Analysing the welfare-improving potential of land in the former homelands of South Africa

Stefania Lovo
Department of Economics, University of Sussex. UK

September 24, 2010

Abstract

This paper investigates the relationship between land and household welfare in the former homelands of South Africa. It uses the year of arrival in the current location as an instrument for land access and size. This identification strategy relies on the fact that African households were forcibly relocated to the homelands since the introduction of the Native Land Act in 1913 to the end of apartheid. Because of increasing population pressure in the homelands, later arrivals were less likely to have access to land and to larger plots of land. Results show that access to land positively affects the welfare of rural households in the former homelands. This finding is confirmed by several specifications proposed to deal with the presence of potential confounding effects. Because the homelands are relatively more disadvantaged and less fertile areas, these results are likely to provide a lower bound for the positive effects of land on household welfare.

Keywords: household welfare, land, forced removals.

JEL: O12, Q15.

1 Introduction

South Africa has a large rural population mostly residing in the former homelands. Statistical records reveal that in 1997 about 12.7 million households, that is 31% of the total population, were living in rural areas in the former homelands. Despite the large share of the rural population, 86% of arable land was controlled by large commercial farms while 50% of landed African households had access to less than 1 hectare of land. Although off-farm activities and government transfers are very important sources of income for the rural economy, land-based activities can highly contribute to the overall well-being of rural population by providing a return to family uneducated labour (Carter and May, 1999) and goods and services for home consumption. This argument gains more importance when placed in a broader economic context. According to Lipton and Lipton (1993), South Africa's large endowment of labour suggests the need for more labour-intensive agricultural production that requires a movement toward small scale labour-intensive farming. Eswaran and Kotwal (1986), in fact, show that the utility of landless households is also increased when the distribution of land is moved from a highly unequal setting, with few very large farms, to a scenario characterised by a more egalitarian distribution of smallholders.

Although the effects of land holding extend beyond those on the direct beneficiaries, the analysis proposed here is limited to the relationship between land endowments and household welfare in the rural former homelands. An asset index is used to measure household welfare using two different datasets collecting information on rural farm households. The choice of the indicator is driven mainly by the availability of the data. Nevertheless, the asset index has some advantages over other measures of welfare, which will be explained during the analysis. Moreover, the asset index, constructed using principal component analysis (Filmer and Pritchett, 2001), leads to a welfare distribution that is coherent with that obtained using alternative measures such as income and consumption per capita.

The economic theory of the farm household provides support to a positive relationship between land and household welfare. However, little empirical evidence is available mainly due to the difficulties in identifying the causal relationship between land and a measure of household welfare. This paper investigates the relationship between land endowments and household welfare exploiting historical data on migration to the former homelands. The identification strategy relies on the fact that, since the introduction of the

Native Land Act in 1913, African households have been forcibly relocated to the homelands. The year of arrival in the current location is used as instrument for land endowments since later incomers were less likely to have access to land and to larger plots of land given the increasing population pressure in these areas.

Results show the positive effect of land access on household welfare. Land size is also positively related to household welfare so that an increase of 1 hectare is expected to lift the household into a higher decile of the welfare distribution. A set of alternative specifications control for the presence of confounding effects produced by the potential correlation between the year of arrival and the location of the household, the displacement costs occurring in the early years after arrival and the quality of the land. Further checks ensure that the results are robust to the choice of the welfare indicator and of the historical sub-periods characterising the process of segregation of the African population since 1913. The impossibility of distinguishing between voluntary and forced movements, in particular within the homelands where the first are more likely to occur, however challenges the validity of the instrument. This issue is partially addressed with the support of information on the district of previous residence contained in one of the survey adopted and in the 1996 population census.

This paper proceeds as follow. Section 2 provides an overview of the existing literature investigating the relationship between land endowments and household welfare from both a theoretical and empirical perspective. This is followed by the description of the two datasets used in the analysis and a discussion of the main characteristics of the households in the sample. Section 4 introduces the historical setting underlying the identification strategy proposed in this paper with a focus on the main events and aspects characterising the massive forced removals of the African population conducted during the apartheid legacy. Section 5 outlines the empirical strategy adopted and section 6 and 7 discuss the results. Finally section 8 concludes.

2 Land and household welfare: theory and existing empirical evidence

Several authors have highlighted the importance of land in contributing to the livelihoods of the rural South African population in both financial and

social terms. Most households, for example, derive a direct use value from land-based activities from the provision of goods and services associated with livestock, foods harvested and natural resources for home consumption and for exchange with other goods and services.

The theoretical framework underlying the economic theory on land and household welfare is mainly based on the standard microeconomic theory of the farm household developed by Singh et al. (1986). The household farm is considered a unitary decision unit in which both the consumption and the production side are taken into account. The focus on the household rather than the farm unit is particularly relevant in the presence of market imperfections since consumption and production decisions are jointly determined. Eswaran and Kotwal (1986) and Finan et al. (2005), for example, use a farm household model with imperfect credit and labour market conditions and where access to credit increases with land size. This is based on the argument that larger farms have better access to credit through the collateral use of land. Eswaran and Kotwal (1986), show how household labour allocation decisions are determined by land endowments and that a transfer of working capital, including land, from larger to smaller farm households can be welfare and output improving. Finan et al. (2005) show how household income is positively affected by land endowments through a direct effect (the income generated by the increased production) and an indirect effect when labour and credit markets are imperfect. The magnitude of the overall effect varies across households, in particular, depending on whether the increased demand for inputs is matched by an increased availability of credit due to use of additional land as collateral. In the same vein, Burgess (2001) uses a theoretical household model where land generates a twofold effect on household welfare. Considering imperfections in land and food markets, the author shows that land has the potential to increase household consumption through an income effect, due to increased production, and by providing a cheaper source of food to the household.

Although the economic theory of the farm household gives support to the positive relationship between land and household welfare, with heterogeneous features across households, there is little empirical evidence mainly due to the difficulties in identifying the causal relationship between land and a measure of household welfare. Finan et al. (2005) analyse the impact of land on household welfare, measured using an asset index, using data on Mexican rural households for the period 1997-1998. They propose a linear and a non-parametric specification to capture the non-linearities in the rela-

tionship between land and household welfare. Although the study provides an extensive and rigorous analysis of the heterogeneous correlation between land and welfare across households, little attention is paid to the identification of the causal relation between the two. The authors find that land has a very high marginal welfare value for small farms (less than 1 hectare), and that this effect decreases with land size. Burgess (2001), using data on Chinese households, investigates the relationship between land size and household welfare measured by food consumption expenditure and calories intake. Considering a standard food demand equation, once per capita income has been controlled for, the additional effect of land per capita is to be attributed to its role in providing cheaper food to the household. The effect of land is identified by the fact that land in China is distributed on the basis of household nutritional needs, and therefore of household composition, that is beyond households' discretion given the strict family planning policies.

The majority of papers look at the impact of land transfers obtained through the implementation of land reforms. Besley and Burgess (2000), for example, using a panel dataset on sixteen Indian states for the period 1958-92 found that post-independence land reforms positively contributed to poverty reduction. The potential endogeneity of the land reform variable is addressed by using the composition of past political legislatures as instruments for land reform transfers. Other papers are particularly relevant for the present study given their focus on South Africa land reforms implemented since 1997. Keswell et al. (2010) exploit the quasi-experimental setting of the Land Redistribution and Agricultural Development (LRAD) program, introduced in 2001, and find a positive effect on household consumption for the beneficiaries. The impact is identified by comparing households still in the process of being granted the land transfer and households that have already received it. A previous paper by Valente (2009) looks at the impact of LRAD program on household food security. The results show that the land reform has not been successful in reducing the food insecurity of the beneficiaries. This is mainly attributed to the high displacement costs, since in most cases the assigned land is located far from household's current location, and the lack of organisation. The author uses alternative techniques to deal with observed and unobserved variable biases, although no suitable instruments were available to fully address the endogeneity of the land reform variable.

The existing empirical literature confirms the difficulties in identifying the causal relationship between land and household welfare given the non-random allocation of land and the lack of suitable instruments. In this paper,

I will attempt to address this empirical issue by exploiting historical data on household migration to the homelands.

