Southern Africa Labour and Development Research Unit

Inequality Traps in South Africa:
An overview and research agenda

by

Miquel Pellicer, Vimal Ranchhod, Mare Sarr and Eva Wegner

ISBN: 978-0-9869892-3-0

Inequality Traps in South Africa:
An overview and research agenda

Miquel Pellicer, Vimal Ranchhod, Mare Sarr and Eva Wegner

1   Introduction

There has been considerable effort in ascertaining with confidence the trends in income inequality in South Africa. South Africa has traditionally been among the most unequal countries in the world and continues to be so. Surprisingly, levels of inequality have not decreased despite the transition to democratic rule in the 1990s; if any, they seem to have increased. There has also been considerable work on the proximate causes of these high levels of inequality on the basis of inequality decompositions (See Leibbrandt, Levinsohn and McCrary 2010, Leibbrandt. Woolard, Finn and Argent 2010, and Bhorat et al. 2009 for recent analyses). However, much less is known about the underlying causes of this high level of inequality and of its persistence.

At the same time, there has been growing interest in the international literature in understanding the determinants of inequality dynamics and, in particular, the reasons for its widespread persistence. In this context, the concept of an "inequality trap" has emerged (see Bourguignon et al. 2007). The concept rests on the idea of multiple equilibria. An inequality trap is a situation where some group of individuals is consistently disadvantaged with respect to another, and where there could potentially be a feasible alternative situation where no group is consistently disadvantaged. Inequality traps can be driven by different types of mechanisms, such as political, economic, etc.

This paper is a first step in exploring the potential of this research framework to shed light on inequality persistence in South Africa. We review the theoretical literature containing models that give rise to inequality traps. We briefly discuss the potential of these mechanisms for South Africa. The aim of the paper is to stimulate research in two directions: Use the framework of inequality traps to understand a particular case of persistently high inequality, South Africa; and use the peculiarities of the South African case to help expanding and refining the framework.

The theoretical models reviewed in this paper (models of of inequality traps) are of a specific type. They address both the causes and the consequences of inequality. Typically, they focus

on some factor that simultaneously affects, and is affected by, the distribution of income, with a mechanism of the following type: Income distribution today $\Longrightarrow$ Some factor $\Longrightarrow$ Income distribution tomorrow. In this way, multiple equilibria can be generated. Low inequality may lead to a particular value of the factor in question which, in turn, regenerates low inequality, and similarly with high inequality. The high inequality case corresponds to the inequality trap. In order to generate these type of mechanisms, models of inequality traps need to consider some sort of interaction between the rich and the poor. It is the interaction between the rich and the poor in some sphere that makes the distribution of income play a key role for the determination of that factor that will in turn determine the relative returns of rich and poor in the future.

It is important to note the type of models that fall outside the scope of this review. First, models of poverty traps alone are not considered. Poverty traps are situations where poverty begets poverty. But in these models, the rich and the poor need not interact. For instance, in a nutritional poverty trap, low wages lead to insufficient nutrition which in turn leads to low productivity and low wages. This is a trap, but not an inequality trap. The behaviour of the rich in this setting does not affect the poor. As we shall see, many models of inequality traps include a poverty trap as an ingredient, but inequality traps require the additional feature that the income distribution determines future relative returns. A second type of factors falling outside the scope of this review are those those that affect inequality but that are not in turn affected by it; i.e. that do not lead to traps. Factors such as skill biased technological change or international trade fall into this category. These factors are commonly considered as causes of the increase in inequality in Anglo-saxon countries in recent decades, but the income distribution has not been identified as a key determinant of them.[1]

The paper is organized on the basis of the different realms where the rich and poor interact. We consider economic, political and social mechanisms. In the economic mechanisms, the rich and poor interact either in production or in markets. In the political mechanisms, rich and poor interact in the determination of policy, for instance by voting for redistribution. In

---

[1]An exception is the model by Bénabou (2005), a model where inequality leads to the adoption of skill biased technology.

the social mechanisms rich and poor interact in the social sphere, for instance through peer effects which leads to education externalities.

To our knowledge, this is the first systematic review of these types of models. Piketty (2000), presents an excellent review of inequality dynamics in the Handbook of Income Distribution, but has a different scope. His review is more general in that he focuses on mechanisms that make income differences, not inequality differences, persist. At the same time we cover other types of mechanisms, such as political mechanisms and also update the literature to include the more recent contributions.

The section on economic mechanisms show that through these type of mechanisms inequality traps can emerge in a variety of ways. First, through educational choices where the rich become skilled and the poor unskilled, and where they interact in production as imperfect substitutes. Second, through the price of a market that has the rich and the poor on different sides: in the labour market, the rich may become entrepreneurs, demanding labour, and the poor labourers, supplying it; in the credit market, the rich may become lenders and the poor borrowers, or vice versa, etc. The mechanism of the inequality trap is essentially the same in all cases. Borrowing and non-convex investment opportunities (for example with a fixed cost, such as the cost of college or the cost of starting a firm) imply that the rich end up in good activities or occupations (undertaking "good" investments). When the rich are scarce (as skilled workers, as entrepreneurs, or as lenders) the returns to the rich "activities" are high and to the poor ones are low: high skilled wages, low unskilled ones (or high profits and low wages for the labour market, or high interest rates for the credit market where the rich are the lenders). This implies high inequality and also implies that the poor remain poor so that the rich remain scarce, thus provoking the inequality trap.

The section on political mechanisms briefly reviews the political economy literature aiming to extract insights on how inequality can persist through pro-rich policies and how the latter come about; it also presents a very exploratory look at some data on South Africa, where available. We discuss three types of potential drivers of inequality traps via the political sphere. First, demand factors implying that the median voter cares about factors other

4

than her current income and thus may vote for low taxes. Seeing indications of very high degrees of over-optimism regarding their children's economic status among South Africans, demand factors promise to be a fruitful avenue of research. Second, factors generating a richer pivotal voter than the median, where, effectively, the poor lose voice in the political process. Although poor South Africans do not appear to abstain more than the rich, the rich may come to dominate politics through other ways, such as lobbying. Third, we discuss the impact of clientelism on redistribution, i.e. models where redistribution does not come through a uniform tax rate, but where it takes a particularistic form, targeted to particular groups or individuals. This research area is to date unexplored in South African politics.

The fourth section presents a brief overview of key social mechanisms that explain how inequality may persist across generations. This strand of the literature is mostly concerned with neighbourhood effects, in particular the effects of one's residential neighbourhood on education and income inequality (Bnabou 1993, 1996; Durlauf 1996; Fernandez and Rogerson 1997, etc.). In those models, a child's education is determined among others by school quality and the characteristics of the residents of his neighbourhood.Parental income matters because it determines the choice of the neighbourhood in which families live. Within this setting, relatively unequal economic status may persist across generations in the presence of economic segregation. If rich families concentrate in neighbourhoods with high quality social interaction (good education, presence of role models, peer effects, etc.) while poor families live in poor neighbourhoods (where education performance is poor and children lack role models), the trajectory of these families is bound to diverge in the long run. The purpose of this section is to discuss some of the important contributions that explore how social mechanisms generate inequality persistence. The common feature of all these papers is that social stratification due to the presence of neighbourhood effects (peer effects, local funding of education, etc.) is a key determinant of persistent inequality.

In the empirical section, we start by summarizing the global empirical research on inequality traps to date. Given that this is a relatively new concept, the literature on this topic is quite sparse. We then focus on what has been done in South Africa. We start by summarizing the large body of literature on measuring inequality levels and trends in South Africa. While

there is considerable debate about the level and trend of inequality in South Africa, the consensus view is that inequality is exceptionally high and has been so for a prolonged period. We next focus on what we know about mechanisms through which the persistence of inequality in South Africa might be explained. These include education, financial markets, intergenerational networks and other possible mechanisms.

The remainder of this paper is structured as follows. Section 2 discusses economic mechanisms from a theoretical point of view. Section 3 considers political mechanisms and Section 4 presents on social mechanisms. Section 5 provides a summary of the related empirical research to date, with a particular emphasis on South Africa. The paper concludes with concrete suggestions of avenues for further research on the topic for South Africa.

2   Inequality Traps: Economic Mechanisms

Inequality traps can be generated via economic mechanisms, i.e. mechanisms arising from the poor and rich interacting in production or in markets. The models that study this type of mechanism share two key elements. First, they all generate a poverty trap. All models assume some form of borrowing constraint and a non-convex investment opportunity (i.e. roughly an investment opportunity that yields good returns only when a sufficient amount has been invested, possibly because of the presence of some fixed cost). The combination of these two ingredients generates a poverty trap because the poor never succeed at investing sufficiently so as to reap good returns from the investment. This is because they are poor and moreover cannot borrow. Thus being poor implies they obtain low returns so that they remain poor, a poverty trap.

A second feature that these models share is some realm where rich and poor meet, the outcome of which determines the future returns of the two groups. Typically the realm consists of a market and the condition that regulates the interaction between rich and poor is a market clearing condition (although in the representative model we consider this occurs

via imperfect substitutability in production). In these models, the distribution of income affects the supply and/or demand of some market, and the price of the market in turn affects the distribution of income. For instance, in some models, the poor end up as wage labourers and the rich as entrepreneurs. The rich form the demand for labour and the poor the supply. The relative sizes of the two groups determine the price of the market, i.e. the wage. And the wage in turn affects the relative returns of the rich and poor, i.e. profits and wages.

It turns out that the mechanism for inequality persistence in all the models considered is surprisingly similar. The relative scarcity of the rich individuals implies, via the market clearing condition (or imperfect substitutability in production), high returns for the rich and low returns for the poor. The poor being poor and not being able to borrow cannot invest sufficiently to obtain good returns, so they remain poor and the rich remain few, so that the cycle continues.

Despite sharing this key mechanism for inequality persistence the models differ substantially in their focus. In particular, they differ in the realm in which the rich and poor interact. The realms of interaction as well as the models addressing each of these are the following:

Production/ Market for skills: Galor and Zeira 1993, Mokherjee and Ray 2002, Ljunqvist 1993, Owen and Weil 1998, Moav and Maoz 2000

Labour market: Banerjee and Newman 1993, Ghatak and Jiang 2002, Matsuyama 2005

Credit market: Aghion and Bolton 1997, Piketty 1997, Matsuyama 2000

Product market: Mani 2001, Matsuyama 2002

In what follows, we first consider a simple illustrative model to discuss the key mechanisms of persistence (belonging to the first type of models, where rich and poor end up with different skill levels and interact in production). Second we consider some extensions of that basic model that yield additional insights and finally we discuss the remaining type of models.

2.1 Simple Illustrative Model

2.1.1 Environment

We consider the simplest possible model that is able to generate an inequality trap via economic mechanisms. The model is a simplified version of Maoz and Moav (2000) and Mokherjee and Ray (2003). Consider an economy with a measure 1 of agents, indexed by i. Each agent is member of a dynasty, lives for one period and has an offspring. At each time t there is one member of the dynasty active. Each period, there is an opportunity in this economy to obtain education. Upon payment of a fixed cost F, an individual obtaining education becomes skilled and receives a wage w_t^H. Agents not obtaining education earn the unskilled wage w_t^L. Every period, there are H_t skilled individuals and $1 - H_t$ unskilled.

Demand for skills comes from a representative firm, that produces using skilled and unskilled labour according to a Cobb-Douglas production function:

$$Y_t = H_t^\alpha \left(1 - H_t\right)^{1-\alpha}$$

Labour markets are competitive, so demand for different types of workers comes from equating wages to their marginal products:

$$
\begin{aligned}
w_t^H &= \alpha \left(\frac{1 - H_t}{H_t}\right)^{1-\alpha} \\
w_t^L &= (1 - \alpha) \left(\frac{H_t}{1 - H_t}\right)^{\alpha}
\end{aligned}
\tag{1a}
$$

Two points are worth noting. First, w_t^H may be higher than w_t^L (as will be the case in equilibrium) even if there is no inherent productive advantage for the skilled. Second, both wages depend on H_t so that, as H_t increases, the skilled workers become more abundant and their wage w_t^H falls, whereas the reverse occurs to the unskilled, and w_t^L rises.

