DataFirst

DataFirst Technical Papers

Household formation and household size in post-apartheid South Africa: Evidence from the Agincourt sub-district 1992-2003

by

Martin Wittenberg and Mark Collinson

Household formation and household size in post-apartheid South Africa: Evidence from the Agincourt sub-district 1992-2003[*]

Martin Wittenberg
School of Economics and DataFirst
University of Cape Town

Mark Collinson
(1) MRC/Wits University Rural Public Health
and Health Transitions Research Unit (Agincourt)
School of Public Health, Faculty of Health Sciences,
University of the Witwatersrand, South Africa
(2) Migration, Urbanisation and Health Working Group
INDEPTH Network

October 2014

Abstract

South African national datasets suggest a rapid reduction in household size. However much of this seems to be concentrated over an implausibly short period between 1998 and 2000. We examine the national evidence by accounting for the undersampling of small households in the 1994-1998 period. We also examine the patterns of household change in a more limited context where we have high quality continuous data over this period.

We use the data from the Agincourt Health and Demographic Surveillance System to this end. Our reweighted national data as well as the Agincourt data confirm that households have become smaller over this period,

1

by about half a person, but the process is not as discontinuous as suggested by the "raw" figures.

Because the Agincourt data are longitudinal we are also able to examine some of the mechanisms by which the reduction in household size occurs. To that end we develop a novel decomposition technique. We show that the overall reduction is fuelled by rapid household formation and that much of this seems associated with the public provision of housing and an attempt by households to gain better access to services. Changes in the legal rights of previously marginal groups and in the system of development controls are also likely to have been important.

Key words: South Africa, Agincourt, household size, household formation, survey data

1 Introduction

The end of apartheid led to changes on many fronts: economic, social and political. One dimension which has not received equal attention is in the composition of households and in particular, a reduction in average household size. The core pattern is shown in Figure 1. It suggests that between the late 1990s and 2003 households lost, on average, one full member. Since household size is a ratio of two variables, total population and number of households, this reduction can occur due to changes in the numerator, i.e. population (e.g. increased mortality due to the HIV pandemic) or the denominator (new household formation). Many different social processes are therefore likely to bear on it: demographic processes such as mortality, fertility and age of childbearing (Burch 1970), but also social and economic processes that affect the affordability and desirability of living alone (Börsch-Supan 1986, Ermisch and Salvo 1997, Haurin, Hendershott and Kim 1993). Household size can be seen as prism through which these social processes are refracted.

There is, of course a prior measurement issue. A key question when confronted with such dramatic changes is whether they are "real" or just artefacts of changes in the instrument. Unfortunately there is no independent benchmark at the national level to check these trends against. We do, however, have an extraordinarily rich data source that allows us to analyse these changes in detail in a local area. The MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt) has been collecting information on all households and individuals in a rural area in the east of South Africa since 1992. The data from this Health and Demographic Surveillance System site (HDSS) enables us to go beyond the broad national changes to examine how the process of household size reduction has worked in detail. To that end we develop a novel decomposition technique. So while household size is the prism through which broader social developments are refracted, the MRC/Wits Agincourt HDSS provides the spectroscope through which we can isolate some of the component processes.

The contribution of this paper are therefore threefold. Firstly we provide new evidence against which the national trends in household size reduction can be assessed. Secondly we present a new technique for analysing that reduction. Thirdly we provide a pointer to some of the mechanisms that have been driving that process.

The plan of this discussion is as follows. In the next section, we briefly review some of the literature that has examined South Africa's national data. Section 3 describes the data that we use in more detail. We then describe our methods in Section 4, in particular the new decomposition technique. Section 5 provides the results of our analysis. We provide an interpretation of these trends in Section 6. We conclude by reflecting on what these local processes may suggest about the
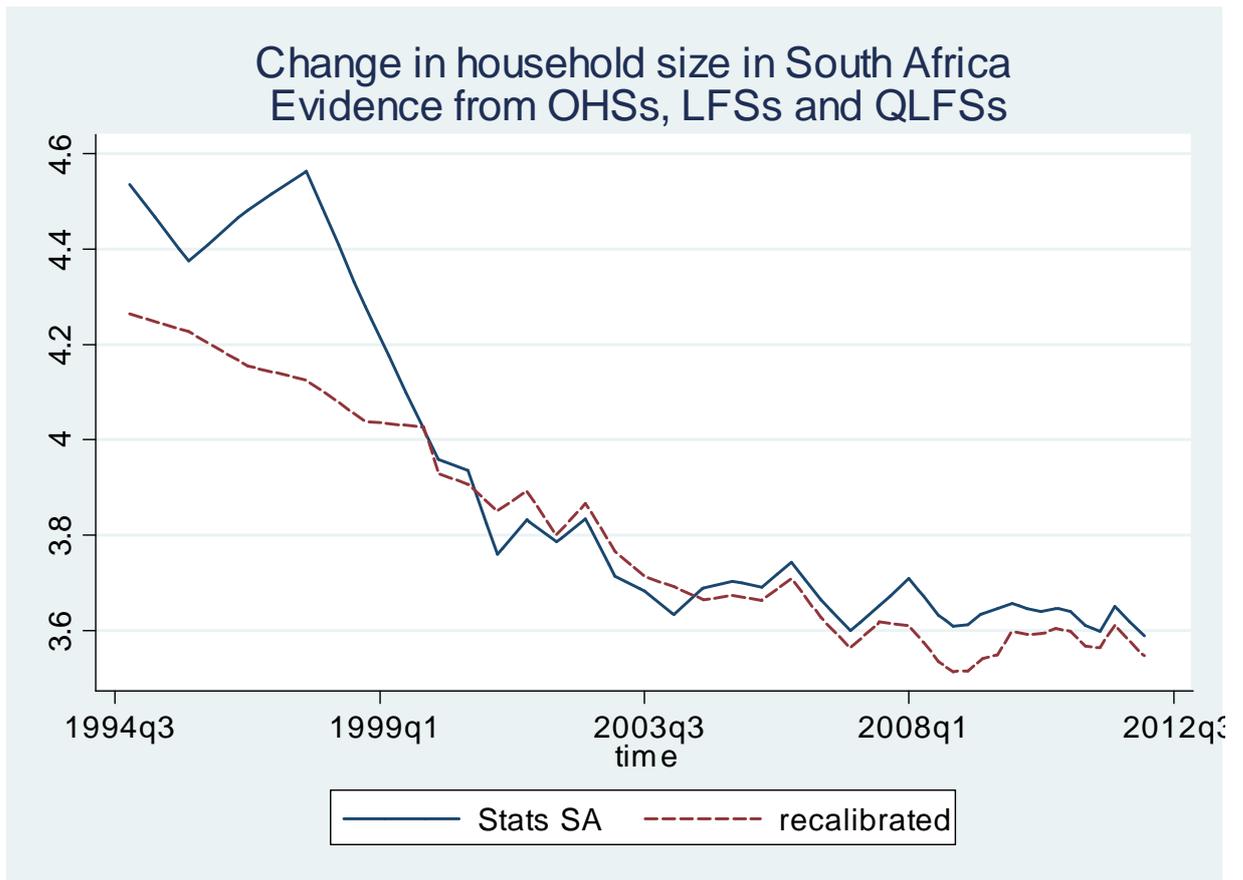
Figure 1: Average household size has decreased dramatically since 1994 according to national surveys.

national picture.

2 Household change and household size in national surveys

The literature discussing the decline in household size has tended to focus on the question whether South African households are becoming more nuclear or "westernised" (Ziehl 2001, Amoateng and Kalule-Sabiti 2008). Russell (2003b, 2003a), however, has argued that it is not clear that the instruments used in measuring household size (the census or sample surveys) adequately cover the complexity of the social connections between people. The problem lies, in particular, with the fact that social surveys tend to take a snap shot of where people are located at that point in time and do not indicate that people tend to move between households and locations. Posel, Fairburn and Lund (2006) point out the importance of such rural-urban linkages in the context of analysing employment and migration behaviour.