3 Data

The data used in this analysis are drawn from two different datasets: the KwaZulu-Natal Development Indicators Household survey (KZN-DIHS) of 1996 and the Rural Survey of 1997. These datasets are the only available datasets that provide information on both land and migration history. I opted for the use of two surveys mainly because neither of them provides exhaustive information for the purpose of this analysis. The Rural Survey 1997 provides data on the amount of land available to the household, the initial focus of this analysis, and detailed information on farming activities. However, it does not provide information on location (distance to the nearest town) and on the previous district of residence, although it does report the year of arrival. This latter information, in particular, is useful to narrow the focus of the analysis to provide further support for the use of the estimation strategy adopted in this study, as will be explained in the following sections. The KZN-DIHS 1996, instead, provides more detailed information on migration to the homeland but is confined to a much smaller sample and provides information only on whether the household has access to land with relatively less information on farming activities. The use of both surveys, therefore, allows me to conduct some exercises to support the instrumental variable strategy adopted in this study. However, because the information on land provided by the two surveys is now of two types, a binary variable indicating access to land and a continuous variable indicating the amount of land available, the analysis is conducted separately for each dataset and is described is section 6.1 and 6.2.

The KZN-DIHS has been conducted by the KwaZulu-Natal Provincial Government and Human Sciences Research Council (HSRC). The complete survey covers 6500 households across the province of KwaZulu-Natal, which incorporates the former homeland of KwaZulu. The sample size has been reduced to consider only the households living in rural areas. This cross section survey has been used mainly because it provides information on both the year of arrival and the previous district of residence, which makes it possible to establish whether a household has moved from a non-homeland area. However, it does not provide information on the amount of land owned by

the household and only reports whether the household has access to land for agricultural purposes. The survey provides detailed information on household living conditions and asset ownership that are useful for the construction of a welfare index. It also provides information on household consumption that will be used in one of the empirical specifications proposed below.

The analysis of the impact of land size on household welfare uses the Rural Survey 1997 conducted by Statistics South Africa, which collected information on 6,000 rural households located in the 10 former homeland territories. This cross-section survey provides information on the hectares available to each household for farming purposes, although less detailed information is available as far as asset holding, income or consumption are concerned. Another drawback of this survey is the lack of information on the previous district of residence, as previously mentioned, so that it is not possible to distinguish between movements to and within the homelands.

Table 1: Descriptive statistics of the main variables of interest

	(1) KZN-DIHS 1996 Mean (sd)	(2) Rural Survey 1997 Mean (sd)
Land (dummy)	0.38	0.65
	(0.49)	(0.47)
Hectares of land (landed households)		1.41
		(3.57)
Education household head (dummy)	4.07	4.66
	(3.37)	(4.63)
Age of household head	49.25	56.25
	(14.23)	(16.22)
Gender of household head (dummy)	0.75	0.51
	(0.43)	(0.50)
Pension eligible members (dummy)	0.28	0.42
	(0.45)	(0.49)
Children	1.61	2.97
	(1.49)	(2.01)
Number of skilled members	0.80	1.08
	(1.15)	(1.22)
Number of unskilled members	2.20	2.00
	(1.45)	(1.31)
Average eduction of adult members	5.21	6.58
	(3.16)	(3.49)

Author's calculation using the KZN-DIHS and the Rural Survey.

The summary statistics of the variables used in the empirical analysis and reported in table 1 offer a general picture of the main characteristics of the households that are the focus of this study. According to the KZN-DIHS, 38% of the households living in rural areas in the KwaZulu Natal province have access to land. Among rural households residing in the former homelands, 65% of those interviewed by the Rural Survey have access to a plot of land. Plots are in general small with an average size of 1.41 hectares, so that as a consequence, only 10% of them produce farm products for sale while the majority work the land to provide food for home consumption. The average household size is between 4 and 5 members. Adult members have on average 5 to 6 years of education, less than the 9 years of compulsory education introduced in 1996.

4 Historical background

Segregation in South Africa started to take shape with the introduction of the Natives Land Act in 1913 stating that black Africans were no longer to be able to own or rent land outside designated reserves. During the apartheid era, which officially started in 1948, the reserves were converted to bantustans or homelands and later some of them into 'independent' states within South Africa. The population was classified into four racial groups ('black, 'white, 'coloured, and 'Indian). From 1958, the black population was deprived of its citizenship, legally becoming citizens of one of ten tribally based self-governing homelands: Lebowa, QwaQwa, Bophuthatswana, KwaZulu, KaNgwane, Transkei, Ciskei, Gazankulu, Venda and KwaNdebele. African people were only temporary resident in the remaining territories for as long as they offered their labour there.

Residential areas were segregated, often by means of forced removals. According to Desmond (1971), the governments object was to return 5% of the African population from the white areas to the homelands every year. Several laws regulated the movements of the African population. The Pass law, introduced in 1923, stipulated that the black population should carry pass books when outside the designated homelands. Several influx controls were introduced to limit the number of African people allowed to live and work in white areas (Platzky and Walker, 1985).

People were relocated from 'white farms, from 'black spots' (area of black settlement surrounded mainly by 'white farmers), from small town locations

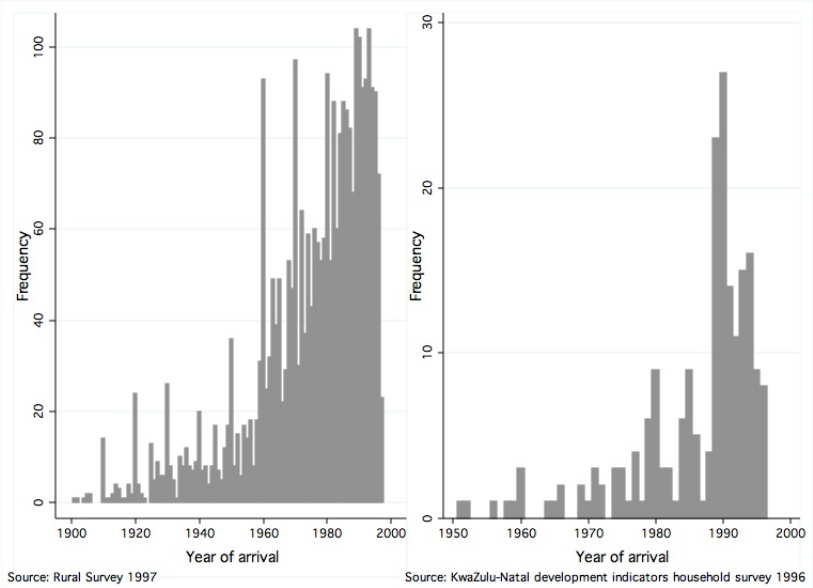
and from metropolitan areas. Removals were initially conducted by direct intervention of government authorities also through arbitrary searches and checks. Later, after 1980, the public emphasis was on people moving 'voluntarily. Removals were, however, the results of indirect coercion by the authorities and the security police through intimidations and threats of arrest and detention (Platzky and Walker (1985), pp 152-76). In many townships and rural areas, for example, new construction was frozen; hospitals, schools and other public facilities for the black population were relocated to the homelands. This was a deliberate tactic to enforce voluntary removal to the homelands (Murray, 1987).

There are no official records of removals and often statistical data were deliberately hidden. However, according to Platzky and Walker (1985), the process of forced removals affected some 3.5 million people in the period 1960-1982 excluding those households forcibly removed within the homelands due to the implementation of the betterment plans described below. Desmond (1971) provides the first attempt to document forced relocations, his narrative description of removals is the results of months spent travelling throughout the country. Simkins (1983) provides some quantitative estimates of population changes and movements for the year 1950, 1960, 1970 and 1980 and estimates a net inflow to the homelands of about a million people in the decade 1960-1970 that had its counterpart outflow mostly in the rural areas outside the homelands.

Figure 1 plots the frequency of arrivals in the current location for households living in the homelands using information from the two household surveys. While it shows that the date of arrival is not always accurately reported, given the high frequency of rounded decades (this issues will be addressed later in the empirical analysis), it shows an acceleration of movements in the early 1990s. This is in line with the fact that evictions accelerated in this period partly in response to commercial farmers' concerns about legislation intended to improve the security and working conditions of their workers (Lyne and Darroch, 2004).

The process of forced relocations also continued within the homeland territories. According to Freund (1984) the initially scattered structure of the homelands, appearing as demarcated islands within South Africa, required a consolidation program that produced another massive wave of removals. Even after this process, because the population within the homelands was of heterogenous ethnic background, an additional reshuffling of people was conducted. Moreover, a series of 'betterment plans' were implemented from

Figure 1: Distribution of arrivals in the homelands



the 1930s onwards to control land usage, which are considered to have produced the numerically largest and most widespread form of resettlement in South Africa. de Wet (1994) argues that if within-homeland relocations are considered, at least seven million African have been resettled for political purposes since 1913. Under this program, designated areas were divided into distinct land use zones: residential, arable and grazing areas. Land regarded as unsuitable for cultivation was removed from use, so that in some areas people were left with less arable land than they had before or they lost their arable land altogether (de Wet, 1987). Finally, households were also removed for strategic and infrastructural reasons, for example to make space for dam projects (Woodstock and Upper Tugela) or for the clearance of South African borders (Freund, 1984). Finally, it is worth noting that forced removals "did not follow a pre-determined and predictable blueprint. Potential victims could not entirely count on the next move of the state" (Freund, 1984) since government removals plans often appeared in contradictory forms in different official publications.

