is a critical sector for economic growth. There have been a number of abuse of dominance cases in the sector, notably those against the state entity Telkom. Similar cases have also been referred and prosecuted against South African Airways (SAA). In the case of Telkom, which we discuss further below, the company had sought to leverage its control over fixed-line telecommunications infrastructure to disadvantage rival downstream internet service providers in favour of its own downstream operation. The intervention of the competition authorities subsequently opened up the downstream market to entrants and competing internet service providers. This level of the market has since become a vibrant one with several firms competing on the basis of making better technology and price options available to customers. In this way, removing a strategic barrier to entry imposed by an incumbent firm led to the inclusion of several other firms into the market with further benefits being shared by end-users as well. Although not documented, it is expected that the entry of new firms in this or other sectors for that matter would also contribute to the creation of employment opportunities.

Critically, the role of addressing barriers to participation does not depend solely on competition law interventions, but on the full spectrum of government agencies and departments, and the underlying policy frameworks, tasked with overseeing various economic sectors. For instance, sector regulators have an important role to play in shaping and enforcing the body of rules which govern the interactions between firms and grant access to those firms wishing to enter markets. The Independent Communications Authority of South Africa's (ICASA) various interventions on call termination rates, including through the introduction of measures to support the smaller mobile network operators that were latecomers to the market, have over time led to significant reductions in the prices of mobile telephony in South Africa.[7] It is however questionable whether new entrants and smaller rivals such as Cell C and Telkom Mobile have been become *effective* competitors in this market, although their opportunities to do so are certainly enhanced by the regulatory changes.

Regulators and authorities in various African countries have played a similarly important role in encouraging growth and competition in mobile money transfer services, for instance through licensing new rivals,[8] which have to date not taken off as expected in South Africa due in part to regulatory rigidity in the banking sector. The domestic banking sector has been characterised by a largely conservative regulatory approach relative to other countries.[9] As it stands, the proliferation of mobile money services in countries such as Kenya and Zimbabwe has made it possible for consumers, including the very poorest, to send and receive money at rates far lower than those that have been historically provided by banks. This presents a more convenient, safe and cost-effective offering to the consumer and facilitates the inclusion of previously marginalised, unbanked consumers especially in rural areas in the payment system in competition with existing banking systems. This also opens the opportunity for the growth of new enterprises based on these more convenient, safer transaction mechanisms. In Kenya, the rates charged by the incumbent operators have just recently been significantly reduced which appears to be in response to the impending entry

[7] See, for example, ICASA Draft Call Termination Regulations wherein the regulator explains the primary objectives of the intervention as being to enhance competition in the sector. The regulator has subsequently revised the specific recommendations including the most recent version in 2014, which recommendations have broadly continued the approach adopted in early versions. Available: https://www.icasa.org.za/AboutUs/ICASANews/tabid/630/post/icasa-notice-of-publication-of-the-draft-call-termintion-regulations-and-an-explanatory-note-to-accompany-the-ddraft-call-termination-regulations/Default.aspx [Site accessed: February 2015].

[8] See, for example, joint interventions by the competition authority and regulators in Zimbabwe: 'Zimbabwe: CTC meets RBZ, Potraz over Econet Inquiry' (10 February 2015), *The Herald*, available: http://allafrica.com/stories/201502100528.html; and Nleya and Robb, 2014.

[9] See, Hawthorne et al, 2014.

of a number of rival providers of mobile money services which is a direct benefit of entry and the competitive process (Nleya and Robb, 2014).

These developments in South Africa and neighbouring countries are not trivial, and result in real benefits to consumers. In the South African case, they are consistent with the objectives of various policy documents and the NDP in particular. However, their attainment in practice relies on high levels of coordination between various government and private actors, focused on a central objective of enhancing rivalry and dynamism in the economy through opening the doors for participation and entry. The following section discussed the theory of barriers to entry, as being central to achieving more competitive markets (and the benefits which can result), in more detail.

3.2 Barriers to entry in theory and practice

We now turn to a more detailed discussion of barriers to entry and their importance in the economy. In economic theory, free entry and exit are important conditions for competition to prevail. When the likelihood of new entry or expansion by existing firms in the market is high, incumbent firms will be constrained by the fear that increased prices would lead to actual or potential rivals expanding output in response to price rises (O'Donoghue and Padilla, 2006). However, if it is difficult, time-consuming or costly for new entrants to come into a market, incumbents may be able to profitably raise prices above cost without a new firm entering the market and driving prices and profits back down. This acknowledges rigidities and the failures of markets. For example, despite the profits available in the market, if there is a high degree of risk involved in entry due to large sunk costs, an entrant may choose not to enter and incumbent firms will be able to continue to charge higher prices and earn positive profits.

