Household responses to adverse income shocks: Pensioner out-migration and mortality in South Africa

Vimal Ranchhod[1 2]

Working Paper Number 133

[1] Post doctoral fellow, SALDRU, University of Cape Town and Population Studies Center, University of Michigan.

[2] I acknowledge comments and contributions from David Lam, James Levinshohn, Charles Brown, Justin McCrary, Taryn Dinkelman and various seminar participants at the University of Michigan, the University of Cape Town, the ERSA labour economics workshop 2006, the PAA 2006 meetings and the African Econometric Society 2007 meetings. Financial assistance was provided by the Fogarty International Center (Grant: D43-TW000657) and the National Research Foundation of South Africa. All errors and omissions remain fully responsibility.

Household responses to adverse income shocks: Pensioner out-migration and mortality in South Africa

Vimal Ranchhod *†

July 6, 2009

Abstract

How do poor households respond to the cessation of cash transfers in developing countries? South Africa's generous social pension system results in most of the poor elderly being the primary 'breadwinner' in the household. I estimate the magnitude of the changes in household composition and labour force activity amongst the resident members of the household, that correlate with a pensioner leaving the household. I use nationally representative matched panel data from several waves of the South African Labour Force Surveys. Compositional changes include the out-migration of school-aged children, and in-migration of middle aged females and older adults of either gender. More than 1 in 4 losing households get an additional older adult. For people who maintain their residency status across waves, I find large and statistically

*Post doctoral fellow, Saldru, Univ. of Cape Town and Population Studies Center, Univ. of Michigan

†I acknowledge comments and contributions from David Lam, James Levinsohn, Charles Brown, Justin McCrary, Taryn Dinkelman and various seminar participants at the Univ. of Michigan, the Univ. of Cape Town, the ERSA labour economics workshop 2006, the PAA 2006 meetings and the African Econometric Society 2007 meetings. Financial assistance was provided by the Fogarty International Center (Grant:D43-TW000657) and the National Research Foundation of South Africa. All errors and omissions remain fully my responsibility.

significant increases in employment rates for middle aged females and males (9.3 and 8.1 percentage points in each case), as well as for older adult females and males (10.3 percentage points in each case). For middle aged adults, this is not accompanied by a corresponding increase in labour supply.

1 Introduction

How do households respond to the cessation of cash transfers in developing countries? I estimate the magnitude of changes in household composition and household labour supply that occur when a pensioner leaves the household, either due to out-migration or death. The data used are nationally representative household level matched data from South African Labour Force Surveys (LFS) from September 2001 to March 2004.

The non-contributory South African Old Age Pension (OAP) forms the backbone of the South African social security system. Recipiency rates are high amongst the elderly, and over 77% of Africans who are age-eligible report receiving the pension. In addition, a means test ensures that the pension disproportionately reaches poorer households. Not only is coverage widespread, but its value is sufficiently high to generally make the pensioner the main breadwinner in their households. Case and Deaton (1998) note that in 1993, the value of the pension was "twice the median household's per capita income" amongst African households. Based on the September 2002 LFS, 19.28% of all households report 'pensions and grants' as their main source of income. Amongst households with a member who is old enough to be eligible, this percentage rises to 63.67% for all households, and 70.17% for African headed households.

Given the importance of the pension, investigating how households cope with its loss is of interest for at least two policy related reasons. First, it can inform us as to how relatively poor families act to mitigate against the effects of adverse economic developments. One dimension of this involves household responses in terms of household composition. If the group that constitutes a 'household' is itself endogenously determined, then careful consideration for

policies targeted at the household level is warranted. Second, sharing of pension income within households might lead to non-recipients deciding not to work. On the other hand, if poor households are liquidity constrained and the pension eases these constraints, we might observe that employment rates actually increase as a result of the pension. Estimating the magnitude of the net effect is the primary contribution of this paper.

To summarize my results, I find significant evidence that the pension does indeed affect both household composition and labour force participation of non-recipients. Households re-organize such that they have more adult time in the labour force, more non-pension aged adult residents who are not employed, and a decrease in the number of school-aged children. I find significant increases in employment amongst resident middle aged and older adults, and significant increases in labour force participation amongst resident older adults.