4.1 Forced removals and access to land

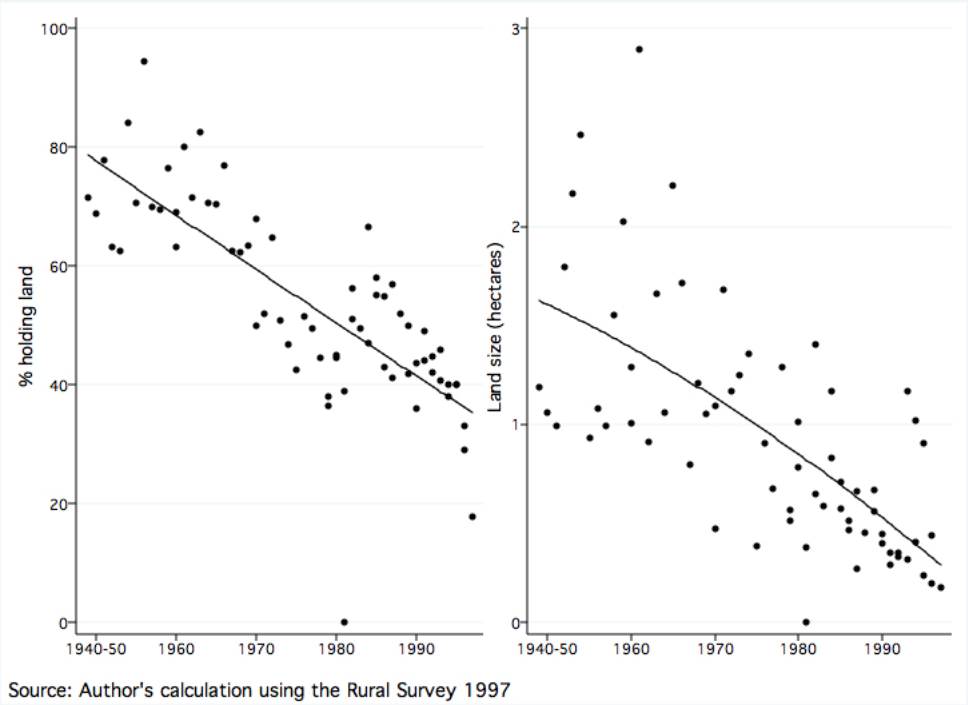
The relationship between removals and access to land is to be found in the increasing population density in the former homelands. The total population density for South Africa almost doubled between 1970 and 1995, from almost 19 people per square km in 1970 to 34 people per square km. The situation was more dramatic in the homeland areas that constitute less than 14% of African territory and hosts a large share of South African population. According to Simkins (1983), while 39% (of a total 'black' population of 11 million) were living in the homelands in 1950, 53% (of 21 million) were in the homelands in 1980. The forced removals and settlement planning were major contributors to the overcrowding in the homelands. In the Qwaqwa homeland in 1983, for example, after a period of massive relocation of people, its population density was estimated to be over 1,000 people per square km, from a population of 24,000 in 1970 to a population of 400,000 in 1983 (de Wet, 1994).

The increased population density in the homeland areas inevitably led to increasing pressure on the available land for farming and residential purposes so that those arriving later in the homelands were less likely to have access to land and particularly to larger plots of land. Using data from the Rural Survey 1997 it is possible to see these patterns in land endowments. Figure 2 shows the negative relationship between both land access and size and the date of arrival in the current location. This negative correlation is at the basis of the identification strategy adopted in this study.

The two surveys were conducted in 1996 and 1997, two and three years, respectively, after the end of the apartheid. Although land distribution has been a major concern since 1995, the first period was mainly characterised by policymaking, consultation and the building of institutions for the delivery of a land reform. Government strategies for reconstruction and development became part of South Africa's Constitution later in 1996 and the final policy framework, the White Paper on South African Land Policy, was implemented in 1997. The available data on land from the two surveys used in this study, are therefore most likely to be unaffected by post-apartheid land reforms. This constitutes an advantage for this analysis since before the implementation of the land reforms, land endowments can be better predicted using historical information on migration to the homelands.

In general, movements to the homelands after 1913 can be attributed to forced removals through coercive actions, intimidation and pressure by

Figure 2: Percentage of household holding land and land size by year of arrival



the public authorities and security police. Case studies discussed in Platzky and Walker (1985), the narrative evidence reported in Desmond (1971) and other anecdotal evidence, in fact, suggest that no households would voluntarily move to the overcrowded and unpleasant homelands. The homeland of residence is also determined by the government according to the language spoken or the ethnic group to which the people apparently belong (Platzky and Walker, 1985), and it is, therefore, excluded from household decision-making. An important distinction need to be made between the timing of relocations and the fact of being removed. The empirical analysis proposed in this chapter considers only those households that report to have moved to the current location during the period 1913-94, therefore, although the results may not be generalised to the entire population of the homelands they are not driven by systematic differences between original inhabitants and new incomers. As far as the timing of relocations is concerned, a specific time pattern cannot be identified since relocations from white rural areas

overlapped with removals from urban areas, black spots, sites allocated to strategic infrastructures and for 'betterment planning'. Therefore, the year of arrival in the homelands cannot be associated to specific causes or conditions. Moreover, because unobservable characteristics were also likely to be unknown to the authorities that enforced the relocations they are likely to be uncorrelated with the timing of arrivals. These circumstances provide a useful setting to analyse the relationship between land and household welfare exploiting the exogeneity of the year of arrival in the homeland to households' welfare-generating ability and its correlation with land access and size.

5 Measuring household welfare using principal component analysis

Household welfare is measured using an asset index. This approach is used mainly to construct a similar measure of welfare across the two surveys, although employing different type of assets, since information on consumption or income is not available in the Rural Survey 1997. Although the choice has been mainly driven by the availability of data, this approach has some advantages. An asset index captures aspects of household welfare that are usually neglected using monetary measures, for example access to basic services such as water and electricity. Moreover, because ownership of assets is easily verified it is expected to be more accurate than consumption expenditure data, for example, which are usually recorded using retrospective recall of information. Given the data available a possible alternative approach would have been the use of the number of assets owned by the households. However this approach give equal weights to all assets and does not taken into account differences in quality. Alternatively, the asset position of the household could also be measured using the value of the assets owned, however asset price are not available in the two surveys used in this study.

The asset index has been constructed using principal component analysis. This approach has been evaluated by Filmer and Pritchett (2001) who demonstrate its suitability for measuring household welfare. Because ownership of different assets is highly correlated across households it is advantageous to collapse information on specific asset ownership into a single new variable (McKenzie, 2005). This artificial variable, W_1, is obtained as the weighted sum of a set of correlated variables, in this specific case variables

indicating asset ownership. Given the vector of asset indicators $(x_1, ...x_N)$ where each vector x_n contains observations on each of the N assets for the H household in the sample, the asset index is represented by the following linear combination:

$$W_1 = f_1\Big(\frac{x_1 - \overline{x}_1}{s_1}\Big) + ... + f_N\Big(\frac{x_n - \overline{x}_N}{s_N}\Big), \quad n = 1...N, \tag{1}$$

where, $\overline{x}_n$ and s_n are the mean and standard deviation of each asset over all households, therefore, the variables are standardised to have zero mean and unit variance. Weights, f_n, are chosen so that this linear combination has the greatest sample variation. In doing so it maximises the heterogeneity across households so that assets which all or none of the households hold receive small weights, since they do not explain the variation in welfare across households.

Table 2: Scoring factors and summary statistics

	(1) KZN-DIHS 1996			(2) Rural Survey 1997		
	Score f	Mean	Sd	Score f	Mean	Sd
Electricity (dummy)	0.27	0.37	0.48	0.22	0.26	0.44
Near water (dummy)	0.14	0.67	0.47	0.04	0.36	0.48
Flush toilet (dummy)	0.34	0.16	0.37	0.03	0.01	0.08
Pit latrine (dummy)	-0.27	0.75	0.43	0.46	0.71	0.45
Other toilet (dummy)	-0.02	0.09	0.29	-0.47	0.28	0.45
Brick structure (dummy)	0.31	0.19	0.39	0.45	0.47	0.50
Traditional house (dummy)	-0.28	0.66	0.47	-0.46	0.50	0.50
Rooms per person	0.14	0.63	0.38	0.22	0.93	0.71
Number of rooms	0.02	2.45	1.23	0.24	4.75	2.52
Own washing machine	0.29	0.23	0.42			
Own washing machine	0.33	0.05	0.23			
Own vacuum cleaner	0.33	0.05	0.23			
Own microwave	0.35	0.07	0.26			
Own car	0.31	0.13	0.33			

The percentage of the covariance explained by the first principal component is 33% in (1) and 29% in (2).