The critical issue is what is meant by difficult, time consuming and costly, and why and how this impacts on potential entry. Most jurisdictions around the world require that for entry to be effective in counteracting the exercise of market power by a firm, it must be likely, timely, and of sufficient nature, scale and scope to constrain anti-competitive effects (ICN, 2004; OECD, 2005). This transcends the static analysis of whether barriers to entry exist, and considers a more dynamic perspective of whether entry would occur in a manner that is easy and sufficient to affect competitive outcomes in the foreseeable future. Barriers need not be found to prohibit the entry of rival firms in perpetuity, they must just serve to retard entry in a manner that is sufficient to prevent (potential) rivals from initiating competitive 'best-responses' to the exercise of market power by incumbents. As the EC Guidelines on Article 82 put it:

> *"For expansion or entry to be considered timely, it must be sufficiently swift to deter or defeat the exercise of substantial market power. For expansion or entry to be considered sufficient, it cannot be simply small-scale entry, for example into some market niche, but must be of such a magnitude as to be able to deter any attempt to increase prices by the putatively dominant undertaking in the relevant market"* (EC, 2009:9).

In the long run, in theory all costs can be seen as variable and there are no fixed or sunk costs (O'Donoghue and Padilla, 2005), and hence no barriers to entry. For example, a lease agreement for a fixed time period may represent a fixed cost in the short run, but in the long-run the lease will end and may not be renewed, or it may be possible to sub-let the property.

However, in the short-run, fixed and sunk costs can be very important, leading to scale economies and high barriers to entry.

Where incumbent firms are able to raise prices above competitive levels they have market power in that they are not disciplined by effective competitors. Market power can only be exerted for a significant period of time if there are some barriers to entry, otherwise, high prices will attract new firms to enter so increasing the actual competitive rivalry. Barriers to entry therefore have a close relationship with the level of competition in a market. Furthermore, as Ezrachi and Gilo (2009; 2010) point out, even where barriers to entry are not high, excessive pricing may not necessarily attract entry. A potential entrant will not necessarily consider the monopoly price when deciding whether or not to enter an industry, but will consider what they expect to be the post-entry price charged by the incumbent. The incumbent can be expected to drop its price in response to entry and therefore the authors argue that supra-competitive pricing may be maintained in some cases even where there are not high barriers to entry.

The Competition Tribunal acknowledged the link between barriers to entry and market power explicitly in its recent decision on excessive pricing by Sasol Chemical Industries.[10] It states that:

> *"in dealing with excessive pricing matters, competition authorities are concerned with pricing in markets characterised by high and non-transitory barriers to entry, i.e. where the dominant firm's position is entrenched."* (Tribunal decision pp. 23 – 24)

It goes on to find that where a dominant firm's position in a particular market is not the result of innovation or risk-taking on its part, and its high prices are a result of it taking advantage of its entrenched dominance, this may be an abuse of dominance to the extent that it harms consumers. It also points out that South Africa's history and the objectives of the Competition Act are relevant in this determination. In other words, context and history matter in the determination of whether conduct by a dominant firm is harmful. This coincides with the literature on social orders which acknowledges the differences between countries and the advantages of incumbents, as do policy documents such as the NDP.

From a theoretical point of view, the economic analysis of barriers to entry has a contested history. In the 1940s and 1950s economists such as Bain argued that a large number of potential barriers to entry existed, following the definition of barriers to entry as factors enabling incumbent firms to earn above-normal profits (Harbord and Hoehn, 1994). The factors cited by Bain included, among many others, economies of scale and scope, capital requirements and product differentiation, and Bain saw these barriers as determining the structure of the industry, prevailing prices and consumer welfare (O'Donoghue and Padilla, 2006). According to O'Donoghue and Padilla (2006), Bain's views were criticised on two main fronts. Firstly, the factors cited by Bain are not the only things which determine the number of firms in an industry, and secondly, the factors cited may be considered endogenous in that they can be altered by investment. For example, there are industries where investment requirements and concentration levels are high (and hence barriers to entry would be considered to be high) but which are actually highly competitive.

[10] Competition Tribunal case number: 48/CR/Aug10.

By contrast, Chicago school economists in the 1970s and 1980s put forward a narrower definition of barriers to entry as cost advantages enjoyed by incumbent firms which new entrants would not benefit from (Harbord and Hoehn, 1994). As Stigler (1968:67) put it: *"a cost of producing…which must be borne by a firm which seeks to enter the industry but which is not borne by firms already in the industry"*. What this means in practice is that unless a potential entrant's long-run costs after entry are higher than those of the incumbent, no barrier to entry exists (O'Donoghue and Padilla, 2006). This sets a much higher bar for the characterisation of a barrier to entry. For example, scale economies would not necessarily be considered a barrier to entry if the incumbent faced similar cost dynamics at the time when it entered, even although the first-mover advantage it enjoys as a result of the scale economies may make it very difficult or impossible for a new entrant to compete. In this scenario, entry may well be deterred resulting in higher prices and lower welfare. This highlights the main problem with Stigler's definition, as noted by O'Donoghue and Padilla (2006), which is that in some cases an incumbent may be able to earn supra-competitive rents even if it enjoys no cost advantage over entrants.

Modern industrial organization (IO) theory suggests that a wider definition of entry barriers than that suggested by Stigler is appropriate, but with a more robust theoretical underpinning than the alternative proposed by Bain. It points out that for markets to be considered perfectly "contestable" such that the possibility of entry is enough to keep prices low, very stringent conditions must hold: entry and exit must be immediate and costless and incumbent firms must not be able to respond to entry immediately (Bishop and Walker, 2010). This latter condition is unlikely to be met, as prices in particular can often be changed at short notice, such that an incumbent can price above the competitive level pre-entry and then immediately lower prices once entry occurs (Bishop and Walker, 2010).