In a different vein Wittenberg and Collinson (2007) have pointed out that the definition and measurement of the "household" is not the only issue in analysing national datasets. They show that there seems to be a major increase in one person households in the period 1998 to 2000 (shown also by the steep decline in household size over that period in Figure 1). They describe this as a "a veritable explosion in solitary living" (Wittenberg and Collinson 2007, p.135) and doubt that it could be a true reflection of national trends. More recently Kerr and Wittenberg (2013) have suggested that in the early national household surveys, i.e. the October Household Surveys up to and including 1998, small households were undersampled. The instructions to fieldworkers was to interview only one household at each address and, if there were more than one, to select the households with probability proportional to size. They find no evidence that smaller households were weighted up to compensate for this undersampling. While this discussion resolves one puzzle, i.e. the reason for the precipitous decline in household size, it raises a whole host of new questions: did household size decline at all over this period? If so, by how much? And what could have produced this trend?

There are several candidate explanations. The increased mortality associated with the HIV epidemic or the decrease in the fertility rate (Moultrie and McGrath 2007) would all be expected to produce declines in the average household size in the long run. Nevertheless the mechanism by which this process would work would not be the one in which new household formation outstrips the population growth rate, which is the pattern that we will show below. Indeed a rapid rate of household formation raises additional issues given that economic conditions

in the late 1990s were arguably tough. Economic approaches to the analysis of the household emphasise that the decision to set up an independent household would tend to go up with income (Börsch-Supan 1986, Ermisch and Salvo 1997, Haurin et al. 1993). In tough economic conditions the reverse would occur: dependent children will delay moving out of the parental home, or might even move back. Indeed one strand of the South African labour literature has argued that unemployment has led to higher levels of co-residence with pensioners than might otherwise have been the case for these sorts of reasons (Klasen and Woolard 2009).

Given these difficulties it is important not only to produce analyses that can confirm what has happened to South African households, but that can also point to some of the mechanisms that might have produced that outcome.

3 The Data

We will use two types of data for our analyses: nationally representative sample surveys collected by South Africa's offical statistical agency and the data from the Agincourt Health and Demographic Surveillance System site.

3.1 South Africa's national household surveys

Statistics South Africa (and its precursor, the Central Statistical Services) has been conducting annual nationally representative sample surveys on a range of socio-economic issues since 1994. All of these surveys are multi-stage stratified and clustered instruments. Typically they survey around 30 000 households in around 3 000 clusters, although there has been some variation in this design. During the 1990s these surveys were conducted annually in October and were therefore referred to as the "October Household Surveys" (OHSs). This series was discontinued and replaced in 2000 with a more focussed bi-annual Labour Force Survey (LFS). The Labour Force Surveys were, in turn, replaced by the Quarterly Labour Force Surveys (QLFSs) in 2008. On top of this the "General Household Survey" has been conducted annually since 2002 and a range of more specialised surveys (such as the Income and Expenditure Surveys, and the Living Conditions Surveys) have been conducted less frequently.

Because they offer an unbroken series (particular over the crucial late 1990s early 2000s) we focus only on the OHS, LFS, QLFS series. There are several known issues in relation to this series. We have already mentioned the evidence from Wittenberg and Collinson (2007) about the undersampling of small households in the OHSs. Branson and Wittenberg (2014) point out that the household weights released with the OHSs are not aligned with the person weights. Furthermore there are also breaks in the demographic model underpinning the person weights, so that
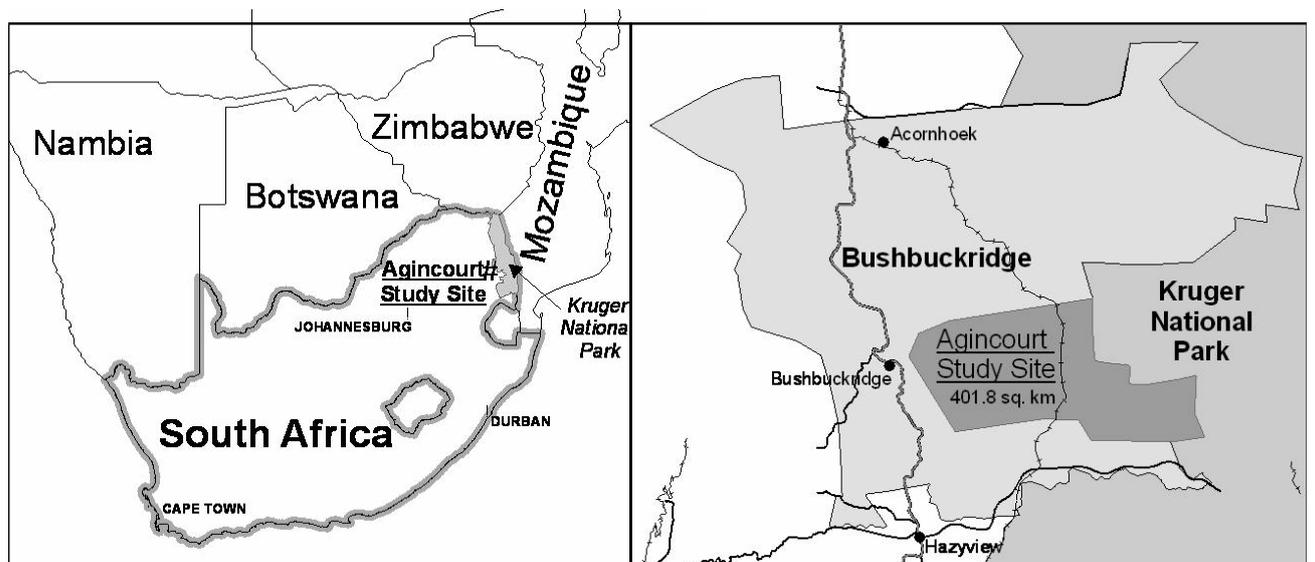
Figure 2: The Agincourt field site covers 21 villages in the Bushbuckridge area

there are shifts in some of the aggregates which are due purely to these changes in assumption. Branson and Wittenberg (2014) suggest that a recalibration of the original Statistics South Africa weights can deal with many of these idiosyncratic shifts. Machemedze, Kerr and Wittenberg (2014) extend this approach to deal with the undersampling of small households in the OHSs.

For the purposes of our analyses we use the PALMS version of the data (Kerr, Lam and Wittenberg 2013) which contains a set of harmonised weights to deal with the shifts pointed out by Branson and Wittenberg (2014). We make the further adjustments suggested by Machemedze et al. (2014). To facilitate comparison with the Agincourt data we will have occasion to restrict our analyses to the "rural" subsample of the national datasets. Unfortunately a change in the mastersample in 2004 means that a "rural" indicator is not available after that period. As Figure 1 indicates, however, the period from 1994 to 2003 is in fact the most interesting one in relation to national trends.

3.2 The Agincourt demographic surveillance data

The MRC/Wits Agincourt Unit was established in 1992 with the aim of addressing issues around the decentralisation of health services and to provide accurate information for planning (Tollman 1999, Tollman, Herbst, Garenne, Gear and Kahn 1999). The strategy was to conduct health and demographic surveillance, underpinning a programme of inter-disciplinary health and population research.

Agincourt was selected in part because it reflects many of the key developmental challenges. The area lacks a functioning vital registration system, thus making on-going demographic surveillance appropriate. Furthermore, the area formed part of the previous Gazankulu homeland and therefore exhibits many of the characteristics of these areas: a lack of infrastructure and a population that has been subject to forced removals and betterment planning (for a discussion of some of these processes see Niehaus 2001).

Agincourt is a sub-district of the Bushbuckridge region of the Limpopo Province (see Figure 2). The site is particularly interesting, since it is close to the Mozambique border and has a significant subpopulation of Mozambican refugees. These refugees arrived in the late 1980s during Mozambique's civil war. They come from the same language group as the South Africans, but they form a distinct subpopulation. Indeed, many of them live in villages which consist predominantly of refugees.

In our empirical work we work with a four-fold categorisation of subpopulations: Mozambicans living in refugee villages; South Africans living in "South African" villages; Mozambicans living in "South African" villages and the RDP village. The latter is a settlement of formal cement-brick houses built with money from the government's Reconstruction and Development Programme (RDP). It was constructed in 1999 and was fully settled by 2002.