In this analysis, the vector of asset indicators contains dummy variables for the ownership of specific assets (fridge, washing-machine, vacuum cleaner, microwave and car), not available in the Rural Survey 1997, characteristics of the house (brick structure, traditional, type of toilet) and access to utilities

(electricity and water), and some numerical variables such as the number of rooms in the house (table 2).

The first principal component explains 33% and 29% of the total variance in the data for the Kwazulu-Natal and Rural Survey respectively. For dummy variables the scores reported in table 2 can be easily interpreted. A movement from 0 to 1 in one of the asset indicator changes the index by the score divided by the standard deviation. A positive score indicates that the ownership of the asset leads to a higher welfare index. For example, a household owning a fridge has an asset index that is by 0.69 higher than that of a household without a fridge. This is in line with low quality assets being attributed a negative score, as in the case of the traditional-type houses and toilets of different types not connected to the sewer system.

Table 3: Descriptive statistics of the asset index by food consumption and income per capita

KZN-DIHS 1996		Rural Survey 1997	
Consumption pc Quartile	Welfare index Mean (sd)	Income pc Quartile	Welfare index Mean (sd)
1	-0.839	1	-0.393
	(1.092)		(1.697)
2	-0.719	2	-0.030
	(1.176)		(1.689)
3	-0.241	3	0.011
	(1.645)		(1.670)
4	2.475	4	0.451
	(3.492)		(1.693)

Regarding the Rural Survey 1997, although the approach uses a reasonable range of assets, the absence of information on household non-agricultural assets such as television, car etc. could lead to an incomplete representation of household living standards. Nevertheless, the asset index constructed using the Rural Survey seems to perform well when compared to an income based measure of welfare. This is reported in the second column of table 3 that shows how higher values of the asset index are associated with higher income per capita. Because information on income is provided only by categories, a better check would be against household expenditure, as it is done for the KZN-DIHS. Nevertheless, the comparison still adds some confidence in the use of this asset index as a measure of household welfare.

6 Empirical specification

This section outlines the empirical procedure for the estimation of the relationship between land endowments and household welfare. The base empirical specification is the following:

$$w_i = \alpha + \beta A_i + \theta X_i + \epsilon_i, \tag{2}$$

where w_i represents the asset index estimated using principal component analysis and X_i is a set of household and district level characteristics that are expected to affect household welfare. These variables include the characteristics of the household head: gender, age and education. The latter two variables, together with variables indicating the highest level of education in the household and the number of skilled members, are expected to capture the human capital contribution to household welfare. Because the pension transfer received from the Old Age Pension Program (OAP) is quite generous for African household and could bias the results if omitted, I also control for the presence of pension eligible members to avoid the potential endogeneity of actual pension take-up. This is done accordingly to the age eligibility criteria of the pension program by including a dummy variable taking the value of one when there is a woman over age 60 and a man over age 65 in the household. Additional controls include the number of children in different age categories, the number of unskilled members and magisterial district level characteristics such as population density, the employment rate and the share of household with access to telephone to measure the level of development of local infrastructures. Further variables are added to address specific empirical issues and will be discussed in the next section. The variable A_i captures land endowments and can be either a dummy variable indicating whether the household has access to land (obtained from the KZN-DHIS 1996) or a continuous variable representing the amount of land available to the household (using the Rural Survey 1997). These two alternative specifications are discussed below.

6.1 Dealing with an endogenous dummy variable

When analysing the impact of land access, the variable A_i indicates a dummy variable that takes the value 1 when the household has access to land and is obtained from the KZN-DHIS 1996. Given the binary nature of this variable, the average effect of land on household welfare, conditional on other

covariates, can be written as follows:

$$E[w_{1i}|X_i, A_i = 1] - E[w_{0i}|X_i, A_i = 1] = E[w_{1i} - w_{0i}|X_i, A_i = 1],$$

where w_{1i} denotes the welfare of the household if it has access to land and w_{0i} represents household welfare otherwise. Because the second term on the left of this equation is not observed, the above effect, usually defined as treatment effect on the treated, cannot be estimated. Thus a comparison of outcomes between treated and untreated is necessary and is reported below:

$$\begin{aligned} E[w_{1i}|X_i, A_i = 1] - E[w_{0i}|X_i, A_i = 0] &= E[w_{1i} - w_{0i}|X_i, A_i = 1] \\ &+ E[w_{0i}|X_i, A_i = 1] - E[w_{0i}|X_i, A_i = 0]. \end{aligned}$$

This effect involves a bias term $E[w_{0i}|X_i, A_i = 1] - E[w_{0i}|X_i, A_i = 0]$ that differs from zero when certain type of households are more likely to have access to land than others. This is the case, for example, if households that have experienced relatively unfavourable circumstances in the labour market are more likely to access land. The bias disappears when access to land is independent of household's ability to generate welfare.

One possible option to correct this bias is the use of an instrument that is correlated with the endogenous dummy variable and independent of household welfare potential, conditionally on the other included covariates. The instrument used in this analysis is the year of arrival in the homelands. As reported in section 4.1, removals conducted by the government were mostly unpredictable and produced a massive movement of people to the homelands. Given the increasing population density in the homelands following the continuous inflows of relocated households, later incomers were less likely to be given access to land (figure 2). This argument underlies the causal relationship between access to land and year of arrival.

The independence assumption requires two conditions: that the instrument is exogenous and that it does not affect welfare other than through its effect on land. First the year of arrival in the homelands can be reasonably considered uncorrelated with unobservable household characteristics, since removals where enforced by official authorities with the aim of segregating the entire black population. Moreover, because household unobservable characteristics were also likely to be unknown to the authorities that enforced the relocations they are likely to be uncorrelated with the pace and timing of the arrivals. In addition, because forced removals were mostly unpredictable no pre-moving arrangements could be undertaken. The date

Table 4: Descriptive statistics of household age structure by decade of arrival

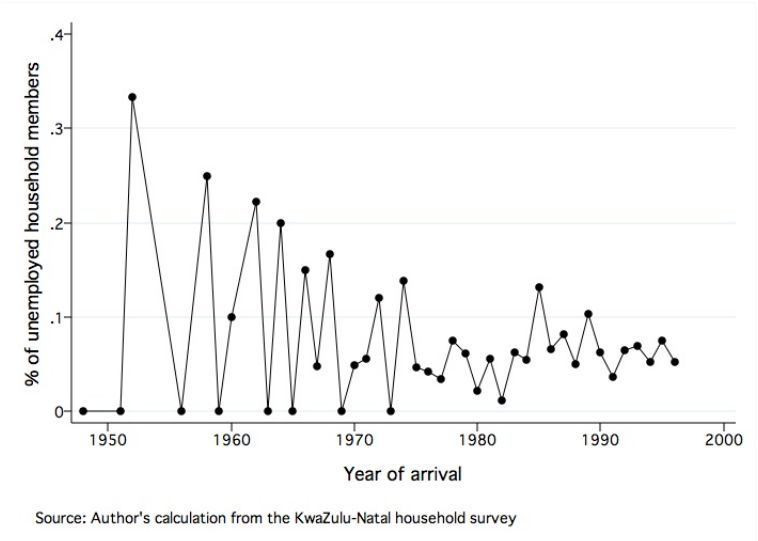
	KZN-DIHS 1996		Rural Survey 1997	
Decade	Household head	Oldest member	Household head	Oldest member
1910			62.333	67.364
			(17.947)	(13.313)
1920			63.543	65.371
			(12.312)	(10.834)
1930			60.734	64.298
			(11.673)	(10.569)
1940			58.105	61.581
			(13.460)	(12.840)
1950	68.571	68.571	59.684	62.538
	(11.013)	(11.013)	(16.102)	(15.584)
1960	57.120	60.080	58.263	61.053
	(11.805)	(13.982)	(15.283)	(14.632)
1970	51.831	52.442	55.989	58.160
	(12.557)	(13.689)	(14.309)	(14.636)
1980	46.373	46.906	51.443	53.770
	(12.257)	(12.856)	(15.342)	(15.708)
1990	41.473	42.068	45.736	47.539
	(13.313)	(14.189)	(15.523)	(16.127)
Total	44.657	45.296	53.454	55.832
	(13.629)	(14.477)	(15.892)	(16.077)