The remainder of this paper is structured as follows: Section 2 describes the value of the pension and some of its eligibility criteria. Section 3 discusses the related literature. Section 4 provides a theoretical framework within to analyze the question. Section 5 describes the data used and related issues. Section 6 discusses summary statistics from the data. In section 7, I discuss the empirical specification of the regression models I estimate. Sections 8 and 9 present the main findings at the household and individual levels respectively. Section 10 presents some caveats and robustness checks to the main results. Section 11 discusses the results relative to some of the related literature. Section 12 concludes.

2 Background

Lund (1993) provides an introduction to the OAP as we see it today. As stated previously, the pension is means tested, and provides a relatively generous cash transfer to recipients. Eligibility depends only on age, nationality and satisfying the means test. The age-eligibility threshold is 60 for women and 65 for men. The level of the means test is set fairly high, so that most of the elderly receive the grant. Moreover, it is based on individual income for the

unmarried elderly, or joint spousal income for married couples, and hence should not have distortionary 'implicit taxation' effects for other household members.

The value of the pension is adjusted periodically, usually on an annual basis, to adjust for inflation. In 2002 and 2003, the value of the pension was set at 620 and 700 rands per month respectively. Adjusting for consumer inflation[1], and converting rands to dollars using the prevailing exchange rate, this equates to between $100 and $125 per month. This is a large transfer relative to potential wage income, and continues for as long as the pensioner remains alive and continues to satisfy the means test.

3 Related Literature

Several researchers have investigated the effects of pension recipiency on various dimensions of household welfare. Case and Deaton (1998) find that the pension is an effective tool for redistribution and that the households it reaches are predominantly poor. Furthermore, the prevalence of three-generation households, as well as 'skip generation'[2] households, results in the pension disproportionately reaching children in poverty.

Some authors have looked at whether the OAP impacts on the health of recipients or their household members. Duflo (2000) finds a discontinuous increase in girls' height for age for children living with pension eligible persons. This increase is significant and is realized on average only when the pension recipient is a woman. Duflo (2003) reports similar evidence that the pension is shared between members of the household. Moreover, the sharing of recipients' income from pensions is differentiated by gender. Case (2001) finds that the health of all household members is improved as a result of the pension.

Others have asked the question; 'How do other members of the household respond when a member becomes pension eligible'? Bertrand *et al* (2003) analyse cross sectional data and

[1]The deflator used is the official Consumer Price Index released by Statistics South Africa.

[2]Households with grandparents and grandchildren but non-resident parents.

find that having a pension eligible person in the household has a statistically significant and negative impact on the labour supply of prime aged resident males in the household. Edmonds (2006) considers the impact of the OAP on child labour supply and schooling attendance. Ranchhod (2006) finds that the pension causes retirement amongst the recipients themselves. Posel *et al* (2006) find that the pension actually increases the labour supply of non-resident household members by financing labour migration towards areas with better employment prospects.

Jensen (2004) questions whether household disposable income increases by the full value of the pension. He estimates that crowding out of remittances by pensions is large and significant. On average, every rand of pension income received by the elderly is met with a 0.25 to 0.30 rand decrease in remittances received from the pensioner's children. Pension income is thus *de facto* shared with family members even when they do not reside with the pensioner.

Edmonds *et al* (2005) find that household composition itself is affected by someone becoming pension age-eligible. They find a decrease in the number of prime working-age women, and an increase in the number of children younger than five and young women of childbearing age. Hamoudi and Thomas (2005) go further, and find evidence that the OAP results in compositional changes consistent with sorting on the basis of unmeasured personal characteristics. This result poses a challenge to much of the prior research conducted using cross-sectional data.