The refugee villages date back to the 1980s, when they were created to house refugees from the Mozambican civil war. They are all located on the fringes of the study area, furthest removed from infrastructure and from economic activities. Not coincidentally they are also located on the border of the Kruger National Park. Indeed most of the refugees came through that park from Mozambique. The "South African villages" go back to the 1950s and 1960s when the villages were laid out in terms of "betterment schemes". Within this category we distinguish between households headed by a South African citizen[1] and households headed by a non-South African (mainly Mozambican). The latter would be mainly ex-refugees that have managed to resettle themselves in more central locations.

Budlender (2003) has suggested that it can be completely misleading to classify households on the basis of the characteristics of a person described as the "head". Such classifications can hide some of the complexities in the nature of the underlying relationships. Indeed these relationships can be quite fluid. Even in our data set there are a few cases where the citizenship of the head changes[2]. We would argue that such a crude categorisation nevertheless captures a significant dimen-

---

[1] There are a few records where we cannot determine the citizenship of the head of the household. These cases have been pooled with the "South African" households.

[2] This happens in 218 households out of a total of 15 856. In most cases, however, the change was from "undefined" to something specific or *vice versa*. In order to maintain a consistent classification we simply ignored the changes and kept the original designation of the household.

sion of local reality. Residents of the area do distinguish between South Africans and the refugees and most households can fairly readily be assigned to one or the other category.

3.2.1 The Health and Demographic Surveillance System

The Agincourt HDSS monitors key demographic events and socio-economic variables in the Agincourt sub-district. A baseline census was conducted in 1992 and since then there have been seven census rounds in nine years. The main demographic, health and socio-economic variables measured routinely by the HDSS include: births, deaths, in- and out-migrations, household relationships, resident status, refugee status, education, antenatal and delivery health-seeking practices (Tollman 1999, Tollman et al. 1999, Collinson et al. 2002). Circular migrants are accounted for by including on the household roster non-resident members who retain significant contact and links with the rural home (Collinson, Tollman, Garenne and Kahn 2001). The "Share common pot" definition of a household is thus expanded to include the temporary migrants who would normally share the same pot on return. The definition of household head is the main household decision maker, as reported by the household respondent.

In the update rounds a trained lay fieldworker interviews the most competent respondent available at the time of visit. Individual information is checked for every household member. All events are recorded that have occurred since the previous census. Where possible, questions are directed to particular household members, for example, maternity history or pregnancy outcome information is asked directly from the woman involved, and a verbal autopsy is conducted with the person most closely involved with the deceased during the terminal illness. Revisits are undertaken when appropriate respondents are not available. Data quality checks include duplicate visits on 2% of households. In addition a number of validation checks are built into the fieldwork and data-entry programme. The software system used consists of a relational database constructed in Microsoft SQL Server.

3.2.2 Tracking households over time

In the HDSS system each household had an identified head. This person also served a reference function for recording relationship information. The variable "relationship to the head of household" was updated annually since 1996 for nearly all members of the population. If a head of household died or out-migrated a new set of references was constructed at the census following the change of household head.

A panel dataset was constructed for this analysis, using HDSS data. The

data, including household membership, were divided into one year intervals for the prospective period, viz. 1993 - 2003. For the study we took a household to be a dissolved if all household members moved out and started as new members moved in. If there was any overlap in membership between successive households we kept the same household identifier.

A limitation of this dataset is that over the observation period the HDSS did not have the capacity to retain a person's HDSS identity number if he or she moved within the study site.

4   Methods

4.1   Descriptive statistics

Our initial approach is purely descriptive, documenting the trends nationally, but particularly within Agincourt. Given the fact that we have a census each year in the Agincourt area, we will be much better able to characterise household formation and dissolution rates. Indeed since we don't see dissolving households in the national cross-sectional datasets we would obviously never be able to assess the balance between formation and dissolution from those sources.

4.2   Decomposing shifts in longitudinal data

There are several ways in which the reduction in household size might arise. It could be that the large households are supplying disproportionately many outmigrants or deaths, i.e. that large households are moving "down" the size distribution. Perhaps due to socio-economic changes, the largest households are being reconstituted, e.g. family groups leaving extended family settings and forming new households. It could also be that larger households simply cease to exist (e.g. due to outmigration) and that the new households that are formed are relatively small. Finally it is possible that if there are many more newly formed households than households going out of existence, and if these are smaller, then the larger proportion of new small households compared to old established ones will bring about a reduction in the overall average.

Given that we have panel data and not just a series of cross-sections, we can look inside households and see how these different mechanisms play themselves out. More particularly we propose an arithmetic decomposition of the change of household size into different effects.

Let $\overline{y}_t$ be the average household size in year t, $\overline{y}_t^S$ be the average size among households surviving to year $t+1$, $\overline{y}_t^T$ be the average among households terminating in year t and $\overline{y}_{t+1}^N$ be the average among households newly formed in year $t + 1$.

Then we have

$$\bar{y}_t = (1 - \theta)\,\bar{y}_t^T + \theta\bar{y}_t^S$$
$$\bar{y}_{t+1} = (1 - \phi)\,\bar{y}_{t+1}^N + \phi\bar{y}_{t+1}^S$$

where θ is the proportion of households surviving to period $t+1$ in the population at time t and ϕ is the proportion of survivors from period t at $t+1$. So

$$\Delta\bar{y}_{t+1} = \theta\Delta\bar{y}_{t+1}^S + (1 - \theta)\left(\bar{y}_{t+1}^N - \bar{y}_t^T\right) + (\theta - \phi)\left(\bar{y}_{t+1}^N - \bar{y}_{t+1}^S\right) \qquad (1)$$

This decomposition is not unique. We could as easily have written

$$\Delta\bar{y}_{t+1} = \phi\Delta\bar{y}_{t+1}^S + (1 - \phi)\left(\bar{y}_{t+1}^N - \bar{y}_t^T\right) + (\theta - \phi)\left(\bar{y}_t^T - \bar{y}_t^S\right) \qquad (2)$$

Unless there is a very rapid increase or decline in the number of households $\theta - \phi$ should be close to zero and the two decompositions should give similar results. In the empirical results we report the first decomposition. The second provides qualitatively similar results and is available on request from the authors.

We term the three effects

- The *within household change* effect $\theta\Delta\bar{y}_{t+1}^S$

- The *replacement* effect $(1 - \theta)\left(\bar{y}_{t+1}^N - \bar{y}_t^T\right)$, since the difference $\bar{y}_{t+1}^N - \bar{y}_t^T$ represents the effects of new households replacing ones going out of existence

- The *dilution* effect $(\theta - \phi)\left(\bar{y}_{t+1}^N - \bar{y}_{t+1}^S\right)$, since $\theta - \phi$ is non-zero only if there is a net change in the number of households and the term $\bar{y}_{t+1}^N - \bar{y}_{t+1}^S$ reflects how newly formed households differ from surviving ones. In a period of rapid household formation, the existing (surviving) households become a decreasing fraction of the entire population of households. Their contribution to the overall mean household size therefore becomes diluted by the new households.

4.3 Disaggregating the decomposition

Because of the complexity of the Agincourt field site – with different subpopulations, each likely to have their own household change dynamics – it will prove useful to break the decomposition down further and apply it to the different household types. Consequently we write each mean given in equation 1 as

$$\bar{y}_t^C = w_{1,t}^C\,\bar{y}_{1,t}^C + w_{2,t}^C\,\bar{y}_{2,t}^C + \ldots + w_{k,t}^C\,\bar{y}_{k,t}^C, \qquad C \in \{N, S, T\}$$

where we assume that we have k different types, $w_{i,t}^C$ is the weight of household type i within category C (i.e. surviving, terminating, newly formed households) and $\bar{y}_{i,t}^C$ is the mean household size of household type i in category C. We can therefore apportion each of the effects as follows:

1. The contribution of household type i to the *within household change effect* is

$$\theta \left(w^S_{i,t+1} \ \overline{y}^S_{i,t+1} - w^S_{i,t} \ \overline{y}^S_{i,t} \right)$$

2. The contribution to the *replacement effect* is

$$(1 - \theta) \left(w^N_{i,t+1} \ \overline{y}^N_{i,t+1} - w^T_{i,t} \ \overline{y}^T_{i,t} \right)$$

3. The contribution to the *dilution effect* is

$$(\theta - \phi) \left(w^N_{i,t+1} \ \overline{y}^N_{i,t+1} - w^S_{i,t+1} \ \overline{y}^S_{i,t+1} \right)$$

We apply this decomposition to four household types within the Agincourt area: Mozambicans living in refugee villages; South Africans living in "South African" villages; Mozambicans living in "South African" villages and the RDP village.