The supply for skills comes from the education decisions of the agents. In the model, education is a parental decision. Agents are born either skilled or unskilled (depending on what their parent decided). They work, earn their corresponding wage, decide whether to send their offspring to school, consume and die. Agents care about their own consumption and about the income of their offspring (gross of education costs), as in the "warm glow" bequest motive. For simplicity, the utility is assumed to be logarithmic and there are no capital markets at all (actually, it is assumed that agents cannot even leave a financial bequest to their offspring). Thus, an agent with wage w^i will send her offspring to school if:

$$Log(w_t^i - F) + Log(w_{t+1}^H) > Log(w_t^i) + Log(w_{t+1}^L)$$

Solving for w_t^i gives the wealth threshold $\tilde{w}$ above which parents educated their children.

$$w_t^i > \frac{F}{1 - \frac{w_{t+1}^L}{w_{t+1}^H}} \equiv \tilde{w}_{t+1} \tag{2}$$

Two points are worth noting. First, it will always be the richer individuals that send their children to school, since the utility sacrifice of paying the fixed cost decreases with income. This is the due to the concavity of the utility function. Second, the threshold $\tilde{w}$ depends on tomorrow's skill premium which, in turn, will depend on tomorrow's scarcity of skills H_{t+1}.

2.1.2 Results

The key equations of the model are 1a and 2. Both relate wages to quantities of skill. The model will be thus discussed graphically in the wage–skill quantity space, as depicted in Figure 1. Since our main focus will be on steady states, time subscripts are omitted.

The figure shows $w^H(H)$ and $w^L(H)$ from production. As mentioned, $w^H(H)$ slopes downward and $w^L(H)$ upwards. Thus, as skills become more abundant, inequality decreases. This is because skilled and unskilled workers are imperfect substitutes in production and

Figure 1: **The Baseline Model**
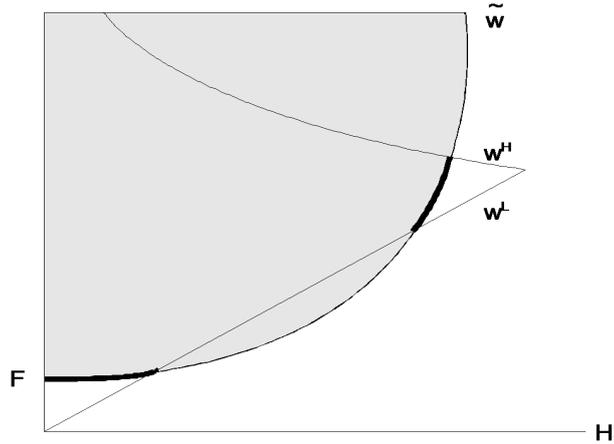
an abundance of skills makes skilled workers less valuable. Eventually the skill premium disappears as H reaches the value α and the two types of workers are equally productive.

The third function shown in Figure 1 is the threshold $\tilde{w}$. Individuals with wages higher than the threshold (shaded area) would educate their offspring. This threshold is increasing in H: as skills become more abundant and the skill premium falls, the incentives to obtain education are reduced. As H grows and the skill premium falls, only richer people are willing to educate their offspring. As the skill premium vanishes the threshold tends to infinity and education is not worth investing in for anyone. Conversely, as H tends to zero and the skill premium tends to infinity, $\tilde{w}$ tends to F: Education would be chosen by all those who can afford it.

The steady states of the model are given by configurations where $\tilde{w}$ lies between w^L and w^H. In that case, at current wages, only currently educated individuals would educate their children. The supply of education would remain the same, in turn giving rise to the same wages. The situation thus perpetuates.

For the particular values of the parameters in the figure there are two sets of steady states. One of them occurs at high levels of H. There, unskilled wages w^L are high but the threshold $\tilde{w}$ is even higher. At those levels of H, unskilled parents are happy to let their children remain unskilled even if they could easily afford to pay the education cost. The key here is the skill

premium, which is very low. The situation remains as a steady state because educated parents, earning slightly more and needing to sacrifice slightly less to pay the cost are willing to educate their children.

A second set of steady states occurs when H is low. There again $\tilde{w}$ lies between w^L and w^H, but the interpretation of the situation is completely different. The skill premium is enormous so that everyone would like to educate their children. However, the unskilled families cannot do so because in a world with borrowing constraints they have no way of paying the fixed cost. The few skilled families, earning very high wages obviously do provide education to their offspring so that the situation remains. This situation corresponds to an inequality trap.

In the example considered in the figure, intermediate levels of education cannot be a steady state. This is illustrated in Figure 2. The w^L line lies above the threshold $\tilde{w}$, implying that at current wages, all the population would like to educate their children. In that case, there is an excess supply of skills and markets adjust wages to make the poor indifferent between educating their children or not. H thus rises and the process continues until we reach the lowest education level among the high education steady states. In this way, the model features a "big push" type of mechanism where a country can remain stuck in an inequality trap. However, if it succeeds in bringing education levels high enough, a virtuous cycle arises where relatively high unskilled wages generate an excess supply of skills. Markets then adjust, reducing the skill premium thereby raising unskilled wages even more.

Several interesting conclusions emerge from this simple model. First, the model does generate an inequality trap, but inequality is largely a by-product. The true driver of the trap is the interplay between scarcity and poverty. In an inequality trap, the scarcity of the rich (of skills) generates inequality and poverty. Poverty, in turn, prevents the poor from taking advantage of the highly rewarding opportunities so that the poor remain poor and the rich scarce. The actual role of inequality in this model is "positive". To see that, consider an exogenous increase in the income of the rich that leaves the unskilled wage unchanged (for instance due to skill biased technical change). What implications would this have on

11

Figure 2: **Dynamics**
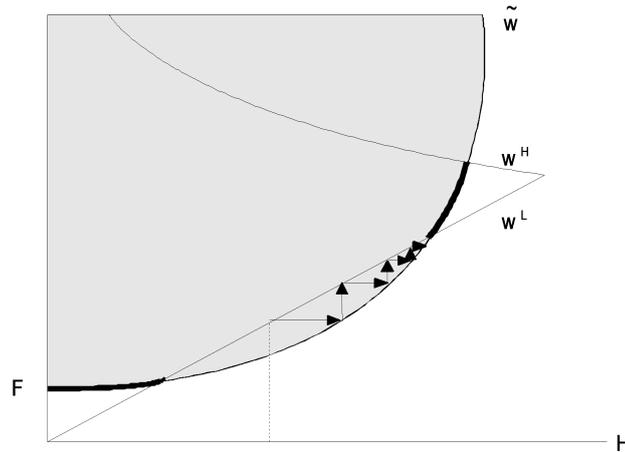
the model? The answer is depicted in Figure 3. The threshold $\tilde{w}$ would shift rightward reflecting the positive incentive effects of a higher skill premium. The region of the trap would shrink so that it would become easier to escape form it. However the effects will be very small. At the inequality trap steady states, inequality is already very high. The "binding constraint" for the lack of education of the poor is the education cost, not the lack of pecuniary incentives from education. An increase in inequality that makes education more attractive would have only a marginal effect on education decisions. The incentive effects of inequality in an inequality trap can hardly be exploited further.

Figure 3: **Role of inequality**
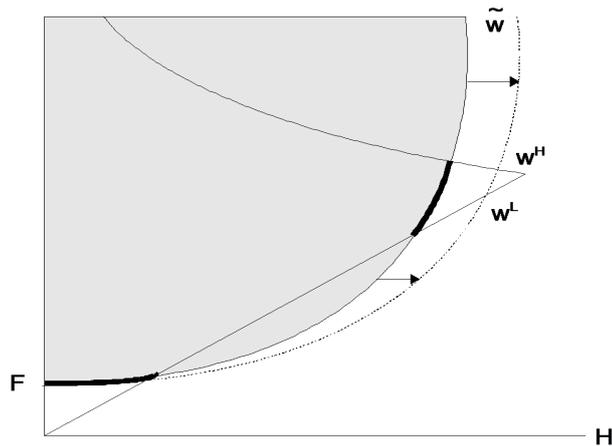
At the same time, that same inequality represents a potentially important source of ineffi-ciency. The presence of large inequality represents very favorable investment opportunities

that are not being undertaken. Efficiency would call for more resources allocated to these opportunities (an increase in education). But this is precisely what does not happen in an inequality trap: the interplay between poverty and a missing market for credit prevents these opportunities for being exploited by a large subset of the population (see Mokherjee and Ray 2003).

A last comment of a more "philosophical" nature is perhaps worth making. The model presented generates inequality even if ex-ante all agents are the same in terms of productive ability. The rich end up being rich, not because they are inherently more productive, but essentially because they are scarce. This scarcity ultimately leads to them being more productive because of the lack of perfect substitutability with unskilled workers. Ultimately, the rich in an inequality trap of this type are rich because the poor are plentiful and because they are too poor to obtain the skills necessary to fully substitute the rich in production.

2.2 Extensions

We will consider two extensions of the model that yield interesting insights into inequality traps via economic mechanisms.

2.2.1 Differences in ability

Individuals may differ in some inherent characteristic that makes them more or less productive, like ambition, motivation, IQ, etc. We consider an extension of the model where individuals differ in a characteristic that we label ability and that is supposed to capture the above mentioned ones. The discussion is based on Owen and Weil (1998) and Maoz and Moav (2000), the latter being the one from whom the modelling strategy is borrowed.

Differences in ability can be parsimoniously incorporated in the baseline model by assuming that ability is reflected in differences in the education cost F, which now becomes indexed by

i, F^i. Suppose there is a uniform distribution of ability, with support $[F_L, F_H]$. The model can be analyzed in a similar way, only that now each individual has a different threshold $\tilde{w}^i$, given by her cost F^i. The highest of these thresholds denoted $(\tilde{w}_H)$ will be the one for the least able people, with costs F_H and the lowest $(\tilde{w}_L)$ for the smartest, with cost F_L. Figure 4 shows this version of the model. The darker area represents wage levels where everyone would educate their children whereas in the lighter area the most able, but not the least able, would.

Figure 4: **The Role of Ability**
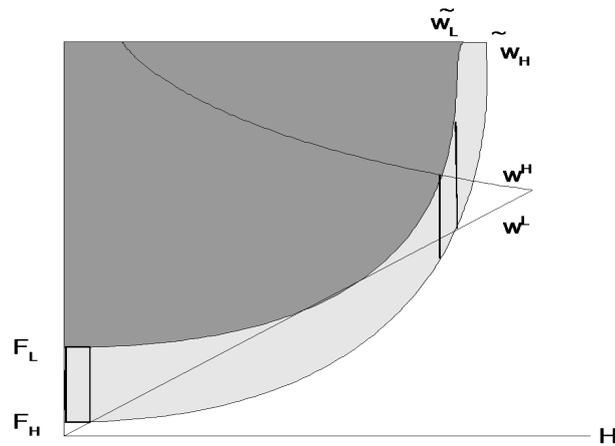
As in the baseline model there are two types of steady states. At low levels of education, there is a steady state where, again, the rich educated parents educate their children and the poor uneducated do not. For sufficiently low levels of H, none of the poor (not even the smartest) send their children to school whereas all the rich (even the least able) do so. Beyond those levels of education, smart poor children start going to school while all rich still do. There is upward mobility but no downward mobility. For high enough levels of education, as the skill premium becomes low enough, rich educated but low-ability dynasties stop finding it worth it to educate their children and downward mobility appears. At even higher levels, the skill premium becomes so low that even the smart kids of uneducated individuals stop finding it worth going to school: upward mobility stops. Between these two levels (where downward mobility starts and downward mobility stops), there has to be some level where upward mobility and downward mobility are equal. That point represents a steady state (corresponding to the high education set of steady states in the baseline model).

14

This extension is important because it adds a novel insight into the inefficiency of inequality traps via economic mechanisms. Inequality traps are inefficient, not only because they leave favorable investment opportunities unexploited, but also because they imply that those that undertake these opportunities are not those that would benefit the most from them. In this model, even if high ability individuals benefit more from education, the criterion for obtaining education is wealth, not ability. Low ability rich children go to school, high ability poor children do not. This compares with the "high" steady state with high education, where schools are composed of high ability individuals, be they rich or poor.

2.2.2 Endogenous education costs

In the baseline model, education costs are fixed. Several studies have analyzed models where these costs are endogenous. There are two main ways to do so, one on the basis of political economy considerations, where the education cost reflects the political decision over the provision of public vs. private education. Second, by considering education as an additional production sector where the inputs are different types of labor and the output is an educated individual. The price of the service is the education cost. Models of this type include Mokherjee and Ray (2003) and Ljunqvist (1993). Here, we adapt the framework of the latter. We discuss the role of endogenizing the education cost by assessing the difference in results it generates relative to the baseline model.