Recent theories show that where there are economies of scale, imperfect information and strategic behaviour by incumbent firms, entry can indeed be deterred, leading to a lessening of efficiency and overall welfare. These theories place particular emphasis on the idea of strategic barriers which are those barriers created when incumbent firms use their dominant position to foreclose or exclude entrants in order to undermine competitive rivalry. According to Bishop and Walker (2010), strategic incumbent advantages arise due to timing (a sort of first-mover advantage) and enable the incumbent to "change the rules of the game" in its favour. There are a range of scenarios in which a dominant incumbent has the incentive to behave in this way as will be discussed in more detail below. For example, Rey and Tirole (2006:8) illustrate a range of ways in which incumbent firms may be able to foreclose competitors, with foreclosure defined as:

> *"a situation in which: (i) a firm dominates one market (bottleneck good); and (ii) it uses its market power in the bottleneck good market to restrict output in another market, perhaps but not necessarily by discouraging the entry or encouraging the exit of rivals"*.

In general, barriers to entry can be classified into two categories: structural and strategic. Structural barriers exist because of factors inherent in the nature of the market.

Structural barriers to entry

The first such barrier is sunk costs which are investments which must be made on entry (for example technology, marketing, and research and development) which the investor will not be able to recoup if the firm exits the market. Such costs obviously increase the risk of entry and, as pointed out by Harbord and Hoehn (1994), they also create an asymmetry between incumbent and entrant since once sunk, costs are no longer opportunity costs. Up until the point that the sunk investment has been made, it forms part of the firm's calculation of the return it will make on entering the industry. However, once the investment is sunk and the firm has entered the industry, these costs become irrelevant to the calculation of future returns and hence to the decision of whether or not to stay in the industry. Thus the expected return to an established firm will always be higher than to a potential entrant which has not yet invested in the sunk cost.

High non-sunk costs of entry are not usually considered as a barrier to entry, since such investments are in theory riskless, as they can be recouped if the firm exits. This is a simplistic interpretation of reality, however, as the extent of recoupment that will be possible will depend on the nature of the asset being invested in. For example, machinery is not necessarily a fully sunk cost if it can be sold and at least part of its value recouped if the firm exits. A good example of a fully sunk cost is advertising, since if the firm is unsuccessful and exits, it cannot expect to recoup any part of its advertising expenditure. A further point relevant to the assessment of sunk costs is that capital market imperfections may mean that firms facing a high up-front investment costs struggle to obtain financing due to perceptions of risk, despite the fact that the investment being made is not really sunk in the true sense. Sunk costs may also be seen as a commitment to stay in the industry (Harbord and Hoehn, 1994), but this is more of a strategic consideration and will be discussed further below.

Scale economies also represent a type of structural barrier to entry. A firm enjoys economies of scale in the production (and/or distribution) of a product when its average costs fall as output increases (O'Donoghue and Padilla, 2006). In an industry where the economies of scale are very high relative to market demand, a large incumbent firm will have an advantage over smaller entrants, since a new entrant selling smaller volumes than the incumbent will have higher costs and make lower margins than the incumbent. Scale economies are linked to fixed and sunk costs which are both types of cost which do not vary with production and therefore which imply scale economies. Where high fixed and/or sunk costs are present, the average unit cost of production will fall as output rises, creating economies of scale.

An absolute cost advantage is present where an incumbent firm has a lower cost of production than an entrant, for example because it has preferential access to raw materials or technology (Church and Ware, 2000). This may be due to a historical advantage in terms of geographic location, rights to certain inputs (such as mines) or preferential contracts with input suppliers. For example, O'Donoghue and Padilla (2006) cite various port cases where the incumbent firm's control over the port infrastructure made it an essential trading party in that port such as in Sea Containers/Stena Sealink, where the port of Holyhead was considered to have unique advantages over Liverpool for ferry travel between Ireland and the UK.

If customer switching is low due to high switching costs or brand loyalty for example, then it may be very difficult for a new entrant to compete initially and, if combined with economies of scale, this may imply a period of loss-making for a new entrant. Switching costs can arise from exogenous factors such as a lack of information, learning or transaction costs (O'Donoghue and Padilla, 2006). If products are complex and customers are generally uninformed, then they may be unlikely to switch even if better or cheaper products exist. This may be the case to some extent with bank cheque accounts, where products are typically structured in a complex way that makes it difficult for customers to compare products across banks (although this feature may not be completely exogenous!). Similarly, if a substantial learning or transaction costs are required in order to switch to a new product, customers may be unwilling to do so. For example, number portability makes a big difference to consumers' willingness to switch from one cell phone network to another, as if numbers are not portable, there are high costs involved in changing network as it means having to change your cell phone number. Endogenous switching costs arise from producers' own technological and commercial product choices (O'Donoghue and Padilla, 2006) which can be designed to increase the costs of switching to competing products.