5 Results: Changes in household size in rural South Africa

5.1 Describing the patterns

Figure 3 presents our first evidence. The top line represents average household size over the eleven year period from 1992 to 2003. It is evident that household size has come down consistently, with perhaps a hint that the process was levelling off at the end. The two lower lines represent the picture from the national surveys, first with the weights as released by Statistics South Africa (the dashed line) and then secondly with the weights as recalibrated according to Machemedze et al. (2014). It is evident that the recalibration reduces the rate at which household size has come down over the period. Indeed the overall reduction is of a similar magnitude to the reduction in the Agincourt HDSS over that period, i.e. around half a person per household.

The second point that is noteworthy is that average household size in the Agincourt district is much larger than it is in the national datasets. There are two reasons for that. The first of these is that Agincourt is a "deep rural" location, whereas there will be other parts of the rural areas (e.g. farming areas in the Western Cape) that will have different characteristics. Secondly, the Agincourt HDSS has a more generous definition of household membership. In the context of analysing "real" changes to households that is a strength, since it will not remove temporary migrants from household rosters, thus producing apparent household size reductions, when in many ways the migrant is still integrally connected to that
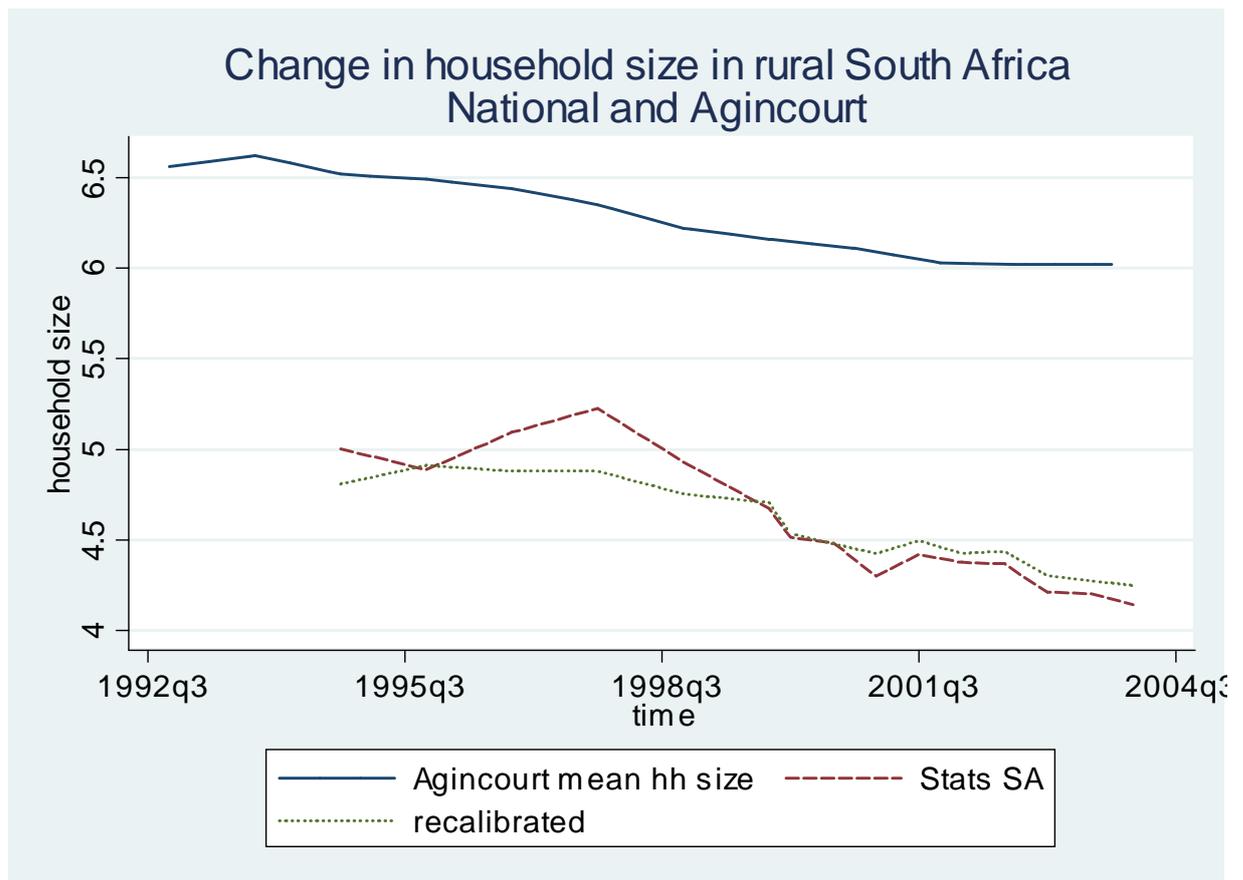
Figure 3: The pattern in the reduction in household size in Agincourt and nationally
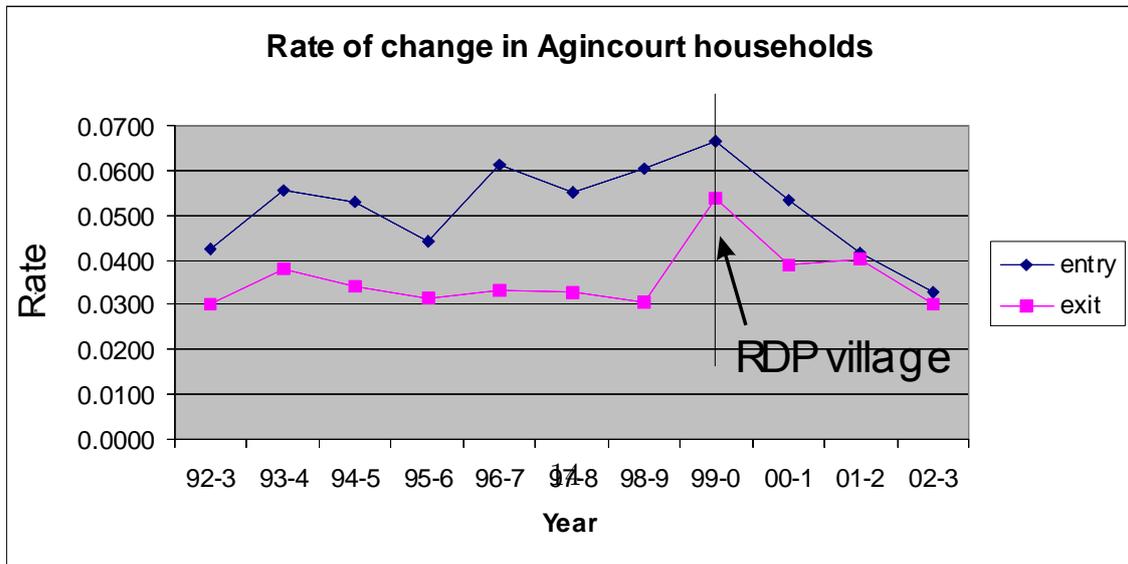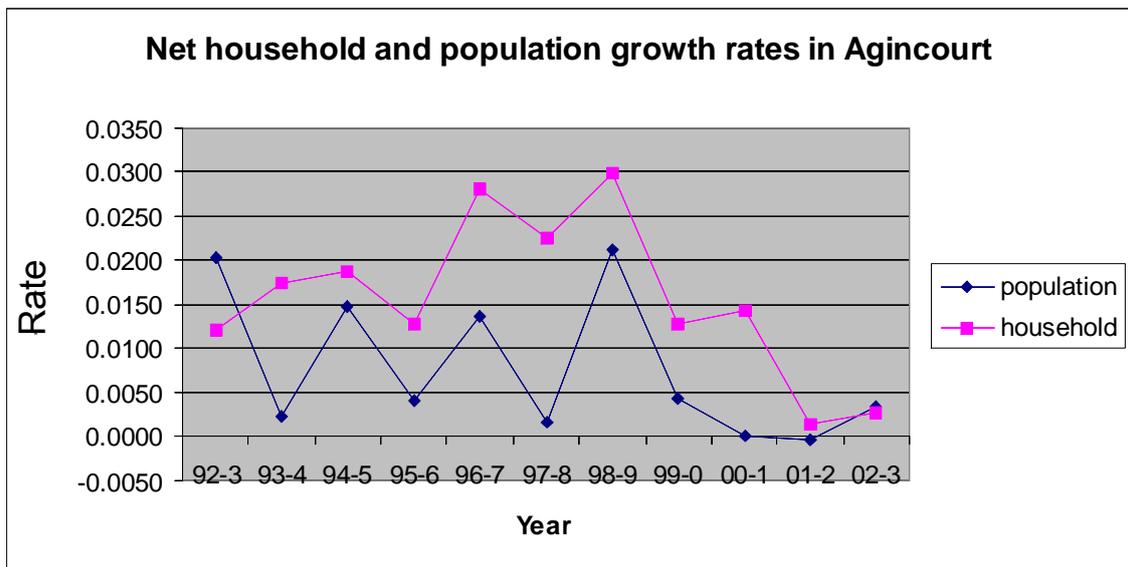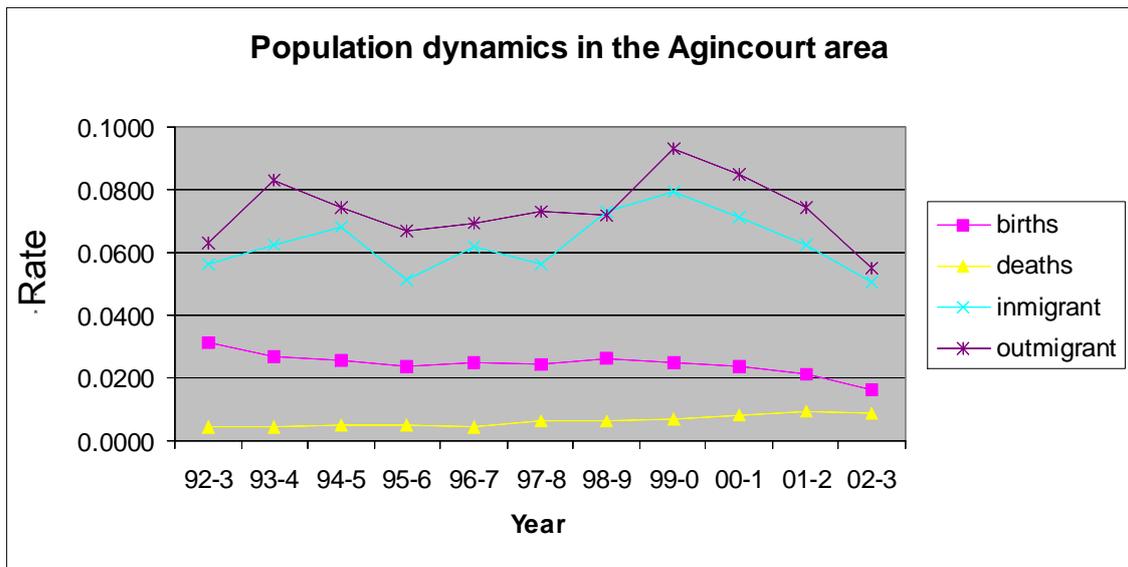
Figure 4: Aggregate Population and Household dynamics in Agincourt