Standard errors are reported in parenthesis.

of arrival can, therefore, be considered independent of households' welfare potential. Second, the year of arrival seems not to affect household welfare through other channels rather than access to land, once additional control variables, described below, are included. Table 4 reveals that households that have moved more recently are in general younger than those that arrived earlier. Because the age structure of the household could itself affect welfare and could be captured by the instrument, regressions include controls for the age of the household head and a polynomial of the age of the oldest member of the household. In addition, the time of arrival could also have affected the location in which the household resides, which could itself have an influence on household welfare. This mechanism can be ruled out by controlling for household road distance to the nearest town as reported in column 6 of table 5. An additional issue arises if, for example, later incomers had access to fewer job and business opportunities given the increasing

population pressure in the homelands. Because this is likely to affect the probability of finding a job, it could lead to a potential correlation between the year of arrival and households' unobservable ability to generate welfare. Unfortunately it is not possible to control for household-specific employment opportunities. However, when plotting the current average share of unemployed members per household by year of arrival, using the KZN-DIHS data (figure 3), there does not seem to be a correlation between the two, suggesting that later arrivals are not worse off in terms of job opportunities[1]. This can also be explained by the fact that most of the people arriving in the homelands became cross-border commuters, living in the homelands and commuting daily to work in "white" areas (Murray, 1987). Additional controls for local population density and employment rate should also capture the availability of job opportunities in the district of residence. Additional controls are introduced to limit the presence of confounding effects and are discussed in the next sections.

Figure 3: Average share of unemployed people in the households by year of arrival (KwaZulu-Natal province, 1996)



Given the binary nature of the land variable, the first stage regression could be estimated using a nonlinear method such as logit or probit. Angrist (2001) argues that the second-stage estimates will be inconsistent if the

[1]The larger variance for period 1950-70 is due to the lower number of observations.

the first stage model is incorrectly specified while a standard two-stage least squares (2SLS) procedure can avoid this inconvenience. The model is therefore estimated using 2SLS. It is worth noting that with a dummy endogenous variable, instrumental variables procedures estimate causal effects for those households whose behaviour is affected by the instrument. This is usually known as local average treatment effect (LATE) (Angrist, 2001). That is, the effect is estimated for those households that obtained land because they arrived early but would not have received it if they arrived later, and for those households that did not receive land because they arrived late in the homelands, but would have obtained it if they had arrived earlier (these two groups are known in the literature as compliers). This means that the results are not informative about the effect on those households that would have never had access to land, or those that would have accessed the land independently of their year of arrival. Unfortunately, in practice, it is not possible to establish the size of this subpopulation of compliers.

6.2 Dealing with a positive continuous endogenous variable

When analysing the impact of land size on household welfare, A_i is a positive continuous variable indicating the amount of land available to the household that is obtained from the Rural Survey 1997. The majority of the households surveyed received the land from the local or tribal authority and about 82% do not possess title deeds. The absence of a land market could lead to the conclusion that land should be considered as an exogenous variable since households cannot easily adjust the amount according to their needs. However, the presence of unobservable household characteristics that could be potentially correlated with both land endowments and household welfare challenges the exogeneity of the land variable. Unobserved land quality, social status, habits and attitude toward agriculture, for example, are likely to be correlated with both household welfare and access to land. To address this issue I use, also in this specification, the date of arrival in the current location as an instrument for land endowments. Similarly to the argument used above, given the increasing population pressure in the homelands, later incomers were more likely to receive smaller plots of land. The correlation between land size and year of arrival can be observed in the right panel of figure 2 where average land size by year of arrival is plotted using data from

the Rural Survey 1997. In line with the previous discussion on the independence of the instrument, the year of arrival can be reasonably considered uncorrelated with unobservable household characteristics, since the timing of forced removals cannot be related to households' habits, attitude toward agriculture or welfare-generating ability. By controlling for the age of the household head and of the oldest members in the household, as previously suggested, it is possible to isolate the effect of the instrument on land from that on the age structure of the household. The year of arrival, however, could also be negatively correlated with the quality of the land plot since better quality land could have been assigned to earlier incomers. In this case the instrument would capture both land size and quality inducing a bias in the estimates. Unfortunately, information on specific plot quality is not available, but this issue is partially addressed by including a proxy for land quality at district level, constructed using average maize production per hectare. This specification is reported in table 6 column 4.

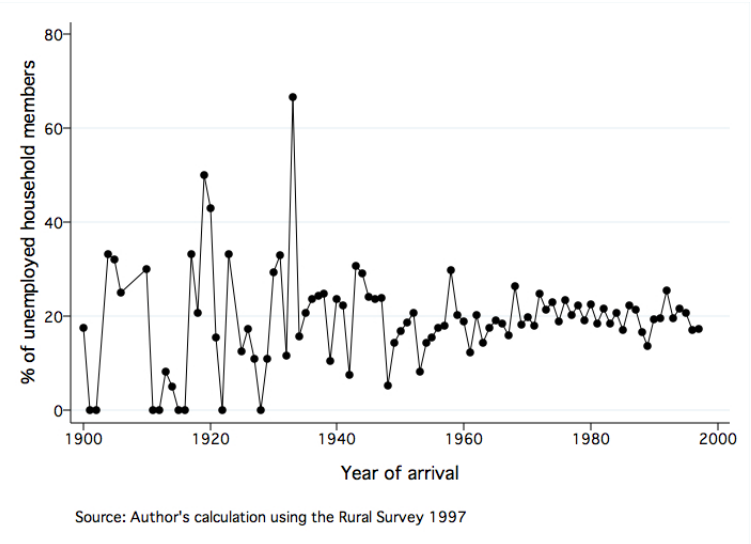
An additional channel through which the year of arrival could affect household welfare is by capturing displacement costs. Specifically, households, in the first years after arrival, may incur high transaction costs, such as search costs and resettlement costs, in the new living area that could affect their ability to generate welfare. This issue is partially ruled out, since I consider only households that arrived before the end of the apartheid, that is only households that have resided in the current location for at least 3 years. This applies also to the previous specification that deals with access to land. Further support is provided by considering only households that had been living in the current location since 1990. This specification is reported in table 6 column 6. Finally, similarly to what has been mentioned in the previous section, when using data from the Rural Survey 1997, the year of arrival does not appear to be correlated with household employment conditions (figure 4), suggesting that the instrument does not affect household welfare through its correlation with the availability of local business and job opportunities. Additional robustness checks are conducted and discussed in the next sections.

The model is estimated using 2SLS and the reduced form is the following:

$$w_i = \alpha + \beta \hat{A}_i + \theta X_i + \epsilon_i,$$

where $\hat{A}_i$ is obtained from a first-stage regression where the year of arrival in the current location is used as an instrument. Although the censored

Figure 4: Average share of unemployed people in the households by year of arrival



nature of the land size variable may suggest the use of a nonlinear first-stage regression, a conventional two-stage least squares model is consistent independently of the non-linearity of the first stage (Angrist, 2001) and is therefore preferred because it gives consistent estimates under a broader set of assumptions.

7 Estimation results: access to land and household welfare in the KwaZulu-Natal province

This section reports and discusses the effects of having access to land on household welfare in the KwaZulu-Natal province. The results have been obtained using an initial subsample of 4,368 African rural households provided by the KZN-DIHS 1996 and are reported in table 5. The subsample is further reduced once the instrument is used in the 2SLS procedure. The first column reports the results of the ordinary least squares estimation of equation 2 and shows a positive correlation between land holding and the welfare level of the household.

As discussed above, a potential endogeneity bias could be driving these

results. The other columns, therefore, report instrumental variable (IV) estimates. The instrument for the land variable is the year of arrival in the current location in all the specifications. Households that moved after 1994, i.e. after the end of the apartheid, are excluded from the analysis. The households considered are, therefore, those that moved between 1948 and 1993 since no households reports a year of arrival earlier than 1948. Unfortunately, when using this instrument the sample size is reduced noticeably to about 700 households. The first-stage regressions are strong with an F statistic above 10 (reported at the bottom of the table). In all IV specifications the first stage estimates show that the year of arrival is negatively associated with access to land[2]. The first-stage regression associated with column 2 is reported in table 9 in the appendix. Column 3 includes additional controls for the age of the oldest members of the household. This is done to ensure that the instrument is not capturing the effect of differences in the age structure of the household and are included also in the subsequent specifications. All regressions include district council dummies to control for differences in environmental and local conditions.