It seems sensible to consider that the education production sector uses skilled labor intensively. The key workers needed to produce education are teachers, who need to be educated themselves. We consider the extreme version where the education sector is competitive and uses only skilled labour. If educating one individual requires γ skilled workers, the education cost will equal γw^H. In this way, the education cost depends, in turn, on inequality.

How does this model compare to the baseline model? The answer is depicted in Figure 5. Suppose γ and F are such that $\gamma w^H (H^*) = F$ at some intermediate H^*. In other words, parameters are such that at some intermediate level of education, the skilled wage is such that

the endogenous education cost is exactly F, the fixed cost in the baseline model. Because the education cost now is increasing in the skill wage, education will be more expensive to the left of that level and cheaper to the right. This implies that the wealth threshold $\tilde{w}$ rotates clockwise around H^*, as shown in the figure. The result is that education costs are higher precisely when they matter the most: when skills are scarce, the skill premium is high and the unskilled wage is low. In contrast, education becomes cheaper where it matters less: when skills are abundant and the poor can afford education.

Figure 5: **Endogenous cost model**
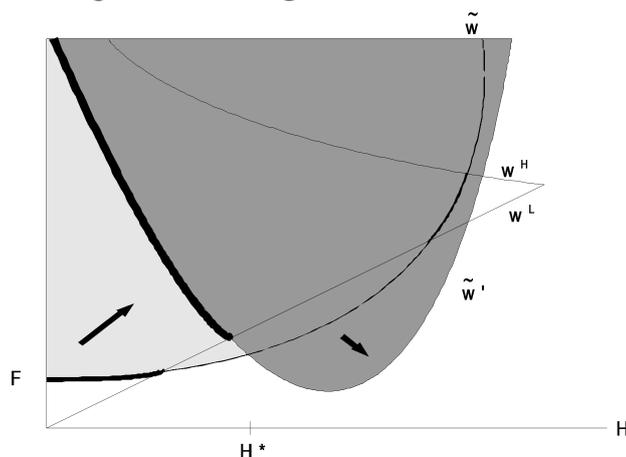
Two points are worth noting of this extension. First, it implies that the inequality trap will be more difficult to exit. As shown in the figure, the inequality trap region becomes larger as education costs becomes higher at low education levels. Second, in this model, inequality matters as such, not just to provide incentives, but perversely, by make it more difficult for the poor to obtain education. If in an inequality trap skill biased technical change would increase the skilled wage, the poor would find it even more difficult to obtain education.

2.3   Other realms of interaction

In the baseline model, rich and poor individuals interact in production. In the remaining economic mechanisms of inequality traps, the rich and poor interact in markets in the sense that they find themselves in opposite sides of some market, the poor in the demand side and

the rich in the supply side, or vice versa. Different models focus on different markets, and interesting insights have been provided by looking at the labour market, the credit market and (in a slightly different manner) the product market. These models are more complex than the previous one in that both sides of the market emerge from occupational choices of agents (as opposed to the baseline model, where labour demand comes from some "firm" established exogenously). As a result, simplifications are made in other respects, such as considering risk neutral agents. Their choices are made on the basis of monetary gains and losses form different options, without considerations for consumption smoothing. This implies that full equality can be achieved where all agents are indifferent between occupations or between being at one side of the market and the other. Most of the models can still be discussed on the basis of the quantity vs. wage/income space. In what follows, we will illustrate heuristically the different models using the type of figure of the baseline model without spelling them out formally.

2.3.1 The labor market

The rich and poor may interact in the labor market. The poor may end up as wage labourers whereas the rich end up as entrepreneurs. This type of situation has been modelled in the seminal work of Banerjee and Newman (1993) and further simplified by Ghatak and Jiang (2002). The latter study provides the basis for the treatment we follow here.

The structure of these models is similar to the one above. Instead of education, there is a generic investment opportunity involving a fixed cost and requiring labour to operate. Those that exploit that opportunity become entrepreneurs, but because of borrowing constraints, only those with sufficient wealth are able to do so. In the model, each entrepreneur uses one labourer and everyone has access to some subsistence technology that provides some floor of earnings (which can also be interpreted as government subsidies in the case of unemployment).

Figure 6 illustrates the model. H in this case is the share of rich individuals (with wealth

17

higher than some threshold) and w^H and w^L represent the profits of entrepreneurs and the wages of labourers, respectively. The threshold $\tilde{w}$ represents the fixed cost of entrepreneurship. When rich individuals make up less than half of the population, there are more would-be labourers than entrepreneurs: i.e. there is excess supply of labour. Wages then need to be kept low in order to make poor people exactly indifferent between employment and subsistence. Wages are thus low and profits high. In the example of the figure, the parameters are such that the threshold $\tilde{w}$ lies between the profits of entrepreneurs and the (subsistence) wages of labourers. The situation thus perpetuates. These steady states correspond to an inequality trap.

Figure 6: **Labour**
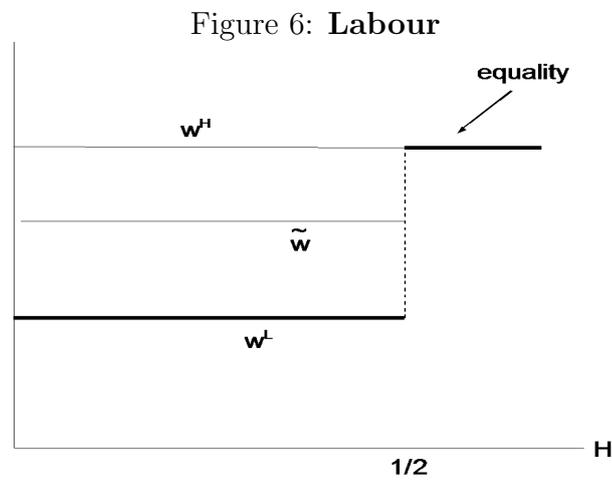
Indeed, there is also an equilibrium with full equality. If there are sufficiently many rich (more than half in this example), there is potentially excess demand for labour. In that case, it is rather the rich that need to be made indifferent between becoming workers and entrepreneurs. Wages will therefore be comparatively high and profits low. In fact, income of the two classes w^H and w^L will be exactly the same and everyone will be indifferent between working or operating a firm.

The key mechanism is thus the same as in the baseline model: Scarcity of rich individuals generates inequality and poverty. Poverty, in turn, prevents the exploitation of profitable opportunities so that the poor remain poor and plentiful, thereby perpetuating scarcity of the rich. Inequality emerges because markets reward scarcity and penalize abundance (in

this case the abundance of labour is penalized by low wages). Inequality perpetuates because the missing credit market makes the previous market rewards be the ones determining that poor and rich remain "segregated" in different sides of the market.

Two extensions of this basic model are worth noting. First, in what actually was the original contribution of Banerjee and Newman (1993), self employment is added as an additional occupational choice. While the key mechanisms remain the same, this addition yields potentially interesting insights regarding, for instance, the size of the informal sector. Self employment is modelled as a investment opportunity requiring a lower fixed cost than full scale entrepreneurship. In that way, in can be interpreted as small scale entrepreneurship in the informal sector. This version of the model can be interpreted as capturing the interaction between poverty and informality. When the poor are very poor and abundant, there will be little self employment/informality. This, in turn, implies that labour supply will be plentiful and wages low, thus ensuring that the poor remain very poor and plentiful. Another equilibrium is possible, however, where an abundant middle class leads to large informality so that the supply of labour is low and wages are high, thereby perpetuating the large size of the middle class.

The second extension (Matsuyama 2005) allows for labour demand from each entrepreneur to be optimally chosen. Labour demand naturally depends on the scale of the firm so that a wealthier entrepreneur will typically demand more labour. In that case, labour demand depends not only on the size of the rich class, but also on how rich they are. A wealthier upper class will demand more labour thereby increasing the wages of workers. This generates a "trickle down" type of mechanism. Under certain conditions, if accumulation is sufficiently rapid, as the rich become wealthier, they pull up the wages of the poor and can eventually bring the poor out of poverty, by making them cross the threshold needed to start a firm and become entrepreneurs. In this case, the inequality trap can be endogenously overcome, for inequality today has a positive effect on the poor in the future.

2.3.2 The Credit Market

Another important market where the rich and poor may interact is the credit market. Depending on the model, the poor end up as the creditors and the rich as debtors or vice versa. This has been studied in the work of Aghion and Bolton (1997), Piketty (1997) and Matsuyama (2000). By focusing on the credit market, these articles address explicitly the microfoundations of capital market imperfections, a key feature present in all the models discussed so far. Because of moral hazard or adverse selection problems, borrowers may shirk and not exert enough effort to make their project succeed or may renege on their debts. These problems are more acute the more a borrower needs to pay back. Thus, the higher the interest rate and the lower the wealth of the prospective entrepreneur, the higher the incentive to shirk or escape. Anticipating this, lenders will thus choose not to lend to the poor, or will lend to them less than what they would want to borrow, and the more so the higher the interest rate. In this way, these type of models naturally endogenize the threshold $\tilde{w}$, separating the choices of the rich and poor. We consider first Matsuyama (2000), which can be discussed naturally within the framework used so far. Second, we comment on some valuable additional insights of Piketty (1997) and Aghion and Bolton (1997).

In Matsuyama (2000), agents again face an investment opportunity requiring the payment of a fixed cost. In this case, the opportunity is profitable and does not run into decreasing returns so that it makes sense to borrow and invest as much as possible. Because of the moral hazard problems just described, lenders only trust (and will lend to) the rich. Thus, the poor end up lending whereas the rich end up borrowing. In the steady state, income determination comes from the interest rate: a higher interest rate is beneficial to the poor (lenders) and detrimental for the rich (borrowers). Market clearing in steady state essentially implies a positive relation between the amount of rich individuals and the interest rate: all else being equal, few rich implies less demand for funds and a lower interest rate. Thus the steady states of the model leads to a graph similar to those used up to now. w^H now represents the income of the rich borrowers and w^L the income of the poor lenders. A larger number of rich individuals is associated with a high interest rate, and thus with lower inequality.

The threshold $\tilde{w}$ is also increasing, although for a different reason as in the baseline model and not necessary in a convex way. The reason in this case is that the larger interest rate associated with a higher H makes the moral hazard problems worse and thus implies that only richer individuals are allowed to borrow to get over the fixed cost and invest. The mechanism is thus essentially similar to the one in the baseline model: scarcity of the rich leads to low demand for funds and a low interest rate which implies that the poor remain very poor and the rich remain few.

Interesting additional insights come from Piketty (1997) and Aghion and Bolton (1997). Both consider situations where decreasing returns eventually limits the profitability of investments so that the very rich become lenders. The poor, in turn, are net borrowers even if they do not succeed in borrowing as much as they would want to or at sufficiently favorable terms. Piketty (1997) emphasizes the potential for inequality persistence. Interestingly the argument here is about the relation between inequality and the severity of credit constraints: scarcity of the rich leads to high interest rates, since the rich are the net lenders; at the same time, high interest rates lead to scarcity of rich, as moral hazard problems make borrowing more difficult so that the poor accumulate so slowly that the rich remain few. Aghion and Bolton (1997), in a similar setting, emphasize instead the potential for "trickle down", as we saw before for the labour market. As the rich accumulate fast enough, their increase in wealth pushes interest rates down so that credit constraints bind less and less for the poor and they eventually exit from poverty.

2.3.3 The Product Market

Finally, the rich and poor may end up in essentially different sides of some product market. When preferences are non-homothetic, expenditure shares are not linear and the distribution of income matters for the demand for different products. If workers of different classes are involved in the production of different products, one may end up with consumer – producer interactions that can generate an inequality trap. Studies such as Matsuyama (2002) and, most explicitly, Mani (2001) consider these type of scenarios.

The model in Mani (2001) can be simplified to be represented in the type of framework considered up to now. One part of the model is exactly as in the baseline model: there is an educational investment opportunity upon the payment of a fixed cost and only individuals who are rich enough can exploit it. The difference here is that the relative returns do not emerge from the demand for different types of labour from a production function, as in the baseline model. Instead, relative returns are generated in the following way. There are two sectors, a manufacturing and a luxury sector. Unskilled individuals work in the manufacturing sector and skilled individuals in the luxury sector, both using a linear production function. The luxury sector is open to trade, so that its price is pinned down in international markets. The manufacturing sector, however, is closed so that the price of the good (and hence the wage of its workers) is determined by domestic demand. Here lies the interaction between rich and poor. The more abundant the rich, the more demand for manufacturing and thus the higher its prices and wages. The resulting graph is essentially as in Figure 1 where w^L is the wage of the unskilled working in the manufacturing sector (in this case the skilled wage w^H would be constant, given by international prices of the luxury good). An inequality trap emerges similar to the ones already discussed.