Network effects imply that there are benefits to consumers to purchasing a product which lots of other people also purchase, making products with larger networks of customers more attractive (O'Donoghue and Padilla, 2006; Shapiro and Varian, 1999). This is the case for example with cell phone networks to the extent that on-net calls are priced more cheaply than off-net calls. In these circumstances, customers have a strong incentive to stick with a large incumbent firm rather than move to a smaller network with fewer customers and this can constitute a barrier to entry.

Legal or regulatory barriers may also exist. A distinction can be drawn between regulation aimed at controlling for other undesirable outcomes and that aimed explicitly at regulating monopoly (economic regulation). The former category includes regulation such as licensing and environmental rules. Licensing can raise barriers to entry if it is associated with onerous requirements on prospective licensees or if there are limits to the number of licences that the regulator will grant. Other types of regulation such as environmental rules can also do so, to the extent that such standards do not apply equally or are more costly to meet for entrants than for incumbent firms (O'Donoghue and Padilla, 2006).

Barriers created by economic regulation are an interesting category of entry barriers as they can be influenced by policy interventions and economic regulation is particularly important as it is explicitly aimed at dealing with a lack of competition. There are two ways in which economic regulation can influence barriers to entry. Access regulations seek to ensure that vertically integrated monopolies provide access to essential inputs or facilities to rivals on fair terms. This is usually necessitated when there is a natural monopoly at one level of the value chain but competition is feasible at other levels of the chain. Access regulations which are inadequate or poorly enforced can allow vertically integrated dominant firms to leverage their market power to restrict new entry, helping them to maintain their dominant position. The other major type of economic regulation which can impact entry is price regulation in that it may limit the margins which can be earned by a new entrant, which then reduces the incentive for new firms to enter the market and makes it difficult for smaller competitors to survive. By contrast, effective economic regulation can encourage entry and competition wherever feasible. These concepts will be explored in more detail in review paper 2.

Strategic barriers to entry

Incumbent firms' own conduct may also create barriers to entry, and these are termed strategic barriers to entry. If the entry of a new competitor is likely to reduce the profits made by the incumbent, either because prices fall or its share in total output is reduced, the incumbent may have an incentive to try to deter entry or ensure that it is unsuccessful. This is explicitly recognised in the EC's guidelines on Article 82, which state that:

> *"The dominant undertaking's own conduct may also create barriers to entry, for example where it has made significant investments which entrants or competitors would have to match, or where it has concluded long-term contracts with its customers that have appreciable foreclosing effects"* (EC, 2009:9).

There are a wide range of strategies which may be employed by incumbent firms to these ends. These fall into three main categories: aggressive post-entry behaviour to deter entry, raising rivals' costs and reducing rivals' revenues (Church and Ware, 2000).

Entry deterrence refers to a situation where the incumbent firm employs a strategy in order to make entry seem unattractive to a prospective entrant. This relies on certain assumptions: critically on sunk costs or scale economies and imperfect information (Bishop and Walker, 2010). An incumbent firm faces a choice of whether to accommodate entry or to fight it, and entry is more likely to be deterred if potential entrants expect the incumbent to fight (Bishop and Walker, 2010; Cabral, 2000). However, the incumbent's threat to fight entry may not be credible – if the most profitable strategy for an incumbent once entry has occurred is to accommodate entry, then its threat to fight will not be credible to the entrant and the entrant will enter anyway. However, if the incumbent is able to incur some sunk costs which alter its post-entry payoff such that it enables it to commit to fight, then entry may be credibly deterred (Bishop and Walker, 2010). An incumbent firm may over-invest in capacity (or in research and development and/or advertising) in order to convince a prospective entrant that it will not be able to compete profitably with the incumbent (Motta, 2004; Cabral, 2000). Dixit (1979, 1981) first showed that over-investment in capacity could be used strategically to avoid deter but in fact a combination of sunk costs and scale economies can form a barrier to entry which allows the incumbent to achieve supernormal profits without being more efficient than an entrant (Harbord and Hoehn, 1994).

Similarly, if the entrant does not have perfect information about the incumbent's costs, the incumbent may be able to create a reputation for fighting entry such that entry is deterred, even though the incumbent's best response to entry would have been to accommodate. One such theory by Kreps and Wilson (1982) shows that where there is imperfect information, it may be profitable for an incumbent firm to charge prices which are below cost for a period of time in order to develop a reputation for being "strong" (very low cost) and persuade potential entrants that entry will not be profitable (see also, Cabral, 2000).

Strategic barriers can also arise from the behaviour of firms through practices that raise rivals' costs and/or induce customers or suppliers not to deal with rivals (reduce rivals' revenue). Again there are a number of ways in which incumbent firms can try to create these barriers. They may do so by restricting competitors' access to inputs or to customers. If the incumbent is vertically integrated and has control over an important input, it may be able prevent the competitor from gaining access to a vital input or charge a very high price for it such that the competitor cannot be profitable. Post-Chicago theories have shown that the incumbent firm will have an incentive to do this if there is some reason why it cannot achieve a monopoly profit in the upstream market such as an inability to commit to charge the monopoly price or price regulation, or if it perceives a threat that an entrant may vertically integrate into its monopoly market (Rey and Tirole, 2006; Carlton and Waldman, 1998).