Given that the pension seems to be so important in sustaining the poor and the elderly in South Africa, a natural question to consider is how do these households cope when the pension income stops. A paper which asks a similar question, *inter alia*, is presented by Ardington *et al* (2007). Using individual level longitudinal data from a poor rural area in the KwaZulu-Natal province, they investigate changes in labour supply amongst both migrant and resident household members as a function of the pension. In their dataset, after controlling for individual level fixed effects, they find that both the migration decision as well as labour supply are positively affected by pension receipt. Moreover, they find asymmetries

in the effects of the pension on household members who are already migrant labourers, as compared to those who are potential migrants.

In this paper, I answer the question: 'How does household composition and labour market activity of resident household members change at the same time that we observe the departure of the pensioner, either due to out-migration or death?' This question has not been thoroughly investigated in the literature to date. A second major contribution is that it is the first study in the OAP literature that does so using nationally representative and longitudinal data.

4 Theory

The most basic model of household formation assumes that households form for the production of some non-tradeable good in which there are economies of scale (Becker, 1973). In this paper, I assume that household composition and labour supply of household members are both endogenous outcomes to changes in non-labour income. Various authors have commented on the fact that inter-household migration occurs in response to the pension; see Keller (2004), Edmonds *et al* (2005) and Posel *et al* (2006).

Economic theory is fairly clear on the effect of a loss of outside income on labour supply in a non-credit constrained household. Assuming that leisure is a normal good, we would expect people to be more likely to work when the pensioner leaves the household. This could manifest in terms of home production or market based work.[3] In this context however, an increase in members' willingness to participate in market based work must depend on their time available to increase their work hours. For example, if all 30 year olds are already engaged in market related work, then we cannot observe an increase in their labour force participation (LFP). We would thus expect the response to be greatest amongst those groups who have time to work and whose wages are relatively high (amongst those household

[3]Data limitations preclude me from considering the effects on home production.

members not currently working).

The story becomes more complicated when one considers the role of the pensioner within the household. Suppose, for example, that the pensioner looked after household children. Then there exists the possibility that a resident adult has to leave the labour market to assist with child care. On the other hand, if the pensioner is ill and requires care within the household, his/her departure would free up some other member's time and possibly allow for greater labour market activities.[4]

Economic theories of the family and household formation are also unclear about what would happen to household composition. If the pensioner provided child care, we might expect non-resident family adults to take up residence in the household. On the other hand, the household may have to send adult members out of the household to become migrant labourers in other regions, which is consistent with the model by Rosenzweig and Stark (1989). In contrast, the model presented by Ardington *et al* (2007) incorporates liquidity constraints and assumes that migrant labourers initially need to draw resources from the original sending household. In this scenario, the loss of pension income would lead to a decrease in the out-migration of adults, and a return of some migrants into the household. A different compositional response could be for the household to send children to live with members of their kin network in other households.[5]

The prediction of the effects on household composition is thus by no means unambiguous. Ultimately, the question remains to be informed by empirical analysis.

[4]This latter scenario seems unlikely given that the amount of time reportedly spent by adults on care giving to the elderly is almost zero. See Ranchhod and Wittenberg (2007) for details.

[5]Fostering of African children is not uncommon in South Africa. See for example, Beittel (1992) and Sagner and Mtati (1999).

5 Data, Sample Selection, and Related Issues

5.1 Sample selection

The data I use comes from the South African Labour Force Surveys (LFS). These are nationally representative household level surveys that are conducted with a biannual frequency, in March and September of each year. They contain a complete household roster, demographic information such as age, race, gender and education for each respondent, and detailed information on labour force participation, employment, occupation, hours worked and earnings for all household residents aged 16 and above. In some waves there is also basic information about the quality of the household's physical dwelling structure, home ownership, and the relative importance of various forms of non-labour related earnings. The latter are household level variables.

I make use of data from wave 4 through wave 9 of the LFS in this paper (i.e. Sept 2001 - March 2004). Since most of the analysis I conducted is at the household level, I collapsed all the relevant information to the household level. Table 1 shows the initial sample sizes in the cross-sections, and the subsequent sample after each additional restriction discussed below is imposed. Initially, there are 163197 unique household numbers (within waves) across all of the waves combined.