3 Inequality Traps: Political Mechanisms

This section studies political mechanisms that can lead to inequality traps. In the models that follow, the interaction between the rich and the poor takes place in the political sphere, in most of them in the electoral arena. The section briefly reviews the political economy literature aiming to extract insights on how inequality can persist through pro-rich policies and how the latter come about; it also presents a very exploratory look at some data on South Africa, when available.

We proceed as follows. First, we present the standard model of redistribution in democracies (Meltzer and Richard 1981). On the basis of this model, inequality traps should not exist, given that more unequal countries should redistribute more. South Africa, in turn, with its

high levels of inequality, should be at the top of the redistribution list. Empirical evidence suggests that the relation does not hold neither in general nor in particular for South Africa.

Second, we discuss a number of mechanisms/ factors that could be drivers of inequality persistence via the political sphere. We do so by relaxing different parts of the Meltzer and Richard model. In a first step, we discuss models/ factors where the median voter cares about factors other than her current income. In a second step, we study what happens when the pivotal voter is not the median voter, but a richer individual, i.e. when, effectively, the poor lose voice in the political process. In a third step, we discuss the impact of clientelism on redistribution, i.e. models where redistribution does not come through an uniform tax rate, but where it takes a particularistic form, targeted to particular groups or individuals.

3.1 The Basic Model

The well known basic model addressing the level of redistribution in democracies is Meltzer and Richard (1981, henceforth M&R). This model describes a world where the available policy is a uniform tax, individuals care about their own income, and the enacted policy will be the one preferred by the median voter. Individual demand for redistribution comes from equalizing the costs and benefits of taxation at the margin. The costs come from the disincentive effects of taxation and are the same for everyone. The benefits, in turn, are determined by the position of the respective individual in the income distribution compared to the average. The farther an individual is below the average, the more she will gain from redistribution, the farther above it, the more she will lose from it.

M&R consider a majority voting rule under universal suffrage. In that case, the pivotal voter is the median, and so her preferred tax rate will be implemented. A poorer median voter would thus imply a higher level of redistribution. Typically, a higher level of inequality in a society implies a poorer median voter relative to the mean, so that more unequal societies should redistribute more according to this framework. In short, in the M&R world, inequality traps are nonexistent: higher levels of inequality would lead to more redistribution and no

vicious circles are possible.[2]

3.2   Cross-country data on progressivity

The predictions of the M&R model have long been contested empirically. To date, no solid evidence has been found that more unequal countries redistribute more (see, for instance, Lindert 2004; De Mello and Tiongson 2003). Here, as an illustration of this, and in order to place South Africa into the picture, we show some data on inequality and redistribution for selected countries. The data come from Milanovic (2003) for OECD and Eastern European countries and from van der Berg (2009) for South Africa. Figure 7 plots the pre-tax Ginis vs. the percentage decrease of the Ginis due to taxes and transfers, a measure of progressivity of redistribution.

Figure 7: **Progressivity of Redistribution**
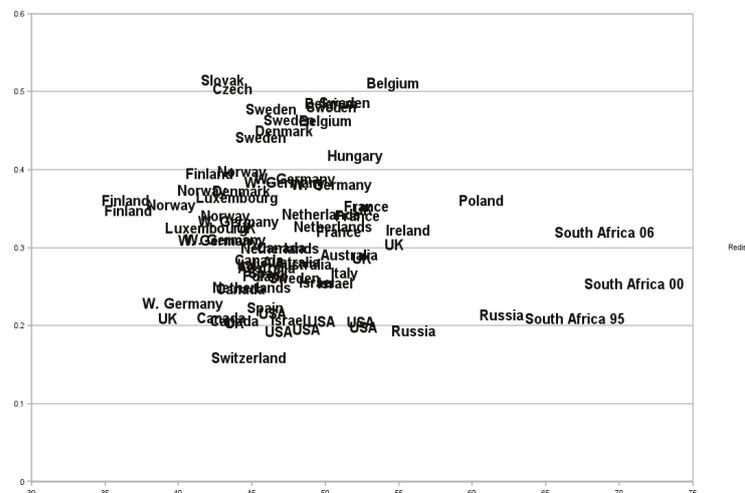
As can be seen in the graph, there is essentially no relation between the level of inequality in a country and the progressivity of taxation: In this sample, more unequal countries do not appear to redistribute more.

[2]Notice that this also holds for models addressing non-democracies. When inequality is high, the poor have more to gain from a revolution and conflict in a society will increase (see, for instance, Acemoglu and Robinson 2000)

What about South Africa? Following the M&R model, we should expect very high levels of redistribution, given its high levels of inequality. Using data from van der Berg (2009), we have plugged in data for South Africa in this graph. While these data are not directly comparable, (the Gini after taxes and transfers includes imputations of indirect transfers, such as education and health), the exercise can be seen as a first approximation of redistribution in South Africa compared to international standards. The data suggest that South Africa reaches, at best, average levels of redistribution.[3] Based on these data, it seems plausible that political mechanisms contribute to a possible inequality trap in South Africa. In the following, we will review a number of factors that could account for "insufficient" levels of redistribution in an unequal country and, where possible, will discuss some descriptive evidence for South Africa from the World Values Survey and the Afrobarometer.

3.3 Political Mechanisms behind Inequality Traps

3.3.1 Demand for redistribution mechanisms

A first type of mechanism that might drive inequality traps relates to the demand for redistribution. While the median voter might still dictate policy, and redistribution might still take the form of a uniform tax rate, voting decisions may be the outcome of factors other than current income.[4] Adding complexity to the individual demand for redistribution can lead to inequality traps if high inequality is associated with an overall depression of demand for redistribution. Several models have been proposed recently in that vein, in particular, addressing the role of effort for demand for redistribution (see Alesina and Angeletos 2005, Cervellati et al. 2007, Bénabou and Tirole 2006). Here, we will focus on Bénabou and Tirole's (2006) model about self-indoctrination and effort.

---

[3]This contrasts with what appears to be the standard wisdom in South Africa, namely that the level of redistribution is very high and that the "problem" is the lack of efficiency in the provision of public services.

[4]There are several excellent general surveys of models of demand for redistribution such as Alesina and Giuliani (2009) and Harms and Zink (2003).

In the Bénabou and Tirole (2006) model, individuals lack willpower regarding effort. The effort they would want to make in the future is higher than the one they will be exerting when the moment comes. Individuals can help solve this problem by investing in optimism. They do this by trying to change the probability that they will "forget" about information that effort might not be crucial in determining income (that, instead, social background for instance might be more relevant). People that have successfully become (over-)optimistic will regard themselves as upwardly mobile and will thus demand low taxes. If many of these over-optimistic people live in a society, there will be low taxes/ little redistribution. In turn, if redistribution is low, low levels of effort (laziness) would pose a severe problem and individuals have high incentives to invest in over-optimism. The reverse story holds for a society with a majority of "realist" individuals who do not forget information about the role of factors other than effort for income. The model provides multiple equilibria: An optimistic low-tax world, which would find itself trapped in high levels of inequality and a realistic high-tax low-inequality world.

The Bénabou/ Tirole model - as well as others addressing the role of effort for the (demand for) redistribution - are typically used to explain differences between the US and Europe, where the US represents the low redistribution/ high effort equilibrium and Europe the high redistribution/ low effort one. How could these stories be applied to South Africa? Starting from the idea that South Africa indeed finds itself in an inequality trap with insufficient levels of redistribution via the political system, it would imply that South Africans are over-optimistic and appreciate effort highly.[5]

We use data from the Afrobarometer and World Values Surveys to evaluate in a very preliminary way the views of South Africans on optimism and effort. The emerging picture is indeed one of over-optimism and a high appreciation of effort. Figure 8 is based on the 2002 (round 2) Afrobarometer. It asks individuals to rank the income they expect their children

---

[5]Notice that this goes against the standard wisdom in South Africa according to which South Africa would rather represent the European equilibrium with high amounts of social grants that de-incentivizes individuals to exert effort.

to have on a scale from 0-10, where 10 represents the rich.[6] As can be seen in the graph, there is a striking percentage of respondents believing their children to be among the rich. More than 30 percent expect their children to have an income of 10, with another 42 percent expecting this income to be in the upper half.

Figure 8: **Optimism**

Cases weighted by adjusted within country weight
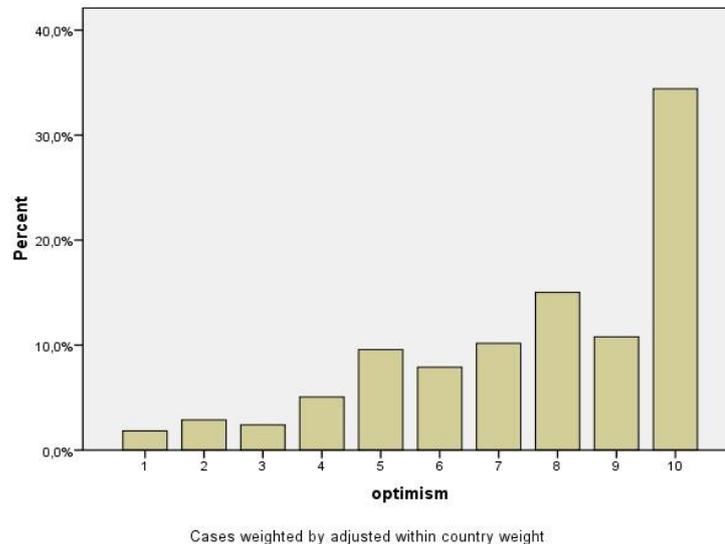
A similar picture emerges when looking at South African views on the role of "hard work" for determining success in life. Compared to other countries, South Africans believe that hard work plays an extraordinarily high role for success. Figure 9 is based on the 2005-2006 World Values Survey. It asks respondents to rank on a scale from 1-10 if it is hard work (1) or luck and connections (10) that determines success.[7] The horizontal axis reproduces the 1-10 scale of these beliefs, the vertical axis shows the percentage of respondents with any of these opinions. The graph compares the beliefs of South Africans with those of US and Swedish citizens, where the US stands for a country where "hard work" for success is emphasized and Sweden for one where the importance of other factors for success is recognized. Indeed the responses of US and Swedish citizens are as expected, with the US curve starting way above the Swedish one with around 20% of Americans believing that it is all about hard

---

[6]The exact wording is the following: "On a scale of 0 to 10, where 0 are "poor" people and 10 are "rich" people: Which number do you expect your children to attain in the future?"

[7]The exact phrasing is: "In the long run, hard work usually brings a better life" (1) and "Hard work doesn't generally bring success—it's more a matter of luck and connections" (10).

work as compared to around 8% of Swedes. Beyond the value of 6, the Swedish curve takes the upper position. The South African curve is astonishing. The data suggest that South African citizens exceed by far US citizens in their belief that hard work determines success: more than 30% of South Africans believe that this is the case.[8] These beliefs are indeed surprising, given that economic success in South Africa has largely been the outcome of apartheid legislation - that is, of political choices - and that race is still an important predictor of income.

Figure 9: **Perceptions about the Role of Effort in Determining Success**

Cases weighted by Weight [with split ups]
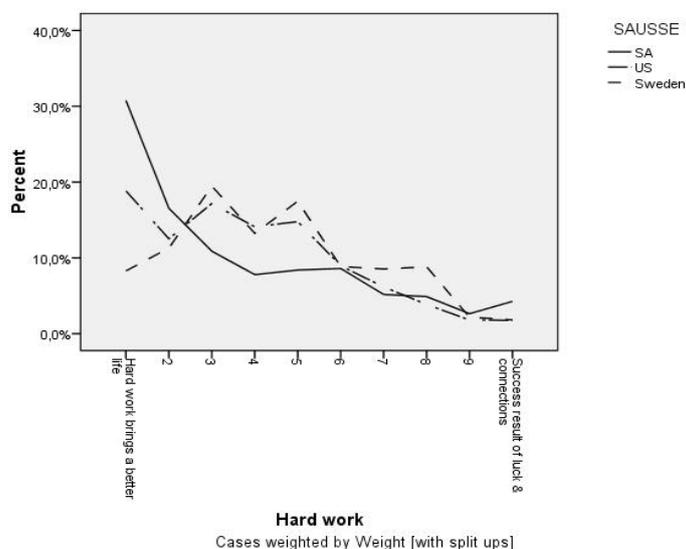
In sum, both sets of data suggest that part of South Africa's inequality trap might come from the demand side. Over-optimism and beliefs that effort will bring about success might impede a demand for redistribution strong enough to break the vicious circle of high inequality leading to low redistribution, leading to high inequality.