An example of this is the conduct of Telkom which has twice been found guilty of anti-competitive conduct relating to the use of its natural monopoly position in fixed line telecommunications to foreclose downstream competitors by raising their costs.[11] In the first case against Telkom the Tribunal found that it had leveraged its upstream monopoly in the facilities market to advantage its own subsidiary in the competitive value added network services market, causing harm to both competitors and consumers. The Commission later found that Telkom had also engaged in a margin squeeze where it had charged prices for wholesale services to internet service providers which precluded cost-effective competition with Telkom Retail's own services. This is an example of strategic conduct, used to frustrate entry and effective competition.

Another strategy which an incumbent may employ is to tie up key customers into exclusive contracts or provide incentives that an entrant cannot match to prevent customers from switching, such that an entrant cannot acquire sufficient customers to reach an efficient scale of production. Post-Chicago theories[12] show that an incumbent may have the incentive and ability to act in this way where there are scale economies in the industry and where consumers are dispersed, have imperfect information and cannot coordinate their actions (O'Donoghue and Padilla, 2006). An anti-competitive outcome is more likely if the incumbent can contract sequentially with customers and if the contracts are of long duration.

South African Airways has twice been found guilty of this type of conduct by the competition authorities.[13] In the first case, SAA was accused by Nationwide, a competing airline, of abusing its dominance through the incentives it provided to travel agents to sell SAA tickets. Travel agencies received commissions on an incremental and individualised basis that provided them with very strong incentives to sell SAA tickets rather than those of competing airlines, and which a smaller competitor would be unable to match. The Tribunal considered that this would lead to two harms, firstly that consumers in the short run would be flying on more expensive tickets and at less preferable times than if the ticket offering had been unbiased. Secondly, SAA was able to perpetuate its existing dominance by restricting new entry into the market and to inhibit its rivals from expanding in the market. Subsequently, Nationwide and Comair brought a new complaint against SAA for similar conduct and SAA was found guilty for a second time.

Alternatively, the incumbent may make the purchase of a product in which they enjoy market power conditional upon the purchase of the product which the entrant is offering so that customers who want to purchase the monopoly product from the incumbent are forced to also purchase the competitive product from it. An incumbent may employ a combination of these strategies in order to deter or defeat entry. The theoretical underpinnings for this theory are similar to those discussed above in respect of vertical foreclosure, whereby a dominant incumbent may have an incentive to leverage its market power into an adjacent market in order to protect its dominant position in the original market (Rey and Tirole, 2006).

[11] Tribunal case numbers: 11/CR/Feb04 and 016865.
[12] See Rasmusen et al (1991) and Segal and Whinston (2000) for the seminal papers.
[13] Tribunal case numbers: 18/CR/Mar01, 83/CR/Oct04.

Strategic barriers to entry are often relevant where there is a vertically integrated monopolist who has an incentive to protect the rents being earned in the monopoly market through attempting to frustrate entry at another level of the market, such as in the case of Telkom discussed above. This is particularly acute where the incumbent firm or firms have control over key inputs required by entrants. In these circumstances, the incumbent firms may find it profitable to engage in strategies to raise rivals' costs or reduce rivals' revenues. They may also choose to accommodate entrants but to attempt to force them into a particular market niche where they can operate at a smaller scale without threatening the incumbents' main market.

Strategic barriers to entry are considerably more difficult to evaluate partly because their effects first have to be demonstrated and the boundary between fair and unfair competition is a fine line. The assessment of the competitive significance of strategic barriers to entry is also confounded by the fact that these practices can be pro-competitive in some cases by incentivizing investment, for instance. For example, an incumbent has the incentive to invest in additional capacity, technology, or research and development activities if they believe that they will be able to earn a return on those investments in future and not suffer losses due to free-riding. Similarly, advertising can be seen as a pro-competitive strategy by an incumbent firm, however 'too much' advertising is sometimes considered to be a barrier to entry if it 'effectively imposes an obligation on entrants to advertise their products to a similar extent' (OECD, 2005), rather than in proportion to sales. As discussed above, advertising is a sunk cost as its value cannot be recouped when a firm exits.

Ultimately, the expansion of capabilities will occur where opportunities for participation are provided and where dynamic rivalry is able to drive innovation and growth. This in turn relies on the ability of entrants to be successful competitors 'on the merits'. It was argued above that the capabilities of the South African economy were narrowly developed in the past in order to serve the interests of the ruling elite, and have remained so, at least partly due to the lack of progress in tackling entrenched dominance in many sectors of the economy. Here we argue that what is required in order to drive capabilities development and growth in the future is greater access to markets and to opportunities, which implies a reversal of these trends and lowering of barriers to entry. Thus the process of capabilities development is not just about coordination by the state, but rather about ensuring wider opportunity throughout the economy. It is not about 'removing' barriers where the underlying reasons for the barriers are intrinsic features of industries but about positive steps for firms to overcome barriers, most obviously in access to finance, but also in relation to the various structural and strategic barriers which have been described above. A critical element of this intervention is to focus economic regulation on opening up access for entrants to critical inputs and facilities which they need in order to compete successfully, in particular in industries which feature network effects and vertically integrated natural monopolies. This will be a theme of review paper 2.