household. The reduction shown in the Agincourt HDSS is therefore indisputably real.

In Figure 4 we present various ways of looking at population changes in the Agincourt sub-district over the period 1992–2003. The top panel gives the birth rate, death rate and in-migration and out-migration rates. We note that the latter two figures have to be treated with some caution, since movements within the Agincourt study site were not tracked. Every such move should be recorded twice on the HDSS data base: as an outmigration and as an inmigration. We observe that the birth rate is decreasing, while there is a noticeable increase in the death rate from 1998.

The middle panel of Figure 4 shows that the net population growth rate is around 0.8% per annum. This is, however, markedly lower than the growth rate in the number of households which averages around 1.7% per annum. The lowest panel shows that the turnover at the level of households is much higher than this net increase would suggest. Household formation rates seem to be around 5% while household dissolution rates are around 3.6%. As in the case of individuals, some of these dissolution and formation events will be entire households that relocate within the study site. Indeed we will suggest later that such internal moves are quite important and that they may provide the opportunity for some households to restructure themselves.

We can break these descriptive stats down by subpopulation. In Table 1 we provide information on the evolution of four settlement/household types within the Agincourt district. We distinguish between three kinds of villages: the RDP village, the "refugee" villages and the "South African villages". Within the last settlement type we distinguish between "South African" and "Mozambican" households.

The most startling implication to flow from Table 1 is that there seems to have been a gradual transfer of households from the refugee villages to the South African ones. We cannot draw this conclusion with certainty, since we cannot track people in our database from one location to another. Nevertheless the aggregate numbers are highly suggestive. While the refugee villages have lost 235 households between 1992 and 2003, there has been a net gain of 562 "Mozambican" households in the South African villages. Looking at the year by year figures is even more suggestive. In the year 1992 to 1993 there was a loss of 104 households from the refugee villages while there were 115 new Mozambican households in South African villages. In 1993-1994 the figures were 149 and 145 respectively and in 1994-95 they were 147 and 162.

5.2 Decomposing the changes

The decomposition of the change in household size for Agincourt is given in Table 2. The total change $\Delta \overline{y}_{t+1}$ is given in the first column. It suggests that average household change has dropped by 0.6 over this period. This represents a 9.1% reduction of household size over this period. The annual changes in this column suggest a slowing of the decrease in the last two years. Columns two and three give $\Delta \overline{y}_{t+1}^S$ and $\overline{y}_{t+1}^N - \overline{y}_t^T$ respectively. We see that changes in household size among surviving households (column 2) are rather small, whereas there seem to be rather big differences in the average household size of new versus terminating households. In column four we give $\overline{y}_{t+1}^N - \overline{y}_{t+1}^S$. These differences are the largest in the table – suggesting that new households are smaller than surviving ones by about 2.7 people.

The decomposition is given in the next three columns. At the bottom we have summed the contributions over the period 1993 to 2003. We note that the "within" household change effect is very small in all years and its total contribution is negligible. Furthermore it changes signs in a number of years. The "replacement" effect is more consistently negative (showing that new households are smaller than the ones they are replacing). It is also a little bit bigger than the "within" effect. Over the entire period it accounts for around 0.2 of a person. The "dilution" effects are markedly larger and account for 0.42 of the decline, which is over 70% of the overall reduction. It is the mushrooming of many new, smaller households rather than a reduction in the size of the existing households that has led to the overall reduction in household size.

5.3 Disaggregating the decomposition

We apply the aggregate decompositions (equation 1) to the four subpopulations within the Agincourt area. These decompositions are given in the middle of Table 3. Several points stand out:

- The South African households within the South African villages seem to have been actively shedding members over this period. Roughly one third of the reduction in household size within this group (from 6.36 to 5.9, as shown in Table 1) is due to surviving households actually becoming smaller. The remainder of the effect is due to dilution - the very rapid formation of smaller households within this category.

- The Mozambican households within the South African villages, by contrast, seem to have been absorbing members. This process is, however, offset by a very strong dilution effect so that the average household size among the Mozambicans in the South African villages has also come down.

- The Refugee villages show a strong negative "replacement effect". This shows that the newly formed households are significantly smaller than the dissolving ones. Indeed, the second row of Table 3 shows that the difference in size between newly formed and dissolving households is largest in the refugee villages. This suggests that the dissolving households in these settlements are relatively larger in size. The positive "dilution" effect is due to the fact that these settlements are not growing - they are shrinking, i.e. $\theta - \phi$ is negative.

- The RDP village shows a massive replacement effect. This is mainly due to the fact that in 1999 the newly formed households (i.e. **all** households in that village) had an average household size of 2.73 (see Table 1), while the previous household size was zero! The once-off change due to the formation of the village accounts for almost all of the replacement effect. More interesting is the positive "within" household change effect. This suggests that many of the initially small households acquired additional members. Indeed, as we noted earlier many of the one-person households seem to have been joined by partners or children.

The most important insight to be gained from these is that there is considerable diversity within the study site. Larger households within the refugee settlements are dissolving. There is the rapid formation of smaller households within the South African villages and within the RDP village. South African households within the South African villages are shedding members, while Mozambican households within these villages and households in the RDP village are absorbing members.