3.3.2 A Richer Pivotal Voter

The predictions of the standard model regarding the positive relation between inequality and redistribution hinge critically on the assumption that the median individual dictates

---

[8]The high figure also assures that this does not simply represent the "white" part of the sample, trying to justify income inequality in South Africa.

policy choices. As pointed out in Saint-Paul and Verdier (1996) and Bénabou (2000), if the pivotal voter is sufficiently rich, the standard positive inequality-redistribution relation reverses, with more inequality implying less redistribution and vice-versa. The reason is that, in that case, as inequality increases, the distance of a richer person to the average increases and she will dislike redistribution more. Therefore, there will be less redistribution in more unequal societies and more redistribution in less unequal ones. In this way, if the pivotal voter is not the median, but a richer individual, an inequality trap can emerge. There are, actually, many reasons why the poor(er) may lose voice in the political process and why richer individuals might exert influence beyond their vote. We now turn to some of these reasons.

The poor vote less. The most straightforward reason why the poor may have less voice is that they may vote less than the rich. Indeed, for Western democracies, it is a well documented finding that the poor and uneducated tend to vote less (cf. Sondheimer and Green 2010). A first look at voting behavior in South Africa, however, suggests that this is not the case in that country. Figure 10 is based on data from the 2002 Afrobarometer survey. It plots income in brackets vs. the percentage abstention in each bracket. As can be seen in the graph, the poor do not appear to abstain more than the rich in South Africa. If there is a relation between income and abstention at all, it appears to be rather the rich who do give up their voting rights.

Parties target swing voters. A second factor that may imply that the poor have less impact on policy is that parties may direct their attention to particular segments of voters - who then become the actual pivotal voters. This can be the case if parties believe that some voters are locked in, i.e. are going to vote for them no matter what, whereas others, the so-called 'Swing Voters', need to be convinced. According to this argument, some voters have strong ideological preferences (Dixit and Londregan 1996, 1998; Cox 2010). These are locked into voting for particular parties.[9] Swing voters, in contrast have much weaker ideological convictions and can be enticed to vote for a party by policies such as a tax rate favorable

---

[9]This is an important difference from the M&R framework where voters do not have ideological preferences.

Figure 10: **Income and Abstention in South Africa**
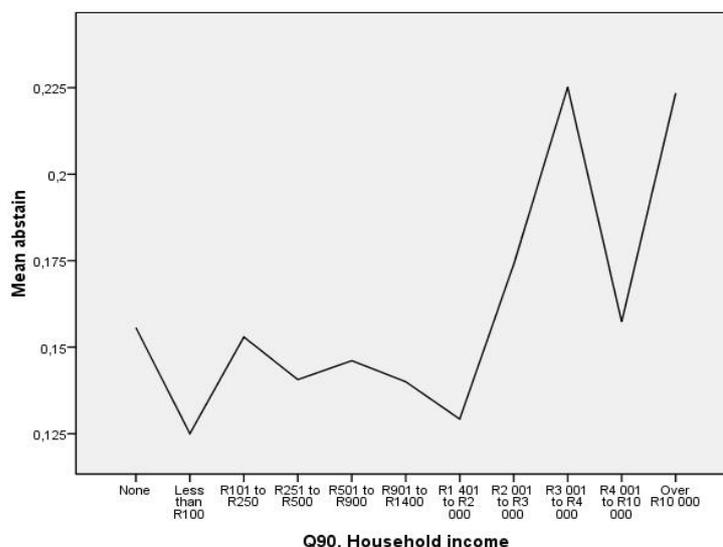
to them.[10] If it were the poor that are locked ideologically, it would imply that the swing voters, the pivotal ones, are richer. For South Africa, a number of questions thus emerge. Is there an inequality trap because the poor are locked into voting for the ANC and thus get less attention by the party? Does the ANC indeed target swing voters and who are they?[11]

Lobbying. There is a large literature on how lobbying distorts electoral outcomes (see for instance, the excellent review on redistribution by Boadway and Keen 2000). In the standard models of lobbying, the policy is the outcome of lobbying efforts of different groups instead of the outcome of voting.[12] The richer pivotal "voter" in this case, is a lobby group that contributes to electoral campaigns or exerts influence on the government. In Becker's (1983) "influence function", for instance, political influence depends on how much money a lobby

---

[10]This can also be done through particularistic transfers to these groups, i.e. club goods, a topic that we address below.

[11]Knoesen (2008) argues that the ANC does not target swing voters but rewards constituencies that overwhelmingly support the party. However, her data - 2000 electoral data to evaluate the reasons behind policies (gas and electricity connections) that were implemented in the 1990s - does not permit this conclusion. It could also be that some constituencies give overwhelming electoral support because they received these new connections, not the other way around.

[12]In this context, parties do not commit to policies they have campaigned on and voters do not have perfect information (Harms and Zink 2003).

spends. Sensibly, the rich will tend to be more successful in their lobbying activities than the poor. This is so for two main reasons: First, because of group size. Being a smaller group than the poor, the rich will overcome free-rider problems more easily than the poor (Olson 1965). Additionally, the smaller group size allows them to distribute higher per capita gains to their members (Peltzman 1976), thus motivating them better. Second, in a world of capital imperfections, they can afford contributions more cheaply (see Esteban and Ray 2006).

3.3.3 Political Clientelism

In the M&R model, redistribution takes the form of a linear tax rate. In many countries, however, redistribution may be targeted to particular groups or individuals. These particularistic transfers can take the form of club goods - for instance, the building of hospitals in neighborhoods that support a particular party or politician, or personal benefits, such as food, a job, or medical assistance in exchange for electoral support. Essentially, clientelism implies vote buying in various ways. It occurs in traditional/ rural settings, where the patron is the local landlord and the client a subordinate living in that constituency as well as in modern ones, where the patron is a political "machine" that employs party brokers to buy supporters.

There is empirical evidence that it is typically the votes of the poor that are bought (Stokes 2005). These votes are cheap because the poor value instant benefits -the transfer from the patron -more than potential public good redistribution later which they might obtain if they voted for a programmatic pro-poor party.

Pellicer (2009) proposes a model where inequality and clientelism feed back into each other, thus creating an inequality trap. In the model, the poor can get organized and implement high levels of redistribution, but this takes time. To prevent that, the rich can provide clientelistic transfers immediately. The immediacy of transfers commands a premium for the poor so that aggregate redistribution is lower in the clientelistic setting. Moreover, the higher

the level of inequality, the higher is this premium, as deeper poverty implies larger marginal utility of immediate income. Thus, whereas low inequality leads to the poor organizing and obtaining high redistribution, high inequality leads to a clientelistic situation with low redistribution, i.e. an inequality trap.

4 Inequality Traps: Social Mechanisms

There is an important strand of literature emphasizing the role of social factors as the main driver of inequality persistence. That social factors play a fundamental role in the perpetuation of poverty and inequality is a standard idea in the social sciences. For instance, sociologists argue that space and community influence individuals' perceptions, aspirations and opportunities. In economics, this strand of the literature is mostly concerned with neighbouhood effects, in particular the effects of residential neighbourhood on education and income inequality (Bénabou 1993, 1996; Durlauf 1996; Fernandez and Rogerson 1997, etc.). In those models, a child's education is determined among others by school quality and the characteristics of the residents of his neighbourhood.[13] Parental income matters because it determines the choice of the neighbourhood in which families live. Within this setting, relatively unequal economic status may persist across generations in the presence of economic segregation. If rich families concentrate in neighbourhoods with high quality social interaction (good eductation, presence of role models, peer effects, etc.) while poor families live in poor neighbourhoods (where education performance is poor and children lack role models), the trajectory of these families is bound to diverge in the long run. The purpose of this section is to discuss some of the important contributions that explore how social mechanisms generate inequality persistence. The common feature of all these papers is that social stratification due to the presence of neighbourhood effects (peer effects, local funding of education) is a key determinant of persistent inequality.

---

[13]Here, a community or a neighbourhood represents a group of individuals who provide and fund education locally to all its members.

4.1 Peer Effects, Local funding of schools and Inequality

The first paper of interest is Bénabou (1996). The author investigates the causes for social stratification (or segregation) and its implications for inequality and productive efficiency. The economy is composed of two groups, the rich R and the poor P who differ by their endowment in human capital ($h^R > h^P$). The city is divided into two communities, one of high quality (e.g. a suburb) and one of low quality (e.g. the inner city). The proportion of rich people is denoted x^S in the suburb and x^I in the inner city. The model has two periods. In period 1, an agent with human capital h consumes, pays his rent ρ and tax $t(h)$ using his initial endowment $\omega(h)$ and debt d. In period 2, the agent works to finance his consumption or bequest and repay his debt $P(h, d)$ so that his income is $y(h) = c' + P(h, d)$. The agent's offspring are provided with human capital $h' = F(h, L, E)$ determined by the parent's human capital h, the quality of social interaction L (local sociological spillover: role model, peer effect, etc.), and spending in education E. One can assume that local neighbourhood spillover is such that: $L'(x) \geq 0$. In addition, in each community, residents fund education through taxation.

In the equilibrium an agent maximising his utility $U(c, c', h')$ will choose to live in a segregated community if:

$$R_x(h, \rho, x) \equiv \frac{d\rho}{dx} \text{ increases in } h \ \text{ OR } \ R_{xh}(h, \rho, x) = R_x(h_A, \rho, x) - R_x(h_B, \rho, x) > 0 \quad (3)$$

In other words, equilibrium in the housing market results in stratification as long as the rich (in human capital) are willing or able to pay a larger rent premium to live in the suburb than the poor. If the proportion of the rich in human capital was even slightly higher in the suburb, i.e. $x^S > x^P$, the suburb becomes more desirable due to its better social environment. Moving to a better community, however, comes at a price captured by the rent premium between the two communities $\rho^S - \rho^I$. As the rich are more willing to pay for this premium to benefit from a better environment, stratification of the communities starts until one of them becomes completely homogeneous.

Various forces combine to drive the process of social stratification. We focus only on two of them. The first and most important determinant is what Bénabou (1996) calls *local complementarities*. Stratification occurs whenever there is complementarity between neighbourhood quality L and an individual adult's human capital h. That is, when rich families value more the effect of neighbourhood quality on their offspring's education than poor families do, i.e. $F_{hL} = F_L(h^R, L) - F_L(h^R, L) > 0$. Second, credit market imperfections may play an important role. In the presence of credit market imperfections, the cost of contracting a debt is higher for poor families so that rich families will find it easier to borrow money to move to a quality neighbourhood thus displacing the poor to lower quality neighbourhoods. Or equivalently, poor families that value education will be prevented from moving to a quality neighbourhood because of the high cost of borrowing.

The process of stratification compounds parental disparities in human capital h with neighbourhood/social disparities L, which results in persistent inequality. The existence of local increasing returns of offspring education $F_{hL} > 0$ is the key mechanism through which income inequality persists. When the marginal returns of parental education increases with the quality of the neighbourhood, the rich will prefer to isolate themselves in suburbs. As a result their offspring's human capital will increase at an increasing rate widening the human capital inequality between rich and poor. This divergence in human capital between inner city and suburb in turn translates into income inequality persistence.

Besides being potentially unequal, the equilibrium is also likely to be inefficient. Efficiency of segregation depends on the trade-off between *(i)* local complementarities between families' human capital and local social spillover, i.e. $F_{hL} > 0$; *(ii)* the decreasing marginal productivity of education with respect to local spillover, i.e. $F_{LL} < 0$; and *(iii)* the relative contribution of the rich to each community: $L''(x) < 0$ indicates that the quality of the neighbourhood increases with the proportion of rich people x at an decreasing rate. In other words, the contribution of the rich to the neighbourhood quality is greater in the inner city than in the suburb. In moving to a suburb, rich families take into account their private benefits from local complementarities between education and neighbourhood quality. However, they ignore the negative external effects their departure imposes on the inner city

(deterioration of the quality of the inner city).