4. Linking competition, barriers to entry and inclusive growth: a research agenda

It is clear that where entry is effective, and entrants are able to affect market outcomes through competitive processes even in smaller local markets, the welfare of consumers (end-users and intermediate users of inputs) is enhanced. In this regard there is an important role for pro-competitive intervention by designated bodies such as the competition authorities and sector regulators in enforcing measures which reduce barriers to entry. This also ties in with other developmental policy frameworks. For instance, competition policy (in so far as it is based on addressing strategic barriers to participation) is a complementary microeconomic tool to industrial development and policy strategies that seek to address structural constraints to growth. In this regard, addressing the concentrated structure of markets, and developing and supporting the introduction of new firms into various (including new) productive sectors is critical in the South African case, including for leveraging and developing productive capabilities. It is especially important for ensuring dynamism and growth in the long-term. In theory at least, this complementarity arises because of a common objective of changing the structure of markets by encouraging or removing barriers to increased participation – creating more inclusive markets where people have the opportunity to participate in both the process of growth and the sharing of its benefits.

Implicit in the principle of increasing participation (through entry), is the assumption that by doing so, increased rivalry between firms will over time result in both static and dynamic gains from competition. Specifically, firms will not only compete on price to win over customers but will also develop their capabilities through investment and innovation in order to compete on product range, quality and efficiency, which are dynamic gains to society overall. Furthermore, the conceptions of inclusive growth above speak directly to participation as being central to sustainable growth. This entails not only participation in terms of welfare and employment, which are important in their own right, but also access to markets and entry in terms of individuals and firms being able to share in the process of growth.

These concerns are the rationale for this programme of research which seeks to identify and assess the types of barriers to entry, in different markets and industries, and to consider practical policies which can address barriers and open markets up to greater participation. It is important to build a strong body of knowledge and evidence of the actual experiences of firms in the country when entering or seeking to expand their operations in selected sectors of the economy. The framework and methodology acknowledges, however, that each sector may be different in terms of the nature of barriers and the history of competitive interactions between firms. This is a critical aspect of thinking about the appropriate set of sectors to consider. While we have argued broadly in the sections above that markets in South Africa are concentrated with limited opportunity for entry in key sectors, it is important to understand the complexities involved in the dynamics of competition, the nature of competitive interactions, main players, the regulatory history and environment applicable in that sector, as well factors related to specific market outcomes.

Consistent with the literature on developing productive capabilities and aligning economic activity with the capabilities specific to a country, as well as the benefits of new entry and competition, it is necessary to be able to compare and contrast firms, even across sectors, to understand factors and capabilities which are common amongst South African firms, and those which are unique to each sector. These aspects can then be assessed together and across sectors to develop a framework for understanding the dynamics of firm entry in South Africa, and the possible interventions by sector or overall which would enhance this process.

Regulation and policy will fail if it is not designed carefully to fit with specific competitive dynamics in a sector, grounded on a sound understanding of aspects which drive private decision making by firms. In many cases, poor choices in terms of policy or regulation can actually reinforce the market power of incumbent firms (See, Das Nair et al, 2012; and Das

Nair and Roberts, 2014). It is therefore important to assess factors such as the modes and costs of entry in the context of the narrow views raised by Stigler and Chicago School theorists around barriers to entry being those costs borne by entrants or potential entrants which are not borne by firms already in the industry, as discussed above. Furthermore, critical to effective entry is the ability to not only overcome structural barriers, with which government can also assist, but also constraints in terms of obtaining customers, or overcoming long-term vertical arrangements between firms which speaks to the ability to obtains inputs. In some sectors such as in agricultural activities where security of offtake is critical to the farmers, the presence of long-term exclusive arrangements can be beneficial; whereas in others such as in retail there are ongoing debates and competition cases surrounding the duration and exclusivity of arrangements between supermarket chains and property owners in malls which may prevent the entry of rival retail offerings.

In some cases, firms are able to overcome barriers, or those barriers may not be prohibitive perhaps where an entrant has access to substantial financial resources. However the reactions of incumbent firms to their entry may make expansion and growth difficult if entrants are foreclosed from accessing key customers or inputs. This is not to say that incumbent firms are somehow obliged to share customers with entrants, or that they are not expected to react to the threat of entry. Critically however, the reaction of incumbents is required to be fair and not anti-competitive, and should certainly not be based on privileged access to or control of essential facilities or key infrastructure, unilateral abuses of market power, or the ability to lobby and shape regulation in their favour unfairly. It is more beneficial to the economy and consumers, as argued above, if firms engage in rivalry on the basis of improving their prices or product offering, innovating and developing their capabilities in order to maintain and win customers. In turn, entrants would be encouraged to develop their own capabilities (and competitiveness) and adapt their entry strategies towards becoming effective rivals.

It is important to understand these dynamics from both a competition lens, which influences the sectors which are of interest to this research, as well as from a wider lens which considers competition and regulatory issues as forming part of the broader set of policy levers and tools which can used to advance economic and socio-economic goals, including industrial development. This understanding can then be used to think about policy, particular in the context of implementing the objectives of the NDP as discussed above. This is the ultimate goal of the research programme on barriers to entry and inclusive growth.

5. References

Acemoglu, D. and Robinson (2012). *Why nations fail: the origins of power, prosperity and poverty.* Crown Business.