The question on whether a person receives the pension or not is asked only of those who are not currently employed. Since the means test is relatively generous, a non-trivial proportion of the working elderly could also be receiving the pension. Moreover, the LFS is structured to classify a broad range of activities as 'employment', which exacerbates the problem.[6] I therefore decided to make use of the legal age requirements as proxies for pension income, which is consistent with what almost all researchers investigating the effects of the OAP have done. For this reason, I excluded all households which had any household members' age as

[6] In September waves, there is a household level module which asks, 'Does any person in this household receive an Old Age Pension?', but this is not present in the subsequent March wave.

unknown. I also focus exclusively on African headed households.[7] Africans comprise the majority of the population, are disproportionately poor, and conditional on age-eligibility, are highly likely to be receiving the pension. In all of the September waves combined (i.e. wave 4, 6 and 8), 88.5% of African headed households that included at least one pension-aged member reported that someone in the household receives the old age pension.

From waves 4 to 9, the LFS contained a 20% out-rotation component of dwellings.[8] Thus, theoretically at least, 80% of dwellings were revisited between any two six month periods. The essence of my analysis is to identify households that we observe in two subsequent waves of the LFSs, identify those which had a pensioner in the 'first' wave and 'lost' that pensioner by the next one, and measure the magnitude of other changes that occur in such households as well. In its most basic form, this is simply a 'before and after' comparison.

5.2 Measurement Error

Measurement Error arising due to false matches

One challenge to the analysis is measurement error arising due to the possibility of false matches. Since identification is based on the idea that a pensioner's departure from a household will lead to additional responses from the remaining household members, it is essential that I do, in fact, observe the same household in each of the two waves. However, this is a dwelling level panel, and is thus not necessarily the same household over time.

To minimize this potential problem, I included only those dwellings where at least one resident member was included in the Statistics South Africa[9] (StatsSA) individual level

[7]Technically, I included households in which the eldest member is an African. Given that the eldest member is generally the household head, and the infrequency with which multi-racial households are observed in the data, this captures the race of the household head accurately in almost all cases.

[8]Source: Statistics South Africa Labour Force Survey metadata documents.

[9]Statistics South Africa is the official national statistics organization, and is responsible for conducting the LFS.

panel, and has the same race, gender, and similar ages in wave t and $t+1$.[10] StatsSA recently invested considerable resources to extract an individual level panel from this rotating dwelling level panel. The match quality is likely to be good, since they use the dwelling identifier information, demographic characteristics, as well as the confidential first and last name of the respondent to identify the person level matches.

On the other hand, the individual level match rate was not quite complete, even after accounting for migration and the rotation pattern.[11] McLaren (2007) reports that between 46.67% and 66.40% of individuals in the cross-sections survive into the matched panel. This is considerably less than the theoretical upper limit of 80%. The author presents evidence that a major reason for the attrition is due to household non-response in a subsequent wave. This is supported by the observation that the distribution of match rates within households is strongly bimodal, with large spikes at zero and one, and a relatively sparse density in between. In addition to non-response of households, is also possible that the dwelling that was 'revisited' was in fact a different physical dwelling to the original one surveyed in the prior wave. This is likely to be more of a problem in regions where addresses are not well captured, such as in some rural areas and shanty towns.

I excluded dwellings where there was more than one household on the property, since this represented a greater risk of false matches than single household properties.[12] The impact of each of these criteria on the sample size is shown in Table 2. Of African headed households living in single dwelling properties, I match about 76% of households across waves, which is close to the 80% I would expect from the rotation pattern.[13] Restricting the sample further to those households with at least one 'good' individual level match reduces the sample size by roughly one third, from 68,413 to 45,582 households.[14] Finally, I included only households

[10] By 'similar' age I required that the $0 \leq age_{t+1} - age_t \leq 1$.

[11] Approximately 5% of residents in waves 5 - 9 reported that they were not living there 6 months ago.

[12] StatsSA indicated that the 'hhid' were maintained by property across waves, but not necessarily for dwellings within properties. Of the households that satisfied every other requirement for inclusion in the sample, this excluded 6.25% of households.

[13] Some of the remaining slippage is likely to occur if a property experiences a change in the number of dwellings between waves.