Bénabou (1996) moreover shows that the existence of a decentralised (local) funding of education tends to amplify the effect of the local spillovers discussed above. Whenever education expenditures and local spillovers are complementary, these expenditures tend to increase with the proportion of rich people in a given neighbourhood. The presence of complementarity between parental education and community quality together with the incentive for agents to move into areas with people of similar preferences[14] generates stratification even when the marginal product of community quality is greater for the poor relative to the rich. Such stratification tends to be inefficient since the net private benefit ignores not only the external effect of the deterioration of the inner city but also the fact that neighbourhood quality L and education expenditures are potentially substitutes, i.e. $F_{EL} < 0$. In such a case, one dollar spent on education in the inner city raises educational outcomes more than in the suburb.

Hence local complementarities (direct peer effects), or local funding of schools (fiscal effects) combined with the housing market are the main drivers for segregation and inequality persistence in this model. This paper has a number of distinctive features. Agents are heterogeneous and differ in their human capital endowments. Credit markets are imperfect although they are not necessary for stratification. However these imperfections compound the effect of local complementarities.

4.2 A model of inefficient segregation

Bénabou (1993) shows that these features are not necessary for inefficient segregation to occur. In his paper, *(i)* agents are identical: there is no heterogeneity in ability, or in endowment; and *(ii)* credit market imperfections are absent. He models how peer effects (education spillovers) affects the composition of skills leading to segregation as high skilled people benefit more from education spillover the higher their number.

---

[14]This allows individuals to be able to set education policies as close as possible to their own preferences.

The paper tries to explain the effect of social stratification on efficiency in production when skilled and unskilled labour are complementary and education is a local public good. The agents in this economy live in a city and choose the community or neighbourhood they want to live in given the cost of rental, local education cost or benefit. They also choose between three occupations: skilled labour (e.g. managers and professionals), low-skilled labour (e.g. assembly line workers) and unemployment. Skilled and low-skilled labour are complementary in production. Production is realised at the city level (global level) while education is a local public good that is provided and financed at the community level.

Skilled labour comes with high wages w_H and high cost of education C_H, while low-skilled labour faces lower wage w_L and lower cost of education C_L. It is assumed that these costs decrease with the proportion of skilled labour x present in a given community: there exists positive education spillover (e.g. peer effects in education) due to the presence of high-skilled workers, i.e. $C'_j(x) < 0$ with $j = H, L$. In addition, the education externality is greater for high-skilled workers than for low-skilled ones. In other words, people investing in high skills benefit more from education spillover than those investing in low skills i.e. $C'_H(x) < C'_L(x)$. Suppose that all agents have initially the same characteristics and that there is no unemployment. The author analyzes how the asymmetric effect of local education spillover between high and low skilled workers drives the endogenous determination of the distribution of occupations, residential locations and land rents.

When agents have identical human capital, complementarity in production requires that some individuals will invest in high skills and others in low skills so that all agent are indifferent between the two occupations i.e. $w_H(x) - C_H(x) = w_L(x) - C_L(x)$.

When education is a local public good financed at the community level, and complementarity in production takes place at the city level, segregation arises whenever $C'_H(x) < C'_L(x)$. If the proportion of the high skilled population is even slightly larger in the suburb, $x^S > x^I$, then people investing in high skills move to the suburb to benefit from more education spillover. This migration increases the rent premium $\Delta\rho = \rho^S - \rho^I$ in the suburb. This process continues until one of the communities becomes perfectly segregated. The segregated equilibrium

is inefficient when the social returns to the concentration of high skilled individuals are decreasing. That is, when the increased costs incurred by those remaining in the inner city as high skilled workers migrate outweighs the reduced costs enjoyed by the high-skilled workers who migrate to the suburb. The outcome is inefficient because those who leave the inner city are only interested in the private benefit from education (due to $C'_H(x) < C'_L(x)$) and do not internalise the social cost that their departure imposes on the residents of the inner city. More generally the convexity of the total cost of education $\Phi(x) = xC_H(x) + (1-x)C_L(x)$ is the main culprit for the inefficiency of the equilibrium. This is because when $\Phi(x)$ is convex, stratification increases the total education costs since the external marginal loss to the inner city outweighs the external marginal benefit that accrues to the suburb. Because the external effect of higher levels of stratification on the low-skilled people is not captured by the market system in the form of a higher rent to the high skilled people who are moving to the suburb (rent price discrimination), the equilibrium is inefficient.

So far, it has been assumed that employment is always preferable to unemployment in that the cost of effort to obtain a low level of education is not too high. Suppose now that, as a result of the stratification process, the inner city is only inhabited by a low skilled population. Suppose also that the returns to low skilled education are negative so that unemployment is preferred to a low skilled occupation. Then the inhabitants of the inner city are better off dropping out of the labour market. The city-wide contraction in the supply of low skilled labour then raises the wage for the low skilled occupation. Therefore there is an incentive for more suburb dwellers to take on low skilled jobs to increase their wages. The process continues until the labour market clears. However, as a result of segregation, output can only be produced by individuals living in the suburb, which results in a significantly reduced level of production and welfare. Thus, whenever unemployment is allowed, segregation can worsen the outcome even further and becomes self-defeating by creating unproductive ghettos, which has a negative impact on the overall productivity of the city.

4.3   Social stratification and inequality dynamics

The models discussed above are not concerned with dynamics. Rather, they may be seen as describing long run steady state equilibria. To the extent that shocks to human capital or income may have long-lasting effects, both initial conditions and transition paths do matter. Durlauf (1996) provides a model where the long run distribution of skills and income are path-dependent. He develops an overlapping generations model in which in each family i, parents allocate their income y_{it} between consumption c_{it} and taxes T_{it}. Each neighbourhood sets its own tax rate proportional to families' income levels according to the median voter rule in order to finance the education or human capital investment h_{it} of their offspring. The education level is also chosen at the neighbourhood level so that all children in a given neighbourhood enjoy the same education irrespective of their parents' income. The existence of a proportional tax system implies that, in an integrated neighbourhood, the rich families subsidize the education of the poor. As a consequence, the rich may have incentive to regroup in segregated neighbourhoods to avoid this redistribution. Since the per capita cost of education $C(H_{it}, n_t)$ decreases with the number of families n_t living in a given neighbourhood, an integrated neighbourhood may nevertheless be appealing to the wealthy if they benefit sufficiently from economies of scale.

When human capital is a local public good and productivity shocks depend both on family income and the neighbourhood income distribution (this is similar to the Bénabou (1993) complementarity assumption), the endogenous stratification of the economy becomes a plausible outcome when the rich set income barriers to entry to prevent families with un-desired characteristics to move to their neighbourhood. This entry requirement is essentially equivalent to housing prices (Bénabou 1993 and 1996) combined with zoning rules (Fernandez and Rogerson 1997). The rich face the trade-off between *(i)* living in an integrated community with low per capita cost of providing education; and *(ii)* isolating themselves in a homogeneous community to avoid subsidising the education of the poor, or to minimise the likelihood of negative neighbourhood shocks. When the latter dominates, rich families end up in quality neighbourhoods, insuring their offspring against negative shocks (and therefore

downward mobility) by a continuous and increasing investment in human capital in each generation. In contrast, poor families end up in poor neighbourhoods and are unable to fund adequately the education of their offspring because of liquidity constraints. This results in low levels of human capital and therefore low future incomes as the occupational choices available to the poor are constrained by their education level. In the long run, this phenomenon may generate poverty traps and increased inequality as the gap between rich and poor neighbourhoods widens. The poor will be trapped in neighbourhoods with low or no prospect of social mobility, where role models are conspicuously absent. The process of homogenisation may ultimately result in persistent inequality.

5 Inequality Traps: An empirical review

Following on from the overviews presented in the previous sections, we turn our attention to what empirical literature exists for the existence and measurement of inequality traps.[15] We first focus on what has been done elsewhere, and then consider specifically the South African environment. As emphasized in the theory, an inequality trap is not the same as a poverty trap, although the two do coincide. Moreover, a specific requirement of an inequality trap is that there is something about inequality itself which leads to dynamic stability in the degree of inequality, even though lower levels of inequality may be sustained in equilibrium.

To empirically identify these characteristics is challenging. One would need to start with existence. Are we currently in a state of equilibrium in terms of inequality? This is very hard to distinguish from a state of disequilibrium with a slow convergence rate. Is there a lower level of inequality that could be sustained in equilibrium? Since we cannot observe counterfactual possibilities in the data, we would need to appeal to comparisons across some combination of time and regions. In terms of actual inequality traps, we would need to identify observable mechanisms that both arise due to inequality, and at the same time

---

[15]For the remainder of this paper, the term "inequality" is used to refer specifically to "income inequality" unless specified otherwise.

perpetuate inequality.

Given the empirical challenges involved with this, both in terms of data available and econometric identification, it is not surprising that the area has not been thoroughly researched in South Africa. Indeed, it is unclear whether the questions of existence are even answerable, regardless of the country under consideration. Thus far, there are only a small number of papers that have seriously attempted to identify this, and it remains unclear whether these have been methodologically satisfactory yet. This being said, the challenges are somewhat moot. Based on all of the evidence that we have, South Africa is a highly unequal society in terms of its income distribution, and it has been so for as long as we have decent national household level income data. It would also be safe to assume that it was a highly unequal society for several decades, if not centuries, prior to the collection of such data. This is sufficient to motivate our research.

The primary focus of this section is to identify causes and consequences of income inequality in multiple dimensions, and where possible, to ask which of these possible dimensions fulfill a dual role of being both a cause and a consequence. We summarize the related literature below and then list a number of avenues which we believe might be useful paths of inquiry.

5.1   Evidence from other countries

One recent paper that posits an empirical test of the inequality trap concept is a short article by Daymon and Gill (2009). To our knowledge, it is the only published research that proposes a formal econometric test for the existence of an inequality trap. The authors use a GMM estimator and use time series variation over 71 countries between 1963 and 2003. Of these, 40 are emerging economies. In essence, they estimate whether the inequality dynamics in countries over time is a function of the initial conditions of that country. That is, they regress the inequality measure at a point in time on lagged values of inequality and a host of other variables for a set of countries. These other variables include measures of political liberty, access to credit, gender inequality, youth literacy, population growth and

infant mortality. Note that this method does not account for an individual's position within their country's income distribution. Their findings are that inequality persistence is a global phenomenon, and is more pronounced in emerging economies. Of the additional variables, the most significant are domestic credit markets, infant mortality (which they use as a proxy for overall access to healthcare), the population growth rate and the youth illiteracy rate.

A second paper which focuses on Mexico is by Guerrero, López-Calva and Walton (2006). Their paper is more concerned with how powerful interest groups, in particular concentrated wealth in the business sector and unions in the labour market, generate inefficiencies in the economy. Their argument essentially involves using multiple examples to support the view that the concentrated power leads to influence over policy in ways that both reproduce the inequality and limits growth. Some of the mechanisms through which this occurs include anti-competitive behaviour on the part of firms and low quality educational outputs due to inefficient performance incentives for unionized teachers.

A third useful document is a proposal for empirical research by Cruces et al (2010), in which they discuss how they will attempt to identify inequality traps in Latin American Countries (LAC). The authors start by tightly defining what they mean by an inequality trap, and emphasize that inequality in outcomes is not sufficient for identifying a trap as a trap must also lead to persistence of inequality over time. Instead, they focus on inequality of opportunities, which do have an effect on outcomes. Thus, they argue that one needs to observe both opportunities and socioeconomic mobility, i.e. outcomes, in order to be able to identify inequality traps. They expand further by providing various different measures of opportunity, as well as the datasets and variables they will use. Even though there are no results per se, this proposal provides a succinct summary of challenges for empirical work in this field, as well as practical ways forward that have been suggested to date.

We next turn our attention to the South African literature. It is worth stating up front that no researcher has yet attempted to identify whether or not South Africa is experiencing an inequality trap within the framework discussed.

41

5.2 What do we know about the levels and trends in inequality in South Africa?

To date, a large body of literature has evolved primarily concerned with measuring inequality. This is sometimes done in a static sense, and sometimes in a relative sense to obtain a measure of the direction, degree and rate at which inequality has been changing.

One of the papers is by Leibbrandt, Bhorat and Woolard (2001), who make use of data from the Income and Expenditure survey of 1995 (IES95). They also compare their results to analogous estimates from the 1993 PSLSD survey. The authors analyse household level income inequality using a variety of techniques. They first consider the importance of race in understanding aggregate inequality in South Africa, using a variety of measures that allow for the decomposition of aggregate inequality into within race and between race components. While the magnitude of the relative contribution of these two components is ambiguous depending on which dataset and which measure is being used, the contribution of between race inequality to aggregate inequality is very high by international standards. The authors then focus on possible mechanisms that might explain the high levels of observed inequality, namely the labour market, asset ownership and welfare. Of these, the labour market, and in particular unemployment, drives inequality the most. Wage income contributes about 67% of inequality, and almost half of this is driven by households with no wage income. At the same time, state transfers make a negative contribution to poverty.[16] The overall point, after much investigation, is that wage inequality is by far the major contributor to income inequality, and that an important factor to consider is unemployment/non-employment.