[2]Given the availability of only one instrument, it is not possible to test for overidentifying restrictions. However, because any function of the instrument can potentially be a suitable instrument, using both the year of arrival and its square as instruments the model is overidentified. The overidentification test statistic reveals that the null hypothesis of joint validity cannot be rejected and add increases the confidence in the instrument. However it is worth noting that the overidentification test relies on the assumption that at least one instrument is valid. Therefore in this case, if this assumption does not hold for one of the instruments it necessarily does not hold for the other. This reduces the power of the test. Because first-stage regressions are better fitted with the year of arrival only, the results reported here consider only one instrument.

Table 5: OLS and 2sls regressions of the effect of access to land on household welfare

	OLS	IV						
	(1)	(2)	(3)	(4)[a]	(5)[b]	(6)[c]	(7)[d]	(8)[e]
Land (dummy)	0.226***	2.554**	2.542*	3.368**	3.806**	3.205*	122.431**	4.273**
	(0.066)	(1.298)	(1.304)	(1.681)	(1.823)	(1.700)	(51.736)	(1.669)
Education household head	0.187***	0.087***	0.093***	0.112**	0.107	0.127**	9.318***	0.038***
	(0.013)	(0.034)	(0.034)	(0.044)	(0.068)	(0.064)	(2.197)	(0.013)
Age of household head	0.025*	0.003	0.074	0.091	-0.066	-0.248	-6.496	0.020
	(0.013)	(0.041)	(0.149)	(0.192)	(0.288)	(0.259)	(7.363)	(0.029)
Gender of household head (dummy)	0.390***	0.387**	0.412**	0.034	-0.360	-0.160	-7.369	-0.011
	(0.061)	(0.190)	(0.190)	(0.323)	(0.682)	(0.636)	(15.851)	(0.147)
Pension eligible members (dummy)	-0.153	-0.148	0.292	0.324	2.078*	1.507	56.745*	-0.167
	(0.099)	(0.324)	(0.432)	(0.544)	(1.106)	(1.022)	(34.392)	(0.146)
Children age 1-5	-0.194***	-0.098	-0.092	-0.197	-0.230	-0.230	-20.046**	0.004
	(0.044)	(0.173)	(0.172)	(0.217)	(0.279)	(0.250)	(9.735)	(0.046)
Children age 6-17	-0.216***	-0.333***	-0.345***	-0.429***	-0.426**	-0.407**	-17.029***	-0.003
	(0.022)	(0.085)	(0.086)	(0.134)	(0.172)	(0.162)	(5.516)	(0.029)
Number of skilled members	0.009	-0.001	0.005	-0.019	0.113	0.231	-11.024	0.066
	(0.040)	(0.141)	(0.141)	(0.189)	(0.250)	(0.235)	(7.419)	(0.049)
Number of unskilled members	-0.354***	-0.526***	-0.505***	-0.512***	-0.810***	-0.746***	-27.983***	-0.114**
	(0.023)	(0.117)	(0.118)	(0.145)	(0.239)	(0.213)	(8.069)	(0.046)
Highest level of education	0.495***	0.450***	0.451***	0.438**	0.303	0.279	3.327	0.072***
	(0.055)	(0.142)	(0.143)	(0.173)	(0.262)	(0.245)	(10.236)	(0.022)
Employment rate (district level)	1.161***	1.851***	1.841***	1.103	-0.699	-1.718	-0.968	0.958***
	(0.186)	(0.580)	(0.582)	(1.076)	(1.642)	(1.549)	(61.172)	(0.208)
Population density (district level)	-0.000	0.003***	0.003***	0.004***	0.004***	0.004***	0.111**	-0.000
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.049)	(0.000)
Road distance						-0.089***	-0.963	
						(0.027)	(1.057)	
District council dummies	Yes	Yes	Yes	No	No	No	No	Yes (Prov)
Age of oldest members	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4372	694	694	496	199	199	198	2541
Kleibergen-Paap F statistics		14.918	14.866	11.033	10.865	10.820	11.064	12.452
Durbin-Wu-Hausman (p-value)		0.012	0.013	0.006	0.030	0.050	0.019	0.000

Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. All regressions include the squared age of the household head. Tests of overidentifying restrictions performed using both the year of arrival and its square do not reject the null hypothesis that the instruments are valid. [a] consider only households living in the homeland. [b] and [c] excludes households that moved within the homeland. [d] The dependent variable is per capita food consumption. [e] uses the Rural Survey 1997.

The coefficient of the land variable increases notably when land is instrumented with the year of arrival and the Durbin-Wu Hausman test suggests that 2SLS results are to be preferred to standard OLS. The downward bias of the OLS estimates could be explained by the fact that if more disadvantaged households, in particular those facing unfavourable circumstances in the labour market, choose to engage in agricultural activities, and therefore to access the land, the estimated effect of land on welfare would be small or even negative, as suggested in Carter and May (1999). Therefore neglecting this source of endogeneity would provide a more pessimistic view of the relationship between land and welfare. The instrumental variable estimates reveal that the effect of access to land on welfare is large. The increase in welfare of around 2.5 units is sufficient, on average, to shift a household from the lowest to the top quintile of the distribution of welfare.

In the IV specifications discussed so far, about 70% of the households considered live in a former homeland territory. Column 4 reports the results when only household living in the former homelands are considered. The KwaZulu homeland comprises a large number of non-contiguous parts spread throughout the KwaZulu-Natal province. The province created in 1994 incorporates the former homeland of KwaZulu and the surrounding province of Natal. Households in the sample are assigned to the former homeland on the basis of the magisterial district of residence. The identification of the magisterial districts belonging to the former homeland has been done using the information provided by Cox (2004) and the map that overlaps the KwaZulu homeland borders with magisterial district boundaries reported in Pauw (2005). The results reported in column 4 confirm previous findings. Although the sample size is further reduced the instrument maintains its explanatory power. This subsample, however, still considers both households that moved to and within the homeland, and therefore it may include households that voluntary changed location within the KwaZulu former homeland. This issue is expected to have a limited effect, since a large fraction of within-homeland movements are expected to be the result of government 'betterment planning'. According to Platzky and Walker (1985), in fact, more than a million people have been moved as a result of 'betterment plans' in KwaZulu from 1950 to 1985. To provide further support to the results, however, the estimates reported in column 5 are obtained by further restricting the sample to those households which migrated from non-homeland areas given the lower probability of encountering voluntary migration in this subsample. The coefficient of the land variable remains stable and significant and no relevant

changes are observed for the other explanatory variables.

Given the availability of consumption data in the KZN-DIHS survey, column 7 reports the results of the same specification in column 6 where, however, the dependent variable is food consumption per adult equivalent computed using the OECD equivalence scale[3]. When using this alternative measure of household welfare, access to land is still found to have a positive effect. This specification offers the opportunity to provide an economic interpretation of the results. Obtaining access to land has a large effect on household welfare by generating an increase in per-adult equivalent food consumption close to its median value. Finally, the last column reports the results when the same specification is applied to the Rural Survey 1997 dataset. Results appear to be in line with the findings obtained using the KZN-DIHS although some issues related to the Rural Survey dataset need to be further addressed as it is done in the following section.

Although the main focus of this analysis is the impact of access to land on household welfare, some useful insights can be obtained by looking at the effects of the other covariates. Education plays an important role in contributing to household welfare. This is shown by the positive and significant effect, throughout most of the specifications, of the education level of the household head and of the highest educational attainment of the household members. It is also reflected in the negative effect of the number of unskilled members in the households that is also likely to capture the effect of the lack of labour market opportunities for less educated household members.

8 Estimation results: land size and household welfare in the former homelands

In this section I explore the relationship between land size and welfare using the subsample of landed households provided by the Rural Survey 1997. I consider only households with access to land since the aim of the analysis conducted in this section is to analyse the welfare-improving effects of land size on households involved in land-based activities excluding those that could be potentially engaged into completely different livelihood strategies (Finan et al., 2005). In the first column of table 6, I report the estimates

[3]The OECD equivalence scale assigns a value of 1 to the first member of the household, 0.7 to each additional adults and 0.5 to additional children in the household.

of the ordinary least squares estimation of equation 2. The dependent variable is the asset index constructed using data from the 1997 Rural Survey and summarised in the last column of table 3. The results show a positive correlation between the amount of land owned and household welfare. The remaining columns report the two-stage least squares estimates and consider only households that moved to the current location during the period 1913-1994. The first-stage regressions show the negative correlation between land size and the year of arrival. The first-stage regression associated with column 2 is reported in table 9 in the appendix. The F statistics reported at the bottom of the table confirms the relevance of the instrument. The results using instrumental variable show that an additional hectare of land, produces an increase of 0.610 in the welfare index, which is sufficient, on average, to cause a shift to a higher decile of the welfare distribution.