Acemoglu, D., Aghion, P., and Zilibotti, F. (2002). *Distance to frontier, selection, and economic growth*; National Bureau of Economic Research, Working Paper 9066.

Aghion, P., Braun, M. and Fedderke, J. (2008). Competition and productivity growth in South Africa, in *Economics of Transition*, 16: 741-768.

Ali, I. and Son, H. H. (2007). *Defining and measuring inclusive growth: application to the Philippines*; Asian Development Bank, ERD Working Paper No. 98 (July, 2007).

Amsden, A. H. (1989). *Asia's next giant: South Korea and late industrialization*. Oxford University Press: Oxford.

Anand, R. Tulin, V. Kumar, N. (2014). *India: defining and explaining inclusive growth and poverty reduction*, IMF Working Paper, Vol 14, No. 63.

Berg, A. and Ostry, J. (2011). *Inequality and unsustainable growth: two sides of the same coin?* IMF Staff Discussion Note 11/08 (Washington: International Monetary Fund).

Bishop, S. and Walker, M. (2010). *The Economics of EC competition law: concepts, application and measurement*, 3rd Edition. Sweet and Maxwell, 2010.

Boshoff, W. (2013). *Illegal overcharges in markets with a legal cartel history: the South African bitumen market.* Centre for Competition, Regulation and Economic Development Working Paper No. 3/2013.

Cabral, L. M. B. (2000). Chapter 15: Strategic behaviour, entry and exit. *Introduction to Industrial Organization*. MIT Press.

Carlton, D. W. and Waldman, M. (1998). *The strategic use of tying to preserve and create market power in evolving industries*. National Bureau of Economic Research (NBER) Working Paper 6831 (December 1998).

Church, J. and Ware, R. (2000). *Industrial organization: A strategic approach.* New York: McGraw-Hill, 2000. Available at: http://works.bepress.com/jeffrey_church/23

Das Nair, R. and Roberts, S. (2014). *The interface of competition and regulation in energy, telecommunications and transport in South Africa.* Presented at Competition Commission and Tribunal Annual Conference, GIBS, Johannesburg, 4 and 5 September 2014.

Das Nair, R., Mondliwa, P. and Roberts, S. (2012). *The inter-relationships between regulation and competition enforcement: Regulating for competition? Remedying anticompetitive behaviour through regulation?* Presented at TIPS/NERSA Economic Regulators Conference, 21 and 22 August 2012.

Das Nair, R., Mondliwa, P. and Sylvester (2014). *Assessment of the long steel cartel: rebar overcharge estimates.* Presented at Competition Commission and Tribunal Annual Conference, GIBS, Johannesburg, 4 and 5 September 2014.

Department of Trade and Industry (DTI) (2013). *Industrial Policy Action Plan 5: 2012/3-2014/5.*

Dixit, A. (1979). “A model of oligopoly suggesting a theory of entry barriers,” *Bell Journal of Economics*, 10, 20-32.

Dixit, A. (1981). “The role of investment in entry deterrence,” *Economic Journal*, 95, 95-106.

European Commission (2009). *Guidance on the Commission's enforcement priorities in applying Article 82 of the EC Treaty to abusive exclusionary conduct by dominant undertakings.*

Economic Development Department (EDD) (2010). *The new growth path: the framework*.

Evans, P. and Joekes, S. (2008). *Competition and Development: The power of competitive markets*. International Development Research Centre (IDRC).

Ezrachi, A. and Gilo, D. (2009). Are excessive prices really self-correcting? *Journal of Competition Law and Economics*, 2009.

Ezrachi, A. and Gilo, D. (2010). Excessive pricing, entry, assessment, and investment: lessons from the Mittal litigation. *Antitrust Law Journal* No. 3. 2010.

Fourie, F, (2014), *How inclusive is economic growth in South Africa?* Econ 3x3, September.

Gerber, D. (2010). *Global Competition: law, markets and globalisation*, Oxford University Press.

Govinda, H., Khumalo, J. and Mkhwanazi, S. (2014). *On measuring the economic impact: savings to the consumer post cement cartel bust*. Paper submitted for the Competition Commission and Tribunal 8th Annual Conference on Competition Law, Economics and Policy, 4-5 September 2014.

Harbord D, Hoehn T. (1994). Barriers to entry and exit in European competition policy, *International Review of Law and Economics*, Vol: 14, pp. 411-435.

Hausmann, R., Hwang, J. and Rodrik, D. (2007). What you export matters, *Journal of Economic Growth*, 12 (1): 1–25.

Hausmann, R. and Klinger, B. (2008). South Africa’s export predicament. *Economics of Transition*, Vol. 16, Issue 4, pp. 609-637, October 2008.

Hawthorne, R., Goga, S., Sihin, R. and Robb, G. (2014). *Review of the Competition Commission Banking Enquiry.* Centre for Competition, Regulation and Economic Development (CCRED), Working Paper 2014/6.

Hidalgo, A. (2009). *The dynamics of economic complexity and the product space over a 42 year period*; Centre for International Development (CID) at Harvard University, Working Paper No. 189 (December 2009).