A second important paper in the measurement of inequality in South Africa is the one entitled "Not Separate, Not Equal: Poverty and Inequality in Post-Apartheid South Africa", by Hoogeveen and Özler (2005). In essence, they take the paper by Leibbrandt et al forward by comparing income and expenditures from the IES95 to the IES00, which was conducted

---

[16]Note that in 1995, a number of current grants, particularly the Child Support Grant, had not yet been introduced.

in 2000. In addition to the IES data, they merge the IES95 with the corresponding October Household Survey (OHS) from 1995, and the IES00 with the September 2000 Labour Force Survey (LFS). They find that inequality increased during this period, mostly due to an increase in the inequality measured within the African subpopulation. For example, for the entire sample the mean log deviation increased from 0.56 to 0.61 during this period, while the Gini coefficient increased from 0.565 to 0.577. For Africans the mean log deviation increased from 0.37 to 0.436. The share of between group inequality decreased from 38.3% to 33.2%.

The authors also document that poverty, especially extreme poverty, increased. They then investigate how much of the observed changes are due to changes in endowments as compared to the price of endowments. One observation they make is that the returns to education increased particularly for Africans with high levels of education. They posit that this is a major component in understanding the increased inequality within Africans, since the increase in educational endowments amongst Africans was relatively small. This in turn explains the observed change in overall inequality. One of their concluding policy recommendations is to 'focus on improving quality educational attainment for the poor'.

There are several other papers that have attempted to measure poverty and inequality in South Africa. These include the papers by Leibbrandt, Levinsohn and McCrary (2010), Ardington, Lam, Leibbrandt and Welch (2005), Leibbrandt, Woolard, Finn and Argent (2010), van der Berg and Louw (2004), van der Berg, Louw and Yu (2008) and Yu (2010), and even this list is not exhaustive. They differ in terms of the datasets employed, the assumptions underlying how to deal with missing data and whether the analysis is done at the household or individual level. In terms of findings, there remains debate about the actual levels of inequality and the rate at which it is changing.

One of the more recent papers on inequality is by van der Berg (2010). He succinctly summarizes the state of the literature in South Africa as follows: "Thus there was probably a strong upward trend in inequality as measured by the Gini coefficient in the second half of the 1990s, and largely stable inequality since. Inequality is clearly very high, but how high

43

is not clear due to data comparability and measurement issues."[17]

For our purposes, this is a sufficient statement to motivate out research agenda. The overall point is that inequality was very high, and has remained so or even increased over the past fifteen years. We turn next to consider what research has been done to understand why inequality in South Africa has been so persistently high.

5.3 Mechanisms through which inequality traps might develop

5.3.1 Education

As discussed in the theoretical section, one mechanism through which inequality might become persistent would be that only a few wealthy individuals obtain high levels of education, and this scarcity of skills subsequently generates very high rates of return for these highly skilled individuals. This generates a persistent level of inequality.

Several researchers, most recently van der Berg (2010), point to the importance of the labour market and educational attainment as an important factor in understanding South African inequality. He also stresses the importance of the quality of education as being relevant. Lam (1999) finds that highly convex returns to schooling are such that modest improvements in educational attainment are unlikely to significantly affect the overall inequality distribution. Hoogeveen and Özler (2005) and Leibbrandt, Levinsohn and McCrary (2010), at the household and individual levels respectively, both attribute the widening inequality between 1995 and 2000 to increases in the rates of return to education.

Thus, a large part of the inequality that exists could be attributed to inequality in educational attainment. For example, the 2008 wave of the National Income Dynamics Study shows that there are relatively few skilled people. Only 4.3% of adults have a Bachelors degree or higher and only 13.9% have some form of post secondary schooling. This is also reflected

---

[17]Pp. 12

in differences in educational attainment by race. Amongst whites, 19.4% have a Bachelors degree or higher, and over 40% have some form tertiary training, whereas the corresponding numbers amongst Africans are 1.8% and 9.1% respectively.

Keswell (2004) shows that while in 1993 the returns to education were the same for both race groups, by 2002 the returns to education for whites had increased substantially relative to Africans. He points out that this may reflect unobservable differences in educational quality conditional on attainment by race or occupational stratification. Thus, even though the least educated race group has experienced an increase in mean years of schooling between 1993 and 2002, both in absolute and relative terms, the change in their rate of return does not strongly affect overall inequality in earnings.

While Keswell is estimating returns per year of schooling, regardless of the level of schooling, Branson, Leibbrandt and Zuze (2009) focus specifically on tertiary education. When focusing on tertiary qualification over the period 2000 - 2007, they find consistently strong returns to education in both employment rates and earnings. They argue that there are strong incentives to attain a tertiary qualification if possible, but that the ability to do so is constrained. Some of these constraints are financial, while some occur through a process of inadequate preparation at prior schooling levels.

Note that this is by no means an exhaustive survey of the economics of education literature in South Africa. What it suffices to point out is that:

- There are barriers that prevent historically poorer groups from attaining high levels of education,

- There is a relatively scarce supply of highly educated workers in the labour force, and

- There is a large rate of return to tertiary qualifications.

Within the theoretical framework presented above, this has all of the markings of an inequality trap.

5.3.2 Financial Mechanisms

The second broad group of models discussed involves financial markets in some sense. The poor are either credit constrained which prevents them from investing in otherwise profitable investments, due to a lack of collateral or other characteristics that make them more risky for lenders. Alternatively, they pay more for credit, which limits their ability to accrue wealth at a rate faster than their wealthier counterparts.

Empirically, there is much less evidence on this. In part, most analysis in based on household and individual surveys, and the questions generally do not adequately cover these facets. For example, if a poor person has no debt since no one will lend to them, questionnaires do not generally ask about the cost of debt that a person would pay if they did indeed have such debt, when said person does not have any debt. In addition, the non-response in such surveys is likely not random, with wealthier households being less likely to participate. Finally, many people refuse to answer questions about both income and wealth, and the overall picture remains somewhat unclear.

The National Credit Regulator (NCR) maintains a database of individual indebtedness that would likely be a useful source of information. They also have several reports available on their website at www.ncr.org.za .The South African Savings Institute (SASI) in conjunction with FinMark trust commissioned some research on savings behavior for old age among poorer people in South Africa. The main finding they report is that savings rates for the express purpose of retirement are very low. The reasons for this include myopia, affordability, alternative investments such as education of children and housing, inflexible savings products and the state provided old age pension.

One of the few published academic papers on savings, insurance and debt is the review article by Ardington et al (2004). They too observe that while South Africa had a well developed financial sector, it was only households at the upper end of the spectrum that could afford to make use of these services. In particular, both cost and location act to exclude the rural poor from using the formal banking system. Obtaining a more thorough understanding of

the opportunities to borrow, save and manage risk across the income distribution would be a useful though daunting exercise as a sub-component of understanding inequality persistence in South Africa.

5.3.3   Other mechanisms

Some other mechanisms exist that may be both a cause and a consequence of inequality. One of these involves human capital in the form of health. South Africa has a well developed private healthcare system, and a less well resourced public health care system. This might arise due to high inequality as follows: The public sector struggles to keep doctors within the state service as the highly skilled doctors want all the privileges and amenities that come with the high incomes they can earn in the private sector. Poorer people cannot afford the private healthcare and medical insurance, and thus have to use the over-burdened public health system, which leads to a greater degree of health risks and problems among poorer people. This in turn affects their labour market outcomes, which in turn generates inequality. Crucial to this hypothesis is that the supply of doctors and nurses is constrained, which it certainly is for a number of reasons.

A different mechanism could be due to crime. Demombynes and Özler (2005) find evidence that crime is positively correlated with local inequality in South Africa. At the same time, Kingdon and Knight (2001) posit that one reason for the low levels of informal sector activity in South Africa may be due to the high rates of crime. If a business needs to pay some fixed cost for security, then there are likely to be certain thresholds of scale below which the cost of security is prohibitive. This would limit the poor from engaging in what might otherwise be profitable activities, thus reinforcing the income distribution.

Yet another possible mechanism may be more in line with the social stratification and access mechanisms discussed above. Indeed, Magruder (2010) examines the importance of network based intergenerational correlations in employment in South Africa. He finds that geographically present fathers may be responsible for a one third increase in their sons' employment

rates.

6 Conclusion

We conclude with a brief discussion about future research possibilities. From an empirical perspective, there are several interesting questions which further research could feasibly inform. One analysis to perform would be to replicate the research proposal of Cruces et al. Given the amount of time and effort they have invested to operationalize an empirical measure to identify inequality traps, it should be reasonably simple to replicate their study for South Africa using our abundance of micro-datasets. This would then have the very useful characteristic of being able to compare our findings with those obtained from a host of other countries where the same methodology has been used.

Many of the mechanisms that generate inequality traps involve the labour market in some form or another. In addition, all of the empirical studies on measuring inequality in South Africa agree that understanding the labour market is central to understanding inequality. This applies both in terms of wages and unemployment domestically. While exceedingly broad, any information about the puzzle of extremely high and persistent unemployment would be useful. In particular, if the costs of complying with legislation are that onerous, why do people not choose self employment of some sort? A simple survey among the unemployed might be very useful. A more specific question to ask, following on from the paper on Mexico, is whether in a time-space sense higher inequality correlates with higher levels of unionization, and whether that in turn correlates with higher steady state levels of unemployment?

Questions about the production function for education and education quality have been asked. As yet, there are no clear answers about what makes a good school 'good'. How much of quality is really perceived, and how much is genuinely productive quality? This leads to models of imperfect information and statistical discrimination on the part of employers. Moreover, what are the full rates of return to school quality, after accounting for the costs,

the probability of success, and subsequent attainment of a tertiary qualification?

One research method that we could employ would be to use institutional rules on financial aid at a place like UCT to instrument for enrollment. With enough planning and resources, we could find people from the past who did and did not receive assistance, and then estimate the returns to enrollment. This maps into the effect of financial constraints on investment in education, and how this in turn may or may not perpetuate inequality.

When we consider the role of financial markets and access to credit, it seems that a lot of work has been done by non-academic researchers. A starting point would be to update the 2004 review done by Ardington et al (2004). Moreover, data could potentially be obtained from the NCR, SARS and the state housing registration office on property prices. Each of these would yield different types of data that would be less susceptible to problems of non-response that arise in the household or individual level surveys.

An additional research question could be to use something like the NIDS data to look at how people who do and do not have bank accounts accrue assets differently, or manage to navigate adverse shocks.

If we consider health, we can use the incidence of an illness in conjunction with the location of the individual to proxy for variation in the general quality of healthcare received. Particularly in rural areas, the closest clinic may be several kilometers away and might be staffed only by nurses without any doctors. This might yield an unbiased estimate on the returns to health for different people.

Political mechanisms may contribute to inequality persistence in South Africa. Having been explored little, they offer ample research potential. On the demand side, key questions are to determine how strong demand for redistribution actually is and what the drivers behind a potentially weak demand are. Are they to be found in (self-)indoctrination of the sort found in the U.S. or do they lie in an optimism that is due to the recent political transition? It will be equally important to explore factors relating to a lack of influence of the poor on policy such as lobbying by interest groups or the ideological allegiances of the South African

poor. Finally, studies of the strength and form of clientelism in South African politics may give fruitful insights into how (local) politicians may perpetuate inequality by buying off the poor.

Although the papers on social mechanisms analyse situations that are specific to the United States, they uncover mechanisms that are also potentially relevant to South Africa given its history. Here the apartheid system institutionalised racial segregation and created an environment that generated neighbourhood effects and persistent inequality. However, in this particular case the exclusion mechanism was not the housing market, zoning rules or income barriers. Instead, it was the system of institutionalised racial segregation that kept blacks (in the general sense) from the good neighbourhoods. Since the Group Areas Act was repealed in the early 1990s, the process of segregation has become economic rather than purely racial (although race and economic status in SA remain highly correlated). Studies exploring the consequences of the repealment of the Group Areas Act on the migration patterns and educational performance of the various groups of colour defined under apartheid would be fascinating. What is the pattern by race? If the formerly disadvantaged move to better neighbourhoods (formerly white areas or simply areas that were better off than their own), how do the inhabitants of the area respond to this migration? Any such responses may occur in several dimensions, including accommodation, building of income or cultural barriers, migration and the private provision of public goods (e.g. private schools rather than public school system). What is the overall evolution of those areas where the formerly disadvantaged migrated to in terms of neighbourhood quality, educational performance, etc.