This picture is complicated yet further when we note the considerable flux that is evident in the household formation and dissolution rates shown in Table 1. Indeed some of the turnover rates, particularly in the refugee villages and the RDP village, are truly astonishing.

The lowest panel in Table 3 disaggregates the decomposition by household types as discussed in Section 4.3. The results are surprisingly clear cut:

- The overall dilution effect is almost entirely due to the rapid creation of small households by South African citizens in South African villages

- The negative replacement effect is completely due to the dissolution of large households in the refugee villages.

- The lack of a within household change effect is due to offsetting effects within South African and Mozambican households within the South African villages. The former were shedding individuals while the latter were absorbing new members.

17
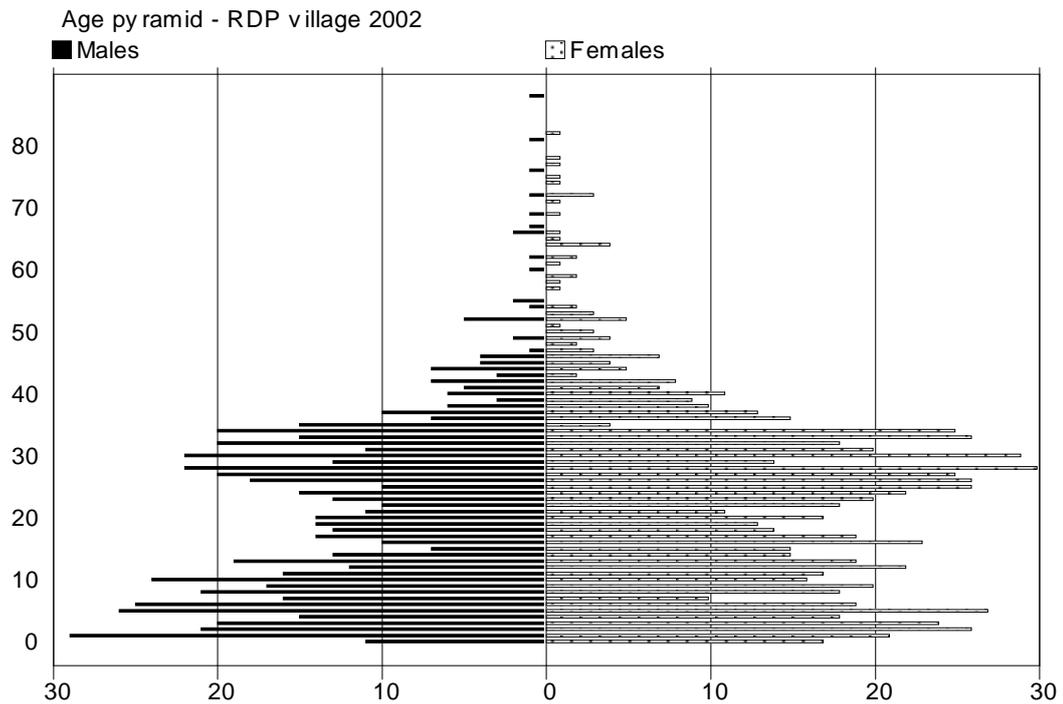
Figure 5: Age pyramid of the RDP village

6 Explaining the shifts

In the previous section we observed that the overall reduction in household size was fuelled by the rapid rate of household formation. As noted in the introduction, this is a rather remarkable finding, given that neither personal mortality, nor a reduction in fertility would have been expected to fuel household formation. Indeed economic conditions in the Agincourt area were not any better than nationally. Unfortunately we do not have direct economic information on these households except for a labour market snapshot that was taken in the year 2000. This module (added to the annual census round) shows high levels of unemployment, particularly in the resident rural population (Collinson and Wittenberg 2001).

6.1 Economic factors: land and services

While incomes might have militated against household formation there were, in fact, economic factors that would have made new household formation much easier.

In particular the costs of access to land and housing seem to have come down strongly. An extreme case is the RDP village within the Agincourt area, where houses were essentially allocated for free through a list system, i.e. a form of rationing. We can get some idea of what sort of individuals have taken possession of these houses by looking at the age structure revealed in Figure 5. It is clear that the houses are occupied by younger children (up to age 12) and adults in their twenties and early thirties. The age pyramid might suggest a settlement of mainly "nuclear" households, but the situation is more complex as is shown in Table 4. This table gives a crude classification of household types for the years 1999 to 2002. 1999 is the year in which residents first took possession of the RDP houses and by 2002 the RDP village was fully settled.

It is evident that compared to newly formed households in other parts of the study site and compared to existing households, there were many more single person households in the RDP houses. Indeed this is true in every one of the years in the table. Furthermore the proportion of "nuclear" households is **lower** among the newly formed households in the RDP village than it is among newly formed households in the rest of the study site.

Anecdotal evidence suggests that some of these "households" might really be seen as subsidiaries of bigger households existing elsewhere in the site. Some families seemed to be putting some of their younger members into the RDP houses as a way of establishing title to an asset that the government was providing free of charge. This raises all the questions about the nature of households introduced above (Russell 2003a). At one extreme one might therefore suppose that these are all "sham" entities, i.e. that within the family there may have been a change in living arrangements, but no substantive change in the social relationships. At the other extreme one could suppose that the external opportunity provided by the government has released some pent-up demand for privacy, which has led to the fissioning of some existing households. The truth is likely to be somewhere between: with some of these "households" more on the independent part of the continuum and others more on the subsidiary one. Undoubtedly there will also be many households somewhere in between, i.e. where the change in living arrangements *does* imply a reconstitution of existing social relationships, without these necessarily being severed, however. Indeed, it is interesting to note that many of the single person households that were established in 1999 must have been joined by partners by 2000 (they were the "existing" households in that year). So although the RDP village is special, it probably exemplifies some of the processes occurring elsewhere.

Indeed there have been other innovations in the local housing market. Collinson, Garenne, Tollman, Kahn and Mokoena (2000), for instance, document the movement of individuals to the adjoining area of Mkuhlu. This shift was enabled by the

breakdown of the "traditional" controls on the development of land. Given that Mkhuhlu had better access to employment, this led to significant local migration. Even within the Agincourt site the power to allocate land has shifted away from the chiefs and headmen to development committees. One of the constraints on new household formation has thereby become loosened.

The reduction in size of the refugee villages can most readily be explained in terms of onward migration to destinations that have better access to services and jobs. Some of the exodus would undoubtedly have been to Gauteng and other areas where job opportunities are concentrated. A move to one of the "South African" villages might, however, also be part of a household strategy to improve access to services. Indeed Cross and Harwin (2000) have argued that there is extensive migration **within** South Africa's rural areas and that much of this can be explained in terms of improving access to publicly provided infrastructure. The migration to Mkhuhlu referrred to above (Collinson et al. 2000) is another example of this strategy.

The broad trends summarised above can all be fitted into a set of economic explanations in which access to land, services and jobs feature prominently. Within the class of these accounts there are two broad competing explanations. It is clear that apartheid artificially reduced the supply of land and services to the majority of the population. It is therefore possible that the rapid rate of household formation is simply due to the release of this pent-up demand. On the other hand, it is possible that certain new policy initiatives of the new government (such as the RDP housing schemes) may have themselves stimulated demand. Our information suggests that both of these may be true. The fact that the rapid rate of household formation predated the creation of the RDP village suggests that there were independent processes leading to the reduction of household size. The creation of the RDP village certainly helped this process along. It seems clear that some of the "household formation" processes around the RDP village were fairly distinctive. On the other hand, a comparison (in Table 4) between the newly formed households within the South African villages and those in the RDP housing scheme suggests that the processes were part of the same continuum.

The economic accounts draw attentiont to the fact that changes in the cost of resources are likely to also change behaviour. We would expect households to act in ways to take advantage of the opportunities that opened up to them with the political, social and economic changes that occurred since 1994. Changes in living arrangements and hence household size follow as a consequence.

6.2   Changing preferences

Of course people will only take advantage of cheaper land to move out, if they (in some sense) prefer to live separately to living with a larger household. More gen-

erally, we noted earlier that there is a debate among sociologists whether African families are becoming more "nuclear" (Ziehl 2001, Russell 2003b, Russell 2003a). The patterns of household formation and dissolution discussed above would certainly suggest that couples or other "minimal household units" (Ermisch and Overton 1985) are leaving larger households and setting up independently.