Ianchovichina, E., Lundstrom, S. and Garrido, L. (2009). *What is inclusive growth?* World Bank Economic Policy and Debt Department (PRMED) (February, 2009).

International Competition Network (ICN). 'Chapter 5: Assessment of market entry and expansion (Barriers to Entry)'. ICN Report on Merger Guidelines (April 2004). Available online: http://www.internationalcompetitionnetwork.org/uploads/library/doc559.pdf

Khumalo, J., Mashiane, J. and Roberts, S. (2014) Harm and overcharge in the South African precast concrete products cartel, *Journal of Competition Law and Economics*, 10(3), 621-646.

Kraay, A. (2004). *When is growth pro-poor? Cross-country evidence*. IMF Working Paper 04/47 (Washington: International Monetary Fund).

Kreps, D. and Wilson, R. (1982). "Reputation and imperfect information," *Journal of Economic Theory*, Volume 27, Issue 2, August 1982, Pages 253–279.;l

Makhaya, G. and Roberts, S (2013). Expectations and outcomes: considering competition and corporate power in South Africa under democracy. *Review of African Political Economy*, Vol. 40, No. 138.

Mncube, L. (2013). The South African wheat flour cartel: overcharges at the mill. *Journal of Industry, Competition and Trade* 13(4).

Motta, M. (2004). *Competition policy: theory and practice.* Cambridge University Press (New York, 2004).

National Planning Commission (2011). *Diagnostic Overview*.

National Planning Commission (2013). *Our future – make it work. National Development Plan 2030*.

Nleya, L. and Robb, G. (2014) Part two: Mobile money in Kenya and Zimbabwe. Centre for Competition, Regulation and Economic Development (CCRED), *Quarterly Competition Review*, November 2014. Available: http://www.competition.org.za/review/2014/11/7/part-two-mobile-money-in-kenya-and-zimbabwe [Accessed online: February 2015].

North, D., Wallis, J., Webb, S., and Weingast, B. (2007). *Limited access orders in the developing world: A new approach to the problems of development*. World Bank Policy Research Working Paper, WPS4359.

North, D., Wallis, J. and Weingast, B. (2012). *Violence and social orders: a conceptual framework for interpreting recorded human history*. Cambridge University Press, 2012.

O'Donoghue, R. and Padilla, J. (2006). *The law and economics of Article 82 EC*. Hart Publishing.

Organisation for Economic Co-operation and Development (OECD) (2011) OECD-EU *Database on Emerging Economies and World Bank, World Development Indicators.* OECD Divided We Stand, 2011. Accessed at: http://dx.doi.org/10.1787/888932535432

Organisation for Economic Co-operation and Development (OECD). (2005). *OECD Competition Committee Roundtables: Barriers to Entry*.

Paelo, A., Robb, G., and Vilakazi, T. (2014). *Study on barriers to entry in liquid fuel distribution in South Africa*. Centre for Competition, Regulation and Economic Development, CCRED Working paper 2014/13.

Page, J. (2012). Can Africa industrialise? *Journal of African Economies*, 21(2): ii86-ii125.

Ranieri, R. and Ramos, R. (2013a). *After all, what is inclusive growth?* No. 188, International Policy Centre for Inclusive Growth, UNDP.

Ranieri, R. and Ramos, R. (2013b), *Inclusive growth: building up a concept*. Working Paper 104, International Policy Centre for Inclusive Growth (IPC-IG).

Ramos, A.R. Ranieri, R. Lammens, J. (2013). *Mapping inclusive growth*, Working paper 105, International Policy Centre for Inclusive Growth (IPC-IG).

Rasmusen, E., Ramseyer, M. and Wiley, J. (1991). Naked exclusion. *The American Economic Review* 81.5 (1991): 1137-1145.Available at: http://works.bepress.com/rasmusen/7

Patrick Rey, and Jean Tirole (2007). A Primer on foreclosure in *Handbook of Industrial Organization*, Mark Armstrong, and Robert Porter (eds.), North Holland, vol. III, 2007, pp. 2145–2220.

Roberts, S. (2010). Chapter 11: Competition policy, competitive rivalry and a developmental state in South Africa in Edigheji, O. (eds) (2010), *Constructing a Democratic Developmental State in South Africa*.

SBP (2014). *Headline report of SBP'S SME growth index: growth and competitiveness for small business in South Africa.*

Segal, I. and Whinston, D. (2000). Naked exclusion: comment. *The American Economic Review*. Vol 90. No.1.

Shapiro, C. and Varian, H. R. (1999). *Information rules*. Harvard Business School Press.

Statistics South Africa (2015). Quarterly Labour Force Survey, Q4 2014. Available online: http://www.statssa.gov.za/publications/P0211/P02114thQuarter2014.pdf

Stigler, G. (1968) Barriers to entry, economies of scale and firm size in GJ Stigler, *The Organisation of Industry* (Homewood, Irwin, 1968), p. 67.

Sutton, J. (2004). *Competing in Capabilities: Globalization and Development.* Clarendon Lectures in Economics, Oxford, UK: University of Oxford, Department of Economics.

Trung, Q. T. (2008). Performance of export-oriented small and medium-sized manufacturing enterprises in Viet Nam in *Asia-Pacific Trade and Investment Review,* Vol. 4 (2008).