In conclusion, there remains many interesting and important questions related to the highly persistent inequality in South Africa. Of these, there are several which can be pursued successfully. This summary will act as a guide in our future research activities.

References

Acemoglu, D. & Robinson, J. (2000), Why did the west extend the franchise? Democracy, inequality, and growth in historical perspective, *Quarterly Journal of Economics*, 115, 1167-1199.

Aghion, P. & Bolton, P. (1997), A Theory of Trickle-Down Growth and Development, *Review of Economic Studies*, 64, 151-72.

Alesina, A. & Angeletos, G. M. (2005), Fairness and Redistribution, *American Economic Review*, 95, 960-980.

Alesina, A. & Paola, G. (2009), Preferences for Redistribution, National Bureau of Economic Research, *NBER Working Paper*, No. 14825.

Ardington, C., Lam, D., Leibbrandt, M. & Welch, M. (2005), The sensitivity of estimates of post-apartheid changes in South African poverty and inequality to key data imputations, *CSSR Working Paper*, No. 106.

Ardington, C., Lam, D., Leibbrandt, M. & Levinsohn, J. (2004), Savings, insurance and debt over the post-apartheid period: A review of recent research, *South African Journal of Economics*, vol. 72(3), pages 604-640.

Banerjee, A. V. & Newman, A. F. (1993), Occupational Choice and the Process of Development, *Journal of Political Economy*, 101, 274-98.

Becker, G. S. (1983), A Theory of Competition among Pressure Groups for Political Influence, *The Quarterly Journal of Economics*, 98, 371-400.

Bénabou, R. (1993), Workings of a City: Location, Education, Production, *Quarterly Journal of Economics* 108, 619-653.

Bénabou, R. (1996), Equity and Efficiency in Human Capital Investment: the Local Con-

nection, *Review of Economic Studies* 63, 237-264.

Bénabou, R. (2000), Unequal Societies: Income Distribution and the Social Contract, *American Economic Review*, 90, 96-129.

Bénabou, R. (2005), Inequality, Technology and the Social Contract, *Handbook of Economic Growth*, Elsevier.

Bénabou, R. & Tirole, J. (2006), Belief in a Just World and Redistributive Politics, *The Quarterly Journal of Economics*, 121, 699-746.

Bhorat, H., van der Westhuizen, C. & Jacobs, T. (2009), Income and Non-Income Inequality in Post-Apartheid South Africa: What are the Drivers and Possible Policy Interventions?, University of Cape Town, Development Policy Research Unit, *DPRU Working Paper*, No. 09/138.

Boadway, R. & Keen, M. (2000), Redistribution, *Handbook of Income Distribution*, Elsevier.

Bourguignon, F., Ferreira, F. & Walton, M. (2007), Equity, efficiency and inequality traps: A research agenda, *Journal of Economic Inequality*, 5, 235-256.

Branson, N., Leibbrandt, M. & Zuze, L. (2009), The demand for tertiary education in South Africa, Saldru, Cape Town.

Cervellati, M., Esteban, J. & Kranich, L. (2007), The Social Contract with Endogenous Sentiments, Unitat de Fonaments de l'Anàlisi Econòmica (UAB) and Institut d'Anàlisi Econòmica (CSIC), *Working Paper*, 702.07.

Cox, G. W. (2010), Swing Voters, Core Voters and Distributive Politics, in Shapiro, I., Stokes, S. C., Wood, E. J. & Kirshner, A. S. (eds.), *Political Representation*, Cambridge, Cambridge University Press.

Cruces, G., Bérgolo, M., Carbajal, F., Conconi, A. & Ham, A. (2010), Looking for evidence

on inequality traps in LAC. Persistent lack of opportunities and social immobility, presented at the 8th PEP general meeting, Dakar.

Daymon, C. and Gimet, C. (2009), An empirical test of the inequality test concept, *Economic Letters*, 105, 165-167.

Demombymes, G. and Özler, B. (2005), Crime and local inequality in South Africa, *Journal of Development Economics*, 76, 265-292

Dixit, A. K. & Londregan, J. (1996), The Determinants of Success of Special Interests in Redistributive Politics, *Journal of Politics*, 58, 1132-1155.

Dixit, A. K. & Londregan, J. (1998), Ideology, Tactics, And Efficiency In Redistributive Politics, *The Quarterly Journal of Economics*, 113, 497-529.

Durlauf, S. (1996), A Theory of Persistent Income Inequality, *Journal of Economic Growth* 1, 75-93.

Esteban, J. & Ray, D. (2006), Inequality, Lobbying, and Resource Allocation, *American Economic Review*, 96, 257-279.

Fernandez, R. & Rogerson, R. (1997), Keeping People Out: Income Distribution, Zoning, and the Quality of Public Education, *International Economic Review*, 38(1), 23-42.

Galor, O. & Zeira, J. (1993), Income Distribution and Macroeconomics, *Review of Economic Studies*, 60, 35-52.

Ghatak, M. & Jiang, N. H. (2002), A simple model of inequality, occupational choice, and development, *Journal of Development Economics*, 69, 205-226.

Guerrero, I., López-Calva, L. F. and Walton, M. (2006), The inequality trap and its links to low growth in Mexico, Harvard University, mimeo.

Harms, P. & Zink, S. (2003), Limits to redistribution in a democracy: a survey, *European*

Journal of Political Economy, 19, 651-668.

Hoogveen, J. G. M. & Özler, B. (2006), Not separate, not equal: Poverty and inequality in post-apartheid South Africa, *William Davidson Institute Working Paper*, No. 739.

Keswell, M. (2004), Education and racial inequality in post apartheid South Africa, *Santa Fe Institute Working Paper*, No. 2004-02-008.

Kingdon, G. and Knight, J. (2001), Why High Open Unemployment and Small Informal Sector in South Africa?, paper presented at the Labour Markets and Poverty in South Africa Conference, 2001, Johannesburg.

Kiminori, M. (2002), The Rise of Mass Consumption Societies, *Journal of Political Economy*, 110, 1035-1070.

Knoesen, S. G. (2008), Public Services Provision in South Africa: Water and Electricity Connections at the Ward Level, Midwest Political Science Association Annual Meeting. Chicago.

Lam, D. (1999), Generating extreme inequality: Schooling, earnings and intergenerational transmission of human capital in South Africa and Brazil, *Population Studies Center Research Report*, No. 99-439, University of Michigan.

Leibbrandt, M., Bhorat, H. & Woolard, I. (2001), Understanding Contemporary Household Inequality in South Africa, Chapter 2 in Bhorat, H., Leibbrandt, M., Maziya, M., van der Berg, S. & Woolard, I. (2001). *Fighting poverty: Labour markets and inequality in South Africa*. UCT Press: Cape Town.

Leibbrandt, M., Levinsohn, J. & McCrary, J. (2010), Incomes in South Africa after the fall of apartheid, *Journal of Globalization and Development*, 1(1), Article 2.

Leibbrandt, M., Woolard, I., Finn, A. & Argent, J. (2010), Trends in South African income distribution and poverty since the fall of apartheid, *OECD Social, Employment and*

Migration Working Papers, No. 101.

Lindert, P. H. (2004), Growing Public, Cambridge University Press.

Ljungqvist, L. (1993), Economic underdevelopment : The case of a missing market for human capital, *Journal of Development Economics*, 40, 219-239.

Magruder, J. R. (2010), Intergenerational networks, unemployment, and persistent inequality in South Africa, *American Economic Journal: Applied Economics*, 2(1), 62-85.

Mani, A. (2001), Income Distribution and the Demand Constraint, *Journal of Economic Growth*, 107133.

Maoz, Y. D. & Moav, O. (1999), Intergenerational Mobility and the Process of Development, *Economic Journal*, 109, 677-97.

Matsuyama, K. (2000), Endogenous Inequality, *Review of Economic Studies*, 67, 743-59.

Matsuyama, K. (2002), The Rise of Mass Consumption Societies, *Journal of Political Economy*, 110, 1035-1070.

Matsuyama, K. (2005), Emergent Class Structure, Center for Mathematical Studies in Economics and Management Science, *Northwestern University Discussion Papers*, No 1407.

Meltzer, A. H. & Richard, S. F. (1981), A Rational Theory of the Size of Government, *Journal of Political Economy*, 89, 914-27.

Milanovic, B. (2003), The median voter hypothesis, income inequality and income, *Econ-WPA*, No. 0305001.

Mookherjee, D. & Ray, D. (2003), Persistent Inequality, *Review of Economic Studies*, 70, 369-393.

Olson, M. (1965), The Logic of Collective Action, Cambridge, MA, Harvard University Press.

Owen, A. L. & Weil, D. N. (1998), Intergenerational earnings mobility, inequality and growth, *Journal of Monetary Economics*, 41, 71-104.

Pellicer, M. (2009), Inequality persistence through vertical vs. horizontal coalitions, *Journal of Development Economics*, 90, 258-266.

Peltzman, S. (1976), Toward a more general theory of regulation, *Journal of Law and Economics*, 211 240.

Piketty, T. (1997), The Dynamics of the Wealth Distribution and the Interest Rate with Credit Rationing, *Review of Economic Studies*, 64, 173-89.

Piketty, T. (2000), Theories of persistent inequality and intergenerational mobility, *Handbook of Income Distribution*, Elsevier.

Saint-Paul, G. & Verdier, T. (1996), Inequality, redistribution and growth: A challenge to the conventional political economy approach, *European Economic Review*, 40, 719-728.

Sondheimer, R. & Green, D. P. (2010), Using Experiments to Estimate the Effects of Education on Voter Turnout, *American Journal of Political Science*, 54, 174-189.

Stokes, S. C. (2005), Perverse accountability: A formal model of machine politics with evidence from Argentina, *American Political Science Review*, 315-26.

Tiongson, E. & Mello, L. D. (2003), Income Inequality and Redistributive Government Spending, International Monetary Fund, *IMF Working Papers*, No. 03/14.

van der Berg, S. (2009), Fiscal Incidence of Social Spending in South Africa, *Stellenbosch Working Paper Series*, No. 10/2009.

van der Berg, S. (2010), Current poverty and income distribution in the context of South African history, *Stellenbosch Economic Working Papers*, No. 22/10.

van der Berg, S. & Louw, M. (2004), Towards time series estimates of distribution and

poverty, *South African Journal of Economics*, 72(3), 546 - 572.
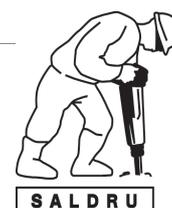
van der Berg, S., Louw, M. & Yu, D. (2008), Post-transition poverty trends based on an alternative data source, *South African Journal of Economics*, 76, 58-76.

Yu, D. (2010), Poverty and inequality trends in South Africa using different survey data, *Stellenbosch Economic Working Papers*, No 04/10.

southern africa labour and development research unit

The Southern Africa Labour and Development Research Unit (SALDRU) conducts research directed at improving the well-being of South Africa's poor. It was established in 1975. Over the next two decades the unit's research played a central role in documenting the human costs of apartheid. Key projects from this period included the Farm Labour Conference (1976), the Economics of Health Care Conference (1978), and the Second Carnegie Enquiry into Poverty and Development in South Africa (1983-86). At the urging of the African National Congress, from 1992-1994 SALDRU and the World Bank coordinated the Project for Statistics on Living Standards and Development (PSLSD). This project provide baseline data for the implementation of post-apartheid socio-economic policies through South Africa's first non-racial national sample survey.

In the post-apartheid period, SALDRU has continued to gather data and conduct research directed at informing and assessing anti-poverty policy.   In line with its historical contribution, SALDRU's researchers continue to conduct research detailing changing patterns of well-being in South Africa and assessing the impact of government policy on the poor. Current research work falls into the following research themes: post-apartheid poverty; employment and migration dynamics; family support structures in an era of rapid social change; public works and public infrastructure programmes, financial strategies of the poor; common property resources and the poor. Key survey projects include the Langeberg Integrated Family Survey (1999), the Khayelitsha/Mitchell's Plain Survey (2000), the ongoing Cape Area Panel Study (2001-) and the Financial Diaries Project.