These patterns cannot reveal, however, whether these changes in living arrangements reflect real changes in the underlying social relationships. As Russell has argued, people are embedded in long-lasting social relationships. Taking a snapshot across these relationships is not guaranteed to reveal the full set of connections. People may be part of an extended family system, even though they spend many years of their life in what looks like a "nuclear" household.

These objections undoubtedly have considerable validity. It is possible that we are observing a moment in which households are reshaping themselves. For instance, it is possible that the "refugee" households are sending out small "scout parties" that try to establish themselves in new locations and that larger households may reconstitute themselves around them in due course. Indeed the strong "within household change effect" among Mozambican households in South African villages might hint at such a process. Nevertheless it is also possible to overplay this sort of objection. What makes our study site interesting is precisely that it allows us to track households over several years. Furthermore it is at the "rural" end of the continuum. Russell's objection makes most sense in the context of urban migrants that are analysed without taking due cognisance of their rural social relationships[3]. Our data set includes the urban migrants provided that they are still identified as household members by the rural household.

Furthermore there are good grounds for believing that rural households may have been under considerable internal social strain. In the late 1980s the Bushbuckridge area saw considerable political conflict which took the form *inter alia* of generational conflict (Niehaus 2001). The "youth" of the area was seen as rejecting many of the "traditional" values of their elders. Given this background one might have expected some changes in the living arrangements.

Besides the generational dimension, there may very well also be a gender dimension. In the old "bantustan" areas, women had no rights to land or housing except through men. With the establishment of democracy in 1994 that pressure on women to stay with a male partner or parent would have been reduced. At the same time the "development controls" implicit in the traditional authority system weakened. The combination of those two forces may also have led to changes in household living arrangements.

---

[3]Even in that context successive cross-sections should capture individuals at all stages of this process.

7　Conclusion

The empirical evidence from this paper confirms that average household size in South Africa has come down. It suggests that the reduction has been of the order of half a person over the period 1994 to 2004 for the rural areas. Our decomposition suggests that the main driver of the reduction in household size is the rapid rate of new household formation over this period. Looking at the disaggregated decompositions, it appears that there are two linked mechanisms operating in the Agincourt area:

- the provision of free housing (in the shape of the RDP village) induced strong household formation. Some of this may have been "bogus" (households sending some members to stake claims to available infrastructure), but certainly not all of it.

- the reconstitution of households to gain better access to services, such as the move out of "refugee" villages

Changes in land rights (particularly of women) may also have played a role. Conflicts around traditional systems of control (of the older generation over the younger; and men over women) may have made setting up of new households more attractive. Arguably all of these forces were also operating nationally over this period. The roll-out of RDP housing, water, sanitation and electricity infrastructure in the late 1990s are likely to have fuelled new household formation. Apartheid was a system that was based on extensive location controls. The end of apartheid removed these. The extension of rights to women and other marginal groups, particularly in the rural areas, would have enabled these to create new living arrangements for themselves.

References

Amoateng, A. and Kalule-Sabiti, I.: 2008, Socio-economic correlates of the incidence of extended household living in south africa, *Southern African Journal of Demography* **11**(1), 75–102.

Börsch-Supan, A.: 1986, Household formation, housing prices, and public policy impacts, *Journal of Public Economics* **30**, 145–164.

Branson, N. and Wittenberg, M.: 2014, Reweighting south african national household survey data to create a consistent series over time: A cross-entropy estimation approach, *South African Journal of Economics* **82**(1), 19–38.

Budlender, D.: 2003, The debate about household headship, *Social Dynamics* **29**(2), 48–72.

Burch, T. K.: 1970, Some demographic determinants of average household size: An analytic approach, *Demography* **7**(1), 61–69.

Collinson, M., Garenne, M., Tollman, S., Kahn, K. and Mokoena, O.: 2000, Moving to Mkhuhlu: Emerging patterns of migration in the new South Africa. Working Paper, forthcoming as a chapter in the African Census Analysis Project monograph series volume 1.

Collinson, M., Tollman, S., Garenne, M. and Kahn, K.: 2001, Temporary female migration and labour force participation in rural South Africa. Working paper, Agincourt Health and Population Unit, University of the Witwatersrand.

Collinson, M. and Wittenberg, M.: 2001, Labour force dynamics in a rural part of South Africa: the Agincourt subdistrict of the Northern Province. Paper presented to the DPRU/FES conference on Labour Markets and Poverty in South Africa, Johannesburg, November 2001.

Collinson, M. et al.: 2002, Agincourt DSS, South Africa, *Populations and their Health in Developing Countries. Volume 1: Population, Health and Survival at INDEPTH Sites*, IDRC Press, Ottawa.

Cross, C. and Harwin, S. J.: 2000, A preliminary outline of migration trends in South Africa's central provinces, based on the October Household Survey 1995. Interim report to CIU Spatial Guidelines for Infrastructure Delivery Initiative. Available from http://www.gautengleg.gov.za/Publish/Public%20Relations/SPATIAL%20DEVELOPMENT(documents/cross%20migrationtrends.pdf.

Ermisch, J. F. and Overton, E.: 1985, Minimal household units: A new approach to the analysis of household formation, *Population Studies* **39**, 33–54.

Ermisch, J. and Salvo, P. D.: 1997, The economic determinants of young people's household formation, *Economica* **64**, 627–644.

Haurin, D. R., Hendershott, P. H. and Kim, D.: 1993, The impact of real rents and wages on household formation, *Review of Economics and Statistics* **75**(2), 284–293.

Kerr, A., Lam, D. and Wittenberg, M.: 2013, Post-Apartheid Labour Market Series 1994-2012 [dataset] version 2.1 of the harmonised dataset based on Statistics SA's OHS and LFS surveys, 1994-2012. available at http://www.datafirst.uct.ac.za/catalogue3/index.php/catalog/434.

Kerr, A. and Wittenberg, M.: 2013, Sampling methodology and field work changes in the October Household Surveys and Labour Force Surveys, *Technical Paper 21*, DataFirst, DataFirst, University of Cape Town.

Klasen, S. and Woolard, I.: 2009, Surviving unemployment without state support: unemployment and household formation in south africa, *Journal of African economies* **18**(1), 1–51.

Machemedze, T., Kerr, A. and Wittenberg, M.: 2014, Recalibrating the OHSs to adjust for sampling changes. DataFirst working paper.

Moultrie, T. and McGrath, N.: 2007, Teenage fertility rates falling in south africa, *South African Medical Journal* **97**(6), 442–443.

Niehaus, I.: 2001, *Witchcraft, power and politics: Exploring the occult in the South African Lowveld*, Pluto Press, London. with Eliazaar Mohlala and Kally Shokane.

Posel, D., Fairburn, J. A. and Lund, F.: 2006, Labour migration and households: A reconsideration of the effects of the social pension on labour supply in south africa, *Economic Modelling* **23**, 836–853.

Russell, M.: 2003a, Are urban black families nuclear? A comparative study of black and white South African family norms, *Social Dynamics* **29**(2), 153–176.

Russell, M.: 2003b, Understanding Black households: The problem, *Social Dynamics* **29**(2), 5–47.

Tollman, S.: 1999, The Agincourt field site - evolution and current status, *South African Journal of Medicine* **89**(8), 855–857.

Tollman, S., Herbst, K., Garenne, M., Gear, J. and Kahn, K.: 1999, The Agincourt demographic and health study - site description, baseline findings and implications, *South African Journal of Medicine* **89**(8), 858–864.

Wittenberg, M. and Collinson, M.: 2007, Household transitions in rural South Africa, 1996-2003, *Scandinavian Journal of Public Health* **35**(Suppl.69), 130–137.

Ziehl, S.: 2001, Documenting changing family patterns in South Africa: Are census data of any value?, *African Sociological Review* **5**(2), 36–62.