[14] This is slightly better than the individual match rate, and occurs since I only require one individual in

with at least one person who was pension age-eligible in wave t.

I am thus able to identify 12,342 households that had a pension-aged member in wave t, where the 'pensioner'[15] is absent in wave $t+1$ for some subset of these households.

There are also 444 households that show a net increase in the number of resident pensioners by the subsequent wave. These I excluded from the analysis for various reasons. First, the substantive question is about how households cope with the loss of the pensioner and the related cash transfer. The question of the impact of pension receipt has been widely studied already, albeit using cross-sectional data. Second, the date of pension eligibility should be fully anticipated, and so only liquidity constrained households are likely to respond. In contrast, the date of the departure of the pensioner, particularly if due to death, is a stochastic variable. Third, it is difficult to separate between actual aging into the pension and age misreporting, which is discussed below. Finally, the sample size is much smaller, which results in limited statistical power.

Measurement Error arising due to 'Age Heaping'

Of the remaining 11,898 households, there are 10266 'Keeper' households which have no change in the net number of pensioners between wave t and wave $t+1$. There are also 1632 'Loser' households, which show a decrease in the net number of pensioners between waves.

A potentially serious measurement error problem arises if people report age imperfectly. Figure 1 shows the 'Age Heaping' phenomenon, whereby people tend to round ages to focal points of multiples of 5 or 10. This results in spikes in the observed age distribution in the data, and is particularly pronounced amongst the elderly. Suppose that a person's reported age in an initial wave is such that I would classify them as pension age-eligible. If, in a subsequent wave, a person's age is then reported as below the pensionable age, I would end up

the household to be matched for the household to remain in the sample.

[15]For the remainder of the paper, I use the word 'pensioner' to refer to a person who is age-eligible to receive the old age pension.

classifying the household as a Loser. In reality, however, the economic environment of such a household has remained the same. This is also not of the classical measurement error form, which necessarily results in attenuation bias, since the composition would simultaneously reflect an increase in the number of older adults who are not yet pensioners.

In response, I only included a subset of Loser households where I can be reasonably confident that a pension-aged individual did indeed leave the residence. I classified a household as a 'strict Loser' if:
(# residents aged $\geq$ ((pension-age)+2)) in $Wave_t$ > (# residents aged $\geq$ pension-age) in $Wave_{t+1}$

Households with someone reported as having 'age $\geq$ ((pension-age)+2)' are unlikely to include households that are not actually pensioner households but may be incorrectly classified due to the age-heaping. On the other hand, households that are actually Keeper households with a pensioner aged $\geq$ ((pension-age)+2)) are unlikely, even with age-heaping, to report the person's age as being strictly less than the pensionable age in a subsequent wave.[16]

All subsequent analysis in this paper includes only the 1220 'strict Loser' and 10266 'Keeper' households.[17] The final sample, then, has 11,486 households, observed once in the 'before' period ($Wave_t$) and once in the 'after' period ($Wave_{t+1}$) each. Unless specified otherwise, the unit of observation is thus a household from a particular panel. Some households are included more than once if they meet all the criteria and appear in more than one panel.

5.3 Selection Correction on Observables

The fact that the match rate was relatively poor introduces the possibility of selection bias. In order for the analysis to be a valid description of what happens to pensioner households

[16] I did not impose a similar restriction on Keeper households. Preliminary analysis revealed that Keeper households are incredibly stable, such that including only 'strict Keeper' households would only limit the sample size but not impact on the magnitude of the estimated coefficients.

[17] For the remainder of this paper, the term 'Loser' implies a 'strict Loser' household.

on average, I need to assume that the households that are included are representative of pensioner households in general. This assumption is unlikely to be true.

The first three columns in Table 3 show how the households that meet all the other criteria[18] but did not feature in the panel compare to those that were included in the panel.[19] I compare these two groups for a host of composition and activity variables. For most variables, t-tests for differences in the means reject the null hypothesis that those included and those that attrited were drawn from the same underlying population.