Table 6: OLS and 2sls regressions of land size on household welfare

	OLS	IV					
	(1)	(2)	(3)	(4)	(5)[a]	(6)[b]	(7)[c]
Land (hectares)	0.009*	0.608**	0.616**	0.612**	0.935**	0.420**	0.380**
	(0.005)	(0.240)	(0.254)	(0.249)	(0.425)	(0.213)	(0.160)
Education of household head	0.071***	0.063***	0.063***	0.063***	0.072**	0.054***	0.059***
	(0.007)	(0.016)	(0.016)	(0.016)	(0.028)	(0.013)	(0.014)
Age of household head	0.028***	0.006	0.003	0.004	-0.008	0.013	0.029
	(0.009)	(0.026)	(0.033)	(0.033)	(0.057)	(0.029)	(0.021)
Gender of household head (dummy)	0.050	-0.259	-0.367*	-0.370*	-0.600	-0.239	-0.385*
	(0.049)	(0.166)	(0.218)	(0.220)	(0.395)	(0.188)	(0.199)
Pension eligible members (dummy)	-0.055	-0.096	-0.065	-0.070	-0.099	-0.066	-0.154
	(0.069)	(0.134)	(0.146)	(0.146)	(0.244)	(0.129)	(0.133)
Children age 1-5	-0.025	0.022	0.024	0.024	0.091	0.019	0.041
	(0.024)	(0.052)	(0.053)	(0.053)	(0.087)	(0.046)	(0.046)
Children age 6-17	-0.013	0.001	-0.001	-0.000	0.039	-0.019	0.004
	(0.016)	(0.032)	(0.032)	(0.032)	(0.058)	(0.025)	(0.028)
Number of skilled members	0.101***	0.067	0.063	0.063	0.035	0.059	0.035
	(0.026)	(0.050)	(0.055)	(0.055)	(0.087)	(0.047)	(0.047)
Number of unskilled members	-0.057**	-0.052	-0.057	-0.057	-0.046	-0.078**	-0.103**
	(0.023)	(0.039)	(0.040)	(0.040)	(0.068)	(0.038)	(0.040)
Highest level of education	0.068***	0.030	0.029	0.029	-0.004	0.053***	0.050*
	(0.011)	(0.034)	(0.034)	(0.034)	(0.059)	(0.019)	(0.029)
Labor market dev index (district level)	0.873***	1.774***	1.786***	1.711***	1.469**	1.426***	1.382***
	(0.110)	(0.479)	(0.495)	(0.421)	(0.655)	(0.363)	(0.318)
Population density (district level)	0.000**	0.001*	0.001*	0.001*	0.002*	0.001	0.001
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
Land quality index				0.070	0.323	0.061	0.057
				(0.127)	(0.224)	(0.124)	(0.094)
Province dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age of oldest members	No	No	Yes	Yes	Yes	Yes	Yes
Observations	3775	2736	2736	2736	2136	2359	2078
Kleibergen-Paap F statistics		12.398	11.016	11.381	6.314	12.898	16.647
Durbin-Wu-Hausman (p-value)		0.000	0.000	0.000	0.000	0.021	0.004

Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. All regressions include the squared age of the household head. Tests of overidentifying restrictions performed using both the year of arrival and its square do not reject the null hypothesis that the instruments are valid. [a] excludes the homelands of Venda and Transkei, [b] Period 1913-90, [c] excludes rounded decades.

Column 3 reports the results when additional controls for the age structure of the household are included, namely a polynomial of the age of the oldest man and woman in the household. This is also included in the subsequent specifications. All regressions include province dummies to control for differences in environmental and other local conditions. In column 4 a variable capturing the variation in land quality across districts is included and is intended to control for the potential correlation between the year of arrival and the quality of the land, although it does not capture plot-specific quality. The results are in line with previous findings and the coefficient of the land quality index is not significant. This can be explained by the fact that the land in the former homelands is in general of poor quality with little variation within the territory (Desmond, 1971).

Because the reasons for the move are unknown, one of the main concern is that forced removals could be confounded with voluntary migration. Voluntary relocations were more likely to occur within the homelands since, as previously mentioned, conditions in the homelands were extremely unfavourable and the available descriptive evidence suggests that no households would voluntary move to these overcrowded and unpleasant areas. Unfortunately, it is not possible to establish whether the household moved to or within the homeland of current residence since the 1997 Rural Survey does not provide information on the previous place of residence. It is worth noting that massive forced relocations were also implemented within the homeland territories, often motivated by 'betterment plans' implemented since 1930. Therefore movements within the homelands are also likely to be the result of coercive government policies although no direct evidence is available. To further address this problem I use the 1996 South Africa population census which provides information on the year of arrival in the current location and the district of previous residence. Using this information it is possible to exclude from the analysis those areas with the highest percentage of within-homeland movements and, therefore, potentially with the highest probability of voluntary movements. Table 8 reports the distribution of movements by homelands and distinguishes between "within-" and "to-" homeland migration. According to these figures, the two former homelands of Transkei and Venda have the highest percentage of within-homeland movements, since 89% and 86% of the households that arrived in the current location during the period 1913 - 1994 were previously residing within the same homeland. In column 5 of table 6, households living in these two homelands are excluded from the sample. The estimates reported confirm previous findings although

the F statistic of the first-stage regression is now lower due to the reduced sample size. Similar results, reported in table 7, are also found when different sub-periods are considered, in particular when households that moved before 1930, 1948 and 1958 are subsequently removed from the sample. These dates correspond to the main events that affected the process of forced removals: 'betterment plans' were introduced in 1930, while apartheid officially started in 1948 and 'black' people were officially assigned to a homeland territory in 1958.

Another potential problem may arise if forced removals have a direct welfare cost, independent of land endowments. As previously mentioned, this issue is partially ruled out by considering only households that arrived before the end of the apartheid, that is only households that have resided in the current location for at least 3 years. A further exercise is conducted by excluding from the sample those households that arrived in the current location before 1990. Although post-arrival tangible and intangible displacement costs can affect household performances and consequently household welfare, it is reasonable to expect that after at least 7 years of residence in the same location the household can overcome the initial difficulties. The results reported in column 6 confirm previous findings. The coefficient of the land variable, although reduced in size is still positive and significant. This confirms that the instrument is not capturing the effect of the displacement costs associated with the length of the residence in the current location.

Because figure 1 reveals that households tend to report the year of arrival in rounded decades, I conduct a further exercise by dropping those households that could be potentially misreporting the year of arrival, i.e. those households that arrived in 1920, 1930, 1940, 1950, 1960, 1970, 1980 and 1990. The instrument improves the ability to predict the amount of land held by the household, so that the F statistic of the first stage is now above 16 and the results are again similar to previous findings. Finally, because those homelands that obtained independence, namely Transkei in 1976, Bophuthatswana in 1977, Venda in 1979 and Ciskei in 1981, were rewarded by the government with the building of roads, shopping centres and hotels (Platzky and Walker (1985), p 23) I also run an additional specification (reported in table 7) including a dummy variable indicating whether or not the homeland obtained independence and the results are almost unchanged.

In line with the results reported in the previous section, the education of the household head positively affects household welfare. The number of unskilled members has a negative effect, although not always significant,

Table 7: Additional results on the effect of land size on household welfare

	1930-1994	1948-1994	1958-1994	1913-1994
	(1)	(2)	(3)	(4)
Land (hectares)	0.609**	0.537**	0.492**	0.628**
	(0.248)	(0.235)	(0.198)	(0.259)
Education of household head	0.063***	0.062***	0.069***	0.065***
	(0.016)	(0.016)	(0.016)	(0.016)
Age of household head	0.009	0.017	0.020	0.005
	(0.033)	(0.031)	(0.030)	(0.034)
Gender of household head (dummy)	-0.368	-0.385*	-0.324	-0.375*
	(0.227)	(0.229)	(0.204)	(0.226)
Pension eligible members (dummy)	-0.057	-0.049	-0.043	-0.057
	(0.149)	(0.145)	(0.143)	(0.147)
Children age 1-5	0.024	0.029	0.031	0.027
	(0.055)	(0.054)	(0.052)	(0.056)
Children age 6-17	0.005	-0.010	-0.012	-0.002
	(0.033)	(0.032)	(0.031)	(0.032)
Number of unskilled members	-0.044	-0.050	-0.049	-0.060
	(0.041)	(0.042)	(0.042)	(0.041)
Number of skilled members	0.069	0.052	0.040	0.061
	(0.056)	(0.054)	(0.054)	(0.056)
Highest level of education	0.030	0.035	0.035	0.029
	(0.035)	(0.034)	(0.032)	(0.035)
Labor market dev index (district level)	1.741***	1.721***	1.621***	1.744***
	(0.432)	(0.414)	(0.369)	(0.439)
Population density (district level)	0.001*	0.001*	0.001**	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Land quality index	0.065	0.021	0.018	0.103
	(0.126)	(0.126)	(0.118)	(0.154)
Independence (dummy)				-0.281
				(0.307)
Observations	2649	2479	2328	2736
Kleibergen-Paap F statistics	10.872	9.576	10.975	10.829
Durbin-Wu-Hausman (p-value)	0.000	0.001	0.000	0.000