Table 1: Aggregate changes in Agincourt by type of household

Year	Mean household size				Number of households				year on year change				new households				terminating households			
	SA	Moz	Refugee	RDP	SA	Moz	Refugee	RDP	SA	Moz	Refugee	RDP	SA	Moz	Refugee	RDP	SA	Moz	Refugee	RDP
1992	6.36	7.29	6.68		6 807	1 663	1 354						225	115	78		101	93	104	
1993	6.39	7.43	6.73		6 931	1 685	1 328		124	22	-26		302	145	106		142	89	149	
1994	6.29	7.41	6.57		7 091	1 741	1 285		160	56	-43		286	162	87		135	64	147	
1995	6.24	7.38	6.65		7 242	1 839	1 225		151	98	-60		258	140	55		143	69	111	
1996	6.18	7.29	6.68		7 357	1 910	1 169		115	71	-56		258	140	55		143	69	111	
1997	6.09	7.16	6.67		7 559	1 993	1 167		202	83	-2		363	170	97		161	87	99	
1998	5.99	6.98	6.47		7 702	2 091	1 157		143	98	-10		300	163	119		157	65	129	
1999	5.97	7.02	6.55	2.73	7 796	2 137	1 151	218	94	46	-6	218	276	127	64	218	182	81	70	
2000	5.96	7.05	6.64	2.81	7 748	2 140	1 122	436	-48	3	-29	218	307	128	68	249	355	125	97	31
2001	5.88	6.96	6.58	2.92	7 785	2 203	1 118	503	37	63	-4	67	260	134	76	140	223	71	80	73
2002	5.88	6.96	6.41	3.01	7 787	2 207	1 128	504	2	4	10	1	214	88	76	104	212	84	66	103
2003	5.90	6.94	6.35	3.10	7 805	2 225	1 119	510	18	18	-9	6	167	69	44	104	149	51	53	98

Year	Net household formation rate				change due to entry				change due to exit			
	SA	Moz	Refugee	RDP	SA	Moz	Refugee	RDP	SA	Moz	Refugee	RDP
1992												
1993	1.8%	1.3%	-1.9%		3.3%	6.9%	5.8%		1.5%	5.6%	7.7%	
1994	2.3%	3.3%	-3.2%		4.4%	8.6%	8.0%		2.0%	5.3%	11.2%	
1995	2.1%	5.6%	-4.7%		4.0%	9.3%	6.8%		1.9%	3.7%	11.4%	
1996	1.6%	3.9%	-4.6%		3.6%	7.6%	4.5%		2.0%	3.8%	9.1%	
1997	2.7%	4.3%	-0.2%		4.9%	8.9%	8.3%		2.2%	4.6%	8.5%	
1998	1.9%	4.9%	-0.9%		4.0%	8.2%	10.2%		2.1%	3.3%	11.1%	
1999	1.2%	2.2%	-0.5%		3.6%	6.1%	5.5%		2.4%	3.9%	6.1%	
2000	-0.6%	0.1%	-2.5%	100.0%	3.9%	6.0%	5.9%	114.2%	4.6%	5.8%	8.4%	14.2%
2001	0.5%	2.9%	-0.4%	15.4%	3.4%	6.3%	6.8%	32.1%	2.9%	3.3%	7.1%	16.7%
2002	0.0%	0.2%	0.9%	0.2%	2.7%	4.0%	6.8%	20.7%	2.7%	3.8%	5.9%	20.5%
2003	0.2%	0.8%	-0.8%	1.2%	2.1%	3.1%	3.9%	20.6%	1.9%	2.3%	4.7%	19.4%

Note: Households have been classified by village type (SA villages, Refugee Villages, RDP village) and citizenship of head of household (SA villages)

Table 2: A numerical decomposition of the changes

	change	Δy^S	$y^N\text{-}y^T$	$y^N\text{-}y^S$	decomposition: within	replace	dilution
1993-94	-0.097	-0.015	-0.982	-2.784	-0.014	-0.038	-0.046
1994-95	-0.026	0.043	-0.689	-2.438	0.041	-0.024	-0.043
1995-96	-0.055	-0.014	-0.32	-2.563	-0.014	-0.010	-0.031
1996-97	-0.086	0.02	-0.907	-2.924	0.019	-0.030	-0.075
1997-98	-0.126	-0.026	-1.3	-2.834	-0.025	-0.043	-0.058
1998-99	-0.064	0.046	-0.535	-3.029	0.044	-0.016	-0.092
1999-00	-0.052	0.008	-0.5	-2.741	0.008	-0.027	-0.033
2000-01	-0.085	-0.036	-0.34	-2.748	-0.034	-0.013	-0.037
2001-02	-0.011	-0.01	0.052	-2.715	-0.010	0.002	-0.004
2002-03	0.003	0.002	0.264	-2.583	0.002	0.008	-0.007
1993-03	-0.599				0.018	-0.191	-0.424
% change	-9.1%			Contribution:	-3.0%	31.8%	70.9%

Table 3: Decomposing the change in household size by type of household

	South Africans in SA villages	Mozambicans in S.A. villages	Refugee villages	RDP vilage	Total
Average annual Δy^S	-0.016	0.057	0.011	0.122	
Average annual y^N-y^T	-0.045	-0.596	-0.981	0.150	
Average annual y^N-y^S	-2.744	-3.012	-2.541	-0.289	
Within category					
within effect	-0.152	0.543	0.094	0.410	
replacement effect	-0.014	-0.221	-0.880	2.858	
dilution effect	-0.330	-0.811	0.405	-0.172	
Contribution overall					
weighted within effect	-0.105	0.098	0.013	0.012	0.018
weighted replacement effect	0.007	0.010	-0.285	0.078	-0.190
weighted dilution effect	-0.410	-0.043	-0.017	0.046	-0.424
Notes: The averages in the case of the RDP village are calculated for the period 1999-2002					

Table 4: Household composition by type of household, 1999-2002

		South Africans in SA villages		Mozambicans in S.A. villages		Refugee villages		RDP vilage	
		New	Existing	New	Existing	New	Existing	New	Existing
1999	Single person	0.293	0.073	0.197	0.048	0.094	0.057	0.367	
	Couple	0.069	0.030	0.039	0.026	0.047	0.026	0.032	
	Nuclear	0.257	0.224	0.331	0.281	0.344	0.258	0.216	
	Single parent	0.159	0.083	0.205	0.071	0.250	0.091	0.174	
	Three generation	0.072	0.223	0.047	0.138	0.109	0.164	0.060	
	Other	0.149	0.367	0.181	0.436	0.156	0.405	0.151	
2000	Single person	0.192	0.070	0.133	0.048	0.132	0.049	0.337	0.246
	Couple	0.081	0.026	0.039	0.021	0.044	0.028	0.068	0.048
	Nuclear	0.342	0.213	0.414	0.265	0.294	0.236	0.249	0.257
	Single parent	0.166	0.077	0.133	0.081	0.250	0.088	0.153	0.187
	Three generation	0.062	0.228	0.094	0.145	0.044	0.171	0.068	0.070
	Other	0.156	0.385	0.187	0.440	0.235	0.428	0.124	0.193
2001	Single person	0.246	0.068	0.052	0.047	0.105	0.043	0.293	0.218
	Couple	0.085	0.026	0.075	0.018	0.013	0.025	0.050	0.072
	Nuclear	0.288	0.206	0.358	0.258	0.382	0.247	0.193	0.278
	Single parent	0.188	0.077	0.209	0.085	0.145	0.095	0.193	0.165
	Three generation	0.073	0.231	0.149	0.156	0.184	0.179	0.071	0.072
	Other	0.119	0.392	0.157	0.436	0.171	0.411	0.200	0.196
2002	Single person	0.271	0.069	0.148	0.047	0.158	0.049	0.308	0.173
	Couple	0.037	0.027	0.045	0.017	0.053	0.027	0.135	0.058
	Nuclear	0.285	0.203	0.295	0.244	0.316	0.243	0.202	0.275
	Single parent	0.182	0.078	0.182	0.093	0.211	0.092	0.183	0.180
	Three generation	0.070	0.233	0.102	0.166	0.092	0.193	0.048	0.090
	Other	0.154	0.391	0.227	0.432	0.171	0.395	0.125	0.225

About DataFirst

DataFirst is a data service dedicated to making South African and other African survey and administrative microdata available to researchers and policy analysts.

We promote high quality research by providing the essential research infrastructure for discovering and accessing data and by developing skills among prospective users, particularly in South Africa.

We undertake research on the quality and usability of national data and encourage data usage and data sharing.

DataFirst