To the extent that such attrition arises for observable reasons, we can correct for this by reweighting our matched sub-sample. For example, if shack dwellers are more likely to move and are thus less likely to be matched, we can adjust the weighting of those shack dwellers who we do manage to match. Thus, non-random matching on observables is not an insurmountable problem per se, as we can use the 'inverse probability weighting' (IPW) method to obtain unbiased estimates.(see Wooldridge 2002, pp 587-590).

I estimated probit models and reweighed the panel sample using the IPW method. The probits were estimated separately for each wave. The probits were estimated only using those observations for which I had a corresponding household that I could potentially match to. This is not too problematic, since the objective is purely a statistical rebalancing one - I want the group in the panel to look more like that from the full cross-section.

The variables I included were original household composition and location variables, employment data of various demographic groups within the household, and information about the ownership and characteristics of the physical home. Since the 'wall type', 'home ownership' and 'dwelling type' questions are only asked in the September waves, for waves 5 and 7 I used the information from the matched household in the subsequent wave.

[18] That is, they had a pensioner, the eldest person was African, there were no observations with age missing, and there was only one dwelling on the property.

[19] All pensioner households from the panel were included in this process, including 'Gainers' and Losers who were not 'strict Losers'.

The regression results suggest that the panel over-represents larger households, urban households, as well as households whose residents owned their home, none of which is surprising.[20] I then predicted the probability of inclusion in the panel, and all results in the analysis are weighted by the inverse of this probability, multiplied by the relevant sampling weights.

At a glance, columns III to VI in Table 3 suggest that the process was reasonably successful at achieving its objective of re-balancing the panel to look like the cross-section.

5.4 Non-random Selection on Unobservables

A more difficult potential problem occurs if we have non-random matching based on unobservable characteristics, which persists even after the selection correction on observable characteristics. If these characteristics are orthogonal to the variables we are interested in, the estimates will still be unbiased in expectation. If, however, an entire household migrates in search of better economic opportunities upon the death of a pensioner, I cannot control or adjust for this. I thus need to qualify my findings to those Loser households where at least one member stays in the same residence.[21]

6 Summary statistics

6.1 Dependent Variables

Table 4 presents mean household characteristics for Keeper and Loser households in the initial $Wave_t$ period, for each of the major dependent variables that I consider. Columns 1 and 2 shows the mean household composition for each age group. The age classification was somewhat arbitrary, with 'kids young' being aged 7 or lower, 'kids school' aged 8 - 15,

[20] These regression results are omitted for brevity, but are available from the author upon request.

[21] It is impossible to determine whether this is a large or small problem in this context.

‘youth’ aged 16 - 20, ‘young adults’ aged 21 - 35, ‘middle adults’ aged 36 - 50, and ‘older adults’ ages 51 - 59 if female, and 51 - 64 if male.[22]

Columns 3 and 4, 5 and 6, and 7 and 8 show the mean proportion of people within these age groups that are ‘working’, ‘in the labour force using the broad definition’, and ‘in the labour force using the narrow definition’ respectively. The ‘broad’ category includes anyone currently employed or willing to work. The ‘narrow’ category includes the employed, and only those unemployed who are willing to work and who have been actively searching for employment in the past month.

To begin with, Loser households are considerably larger, by more than 1 person on average. The differences are most pronounced for the younger age groups, up to and including the young adults. For most groups, the proportion in each of the LFP categories are somewhat similar. The exceptional category is the middle aged adult males in the soon to be Loser households, who are about 10 percentage points more likely to be in the labour force using either definition, and 8 percentage points more likely to be employed. Also of note is that the employment rate in each age group is relatively low, for both Keepers and Losers. Fewer than 1 in 5 young adults, and 1 in 3 middle and older adults are working, in both Keeper and Loser households.