All regressions include also the square of the age of the head of the household. Tests of overidentifying restrictions performed using both the year of arrival and its square do not reject the null hypothesis that the instruments are valid.

probably signalling the presence of constraints in the labour market for less educated household members. The positive and significant effect of the employment rate at district level indicates how the development of the local labour market can positively influence household welfare. Finally, households

Table 8: Movements of households in former homelands

Former homelands	% moved within the homeland area	% moved from other areas
Kwazulu	56	44
Bophuthatswana	62	38
KaNgwane	34	66
KwaNdebele	19	81
Transkei	89	11
Ciskei	43	57
Venda	86	14
Ganzankulu	63	37
Lebowa	73	27
Qwaqwa	22	78

Source: author's calculation from the South Africa Census 1996

with a male head are worse off in comparison to female-headed households. This could be related to the fact that, in rural areas, male heads usually tend to migrate to urban centres, and therefore their presence in the household could signal a lack of off-farm sources of income.

Although the paper finds a positive relationship between land endowments and welfare, it is not possible to identify how these effects are transmitted. The high share of households producing mainly for home consumption suggests that land can benefit them by providing a cheaper source of food. However, other mechanisms could be in action and cannot be disentangled without further investigations.

9 Conclusions

This paper explores the relationship between land endowments and household welfare. Although economic theory supports a positive relationship between land and welfare, little empirical evidence is available mainly due to the difficulties in identifying the causal relationship between land and a measure of household welfare. The potential endogeneity of the land variable is here addressed using historical data on migration to the former homelands. The availability of data on the year of arrival in the current location reveals a negative correlation between land endowment and arrival date that is in line with records on increasing population pressure, and therefore with increasing

scarcity of land, in these areas since the introduction of the Native Land Act in 1913. The year of arrival is expected to be independent of households' unobserved ability to generate welfare. Movements to the homelands, in fact, can be attributed to the massive forced removals conducted by the central government with the aim of segregating the African population into different homelands according to their ethnic background. Movements within the homelands can also be largely explained by government 'betterment planning' for the reorganisation of the territory in the homelands.

The empirical specification adopted in this paper assumes a linear relationship between land size and household welfare that fails to capture the potential non-linear effects of land endowments. Finan et al. (2005), for example, argue that credit and labour market imperfections can affect the ability of the household to maintain production intensity when land area increases. Therefore the relationship between land and household welfare seems to follow a more complex pattern. Non-linear analyses, however, often require non-parametric techniques or non-linear specifications where the presence of potential endogenous explanatory variables requires the use of less conventional and more complex solutions, when possible. Nevertheless the relevance of such heterogenous effects leaves room for further investigations on the relationship between land and welfare across different dimensions of the farm household.

Results show the positive effect of land access on household welfare. Land size is also positively related to household welfare so that an increase of 1 hectare is expected to lift the household into a higher decile of the welfare distribution. A set of alternative specifications control for the presence of confounding effects produced by the potential correlation between the year of arrival and the location of the household, the displacement costs occurring in the early years after arrival and the quality of the land. Results are also robust to alternative specifications. The positive relationship here identified cannot, however, be attributed to one or more transmission mechanisms and again leaves room for further investigations. Nevertheless, these results suggest that reforms aimed at improving access to land, a major concern of post-apartheid governments, have the potential of improving household welfare. Moreover, because the households considered in this analysis are living in relatively disadvantaged and less fertile areas - the homelands - these results are likely to provide a lower bound for the positive effects of land access on household welfare.

References

Angrist, J. D. (2001). Estimation of limited dependent variable models with dummy endogenous regressors: simple strategies for empirical practice. *Journal of Business and Economic Statistic 19*(1), 2–16.

Besley, T. and R. Burgess (2000). Land reform, poverty reduction, and growth: Evidence from India. *The Quarterly Journal of Economics 115*(2), 389–430.

Burgess, R. (2001). Land and welfare: Theory and evidence from China. Working paper, Department of Economics and STICERD, London School of Economics.

Carter, M. R. and J. May (1999). Poverty, livelihood and class in rural South Africa. *World Development 27*(1), 1–20.

Cox, K. R. (2004). Urbanization in South Africa and the changing character of migrant labour. *South African Geographical Journal 86*(1), 7–16.

de Wet, C. (1987). Betterment planning in South Africa: some thoughts on its history, feasibility and wider policy implications. *Journal of Contemporary African Studies 6*(1/2), 85–122.

de Wet, C. (1994). Resettlement and land reform in south africa. *Review of African and Political Economy 61*, 359–373.

Desmond, C. (1971). *The discarded people: an account of African resettlement in South Africa*. Harmondsworth: Penguin.

Eswaran, M. and A. Kotwal (1986). Access to capital and agrarian production. *The Economic Journal 96*(382), 482–498.

Filmer, D. and L. H. Pritchett (2001). Estimating wealth effects without expenditure data - or tears: an application to educational enrollments in states of India. *Demography 38*(1), 115–132.

Finan, F., E. Sadoulet, and A. de Janvry (2005). Measuring the poverty reduction potential of land in rural mexico. *Journal of Development Economics 77*, 27–51.

Freund, B. (1984). Resettlement and the political economy of south africa. *Review of African Political Economy 29*, 49–63.

Keswell, M., M. Carter, and K. Deininger (2010). Poverty and land redistribution: Quasi-experimental evidence from south africa's lrad program. Mimeo, Stellenbosch University.

Lipton, M. and M. Lipton (1993). Creating rural livelihoods: some lessons for South Africa from experience elsewhere. *World Development 21*(9), 1515–1548.

Lyne, M. and M. Darroch (2004). Land redistribution in Kwazulu-Natal, South Africa: Five census surveys of farmland transactions, 1997-2001. Basis crsp, Department of Agricultural and Applied Economics.

McKenzie, D. J. (2005). Measuring inequalit with asset indicators. *Journal of Population Economics 18*(2), 229–260.

Murray, C. (1987). Displaced urbanization: South africa's rural slums. *African Affairs 86*, 311–29.

Pauw, K. (2005). Forming representative household and factor groups for a south african sam. Technical paper series, PROVIDE Project.

Platzky, L. and C. Walker (1985). *The Surplus People: Forced Removals in South Africa*. Johannesburg: Rivan Press.

Simkins, C. (1983). *Four essays on the past, present and possible future of the distribution of the black population in South Africa*. Cape Town: Saldru.

Singh, I., L. Squire, and J. Strauss (1986). *Agricultural Household Models*. Baltimore: Johns Hopkins University Press.

Valente, C. (2009). The food (in)security impact of land redistribution in South Africa: microeconometic evidence from national data. *World Development 37*(9), 1540–1553.

Appendix

Table 9: First stage regressions

	(1)	(2)
	Land (dummy)	Land (hectares)
Year of arrival	-0.009***	-0.012***
	(0.002)	(0.003)
Education of household head	-0.003	0.003
	(0.007)	(0.003)
Age of household head	-0.002	0.030
	(0.010)	(0.036)
Gender of household head (dummy)	0.014	0.556***
	(0.045)	(0.155)
Pension eligible members (dummy)	-0.120	0.054
	(0.081)	(0.190)
Children age 1-5	0.060*	-0.085
	(0.035)	(0.064)
Children age 6-17	0.048***	-0.039
	(0.015)	(0.036)
Number of skilled members	0.063***	0.019
	(0.021)	(0.064)
Number of unskilled members	0.064***	-0.051
	(0.017)	(0.046)
Highest level of education	0.003	0.051
	(0.003)	(0.048)
Labor market dev index (district level)	-0.257**	-1.440***
	(0.111)	(0.004)
Population density (district level)	-0.000***	-0.002***
	(0.000)	(0.001)
Observations	695	2736

Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
All regressions include also the square of the age of the head of the household.