6.2 Income Sources

Non-labour income is captured rather crudely via a question that reads “What is the main source of income for this household?”. One possible response is “remittances”. There is no information on the value of these remittances. Moreover, the question is only asked in waves 4, 6, and 8 (i.e. the September waves). I calculate the distribution of the responses for the Keeper and Loser households in the relevant panels. In order for this comparison to be

[22] At age 7, children should legally be enrolled at school, but enrollment become almost universal by age 8 only. Similarly, 16 is the legal age at which a person may drop out of school or enter employment, while at 21, a person becomes a legal adult.

valid, one needs to believe that Loser (and Keeper) households in waves 5 and 7, were similar to Loser (and Keeper) households in waves 4, 6 and 8.[23] For these reasons, the summary statistics presented in Table 5 are only suggestive.

The Keeper households remain fairly stable, which lends credibility to the aforementioned assumption. About 78% of Keeper households in both time periods report 'Pensions and Grants' as their main income source.[24] Loser households look different from Keepers even in the period prior to their loss. Losers in $Wave_t$ are more likely to report 'Salaries and Wages' as their main income source (26.3% vs. 15.6%). This would be expected if people are anticipating the imminent departure of the pensioner.

In $Wave_{t+1}$, this distribution changes remarkably in Loser households. Pensions and grants decreases as the main income source from 61.7% in the 'before' period, to 35.0% in the 'after' period. This is accounted for mostly by a large increase in the proportion that report remittances as their main income source, which increases from 7.3% to 26.0%. Almost 1 in 5 Loser households experience this transition. Somewhat smaller changes are observed in the fraction of Losers that report salaries and wages as their main income, which increases from 26.3% to 31.0%. Given this, and previous research by Jensen (2003), it seems plausible that there might be offsetting increases in remittances in Loser households.

Changes in both household composition and labour force participation amongst residents may be muted if there is an offsetting increase in remittances to the household to compensate for the loss of pension income. The observed changes in the distribution of main income sources lends credibility to the subsequent analysis and interpretation; that Loser households are indeed experiencing changes that correlate with the loss of pension income.

[23] i.e. Losers at time T_0 in waves 5 and 7 were similar to Losers at time T_0 in waves 4, 6 and 8.

[24] These are simply weighted means. South Africa also has several other grants, the most common being the Child Support Grant.

7 Empirical Specification

Next, I employ multivariate regression techniques to control for additional factors, and test for the statistical significance of changes in household composition and labour force participation. I regress the difference in the 'dependent variable' for households between $Wave_t$ and $Wave_{t+1}$, on an indicator for whether the household was a Keeper or a Loser. The regression that I fit is of the form:

$D_{j,t+1}$ - $D_{j,t} = \beta_0 + \beta_1 losepen_{j,t} + \beta_2 X_{j,t} + \epsilon_{j,t}$
where j, t denotes a household j in $Wave_t$, and D is the dependent variable of interest.

'$losepen_{j,t}$' is an indicator variable that equals one if household j is a strict Loser between $Wave_t$ and $Wave_{t+1}$, and 0 if household j is a Keeper between $Wave_t$ and $Wave_{t+1}$. Additional X variables include an indicator variable for urban areas, provincial dummy variables, wave dummies, household size [25] and a count variable for the number of pensioners in the household in the initial period. I include this last one since losing one of two pensioners potentially has smaller effects than losing the only pensioner in the household.[26]

This specification nets out any unobservable but time invariant characteristics that are specific to a particular household. Moreover, by comparing the change in Loser households relative to Keeper households, I also expect to net out any effects that arise due to the aging of the underlying population, as well as changes in economic conditions that equally affects members of both groups of households. To correct for the non-independence of households across different waves of the panel, I estimate robust standard errors which are clustered at

[25] Household size was not included in the regression where the dependent variable was the difference in household size itself.

[26] If the effect of a loss of a pensioner does indeed differ depending on the number of pensioners to begin with, then the coefficient estimates capture an average effect across the groups with different numbers of initial pensioners. I also estimated the results for the subsample of households with only 1 pensioner in $Wave_t$. Results are similar in magnitude and significance for most of the groups and outcome measures. However, the sample size decreases by more than 2000 observations, from 11486 to 9304. I thus chose not to limit the sample further.

the household level.

8 Changes in composition and aggregate household labour force participation

The coefficient on the *losepen* variable is presented for each of the dependent variables in Table 6. Each coefficient and corresponding standard error is obtained from a separate regression. The coefficients measure the difference in the mean changes in the dependent variables between Keepers and Losers, after controlling for all the other X variables. As stated previously, exploratory analysis revealed that the Keeper households are incredibly stable, so the coefficients reported are identified primarily using variation within the Loser households.

I observe large and significant changes in household composition. Not surprisingly, aggregate household size in Loser households goes down. There is a reduction in the number of school aged children in the household of 0.063, and an inflow of 0.052 middle aged females. The largest change in composition occurs amongst the older adults, with a net increase of 0.279. Thus, more than 1 in 4 losing households get an additional older adult on average. This in-migration is approximately equally comprised of men and women.

In terms of the numbers employed, almost all the groups considered experience a significant increase in the number of employed persons in that group.[27] The largest of these is experienced by the older adults, with a coefficient of 0.102. These are comprised equally of men and women. The increase in number employed is less than the increase in the number of residents for the adult groups that show a significant change in composition. Thus, there is also a considerable amount of more adult time for home production activities in Loser households.

[27] The exception, once again, are the middle aged men.

With respect to the number in the labour force, only for the older adults is the increase significant, with a coefficient of 0.118 for both genders combined. This is also the only group for whom more people enter the labour force than find jobs.

An alternative and potentially more intuitive way to interpret these findings is to consider the proportionate changes in the mean levels of the various outcomes that these regression coefficients suggest. The numbers presented in Table 7 reflect this.[28] I restrict commentary to the numbers that correspond to significant coefficients from Table 6. With regard to compositional changes, the household size decreases by 13.5%, and the number of children of school going age decreases by 4.8%. There is a 10.9% increase in middle aged adults, with middle aged females increasing by 15.1%. The largest changes are observed amongst the older adult group. For both genders combined as well as for each gender considered separately we observe an increase of greater than 100% in this age group.

Turning now to the proportion that are employed, there is an increase in the proportion employed for all the adult groups considered of either gender. The change is largest amongst middle aged adults, with the proportion of employed females increasing by 5 percentage points. For young adults, the change is 3.9 percentage points, while the changes are smallest for the older adults, at 2.5 percentage points. In considering how the household copes with the loss, there are changes in the number of people available for work, the probability that a person in a particular group will be employed and the total number of additional workers in the household.

The only significant coefficients from Table 6 with regard to being in the labour force relates to the older adults. Using the broad definition, the corresponding mean changes in proportion

[28]It is undesirable to use the proportion in a particular labour market category in a particular household as a dependent variable directly in a regression. Any household with no person in a particular age group will have an undefined proportion in that labour market category. Moreover, a majority of compositional changes are from 0 to 1 or vice versa. The difference in these proportions will thus be undefined for every household with no members in an age group in **either** wave. This would result in the most dynamic households, in terms of composition, being dropped from the estimation sample. Hence, we report the proportionate changes in the mean, rather than the regression adjusted mean of the proportionate changes.

in Table 7 are -4.7, -2 and -6.9 percentage points for the older adults, the older female adults and older male adults respectively. The corresponding numbers using the narrow definition are 1.4, 2.9 and 0 percentage points. The point estimates from the regressions in Table 6 are almost identical for each group regardless of the definition used, hence the corresponding difference in magnitudes and signs in Table 7 arise almost entirely due to differences in the initial proportions in these demographic groups in the two labour market categories.

In sum, I find significant evidence that the household re-organizes itself in conjunction with the departure of a pensioner. Aggregate household level labour supply when measured in number of workers increases primarily amongst the older adults, although the number of employed adults increases for most categories. In proportionate terms, however, the increase is largest amongst the middle aged adults in general and the middle aged females in particular. That said, the change in composition makes it difficult to identify an important policy concern. Are the changes in labour supply and employment arising due to a behavioural change from already resident household members, or are they the result of the in-migration of people who were already in the labour force or employed, and simply maintained their status? I partially inform this question in the next section.

9 Changes in labour force participation conditional on maintained individual residency

Table 8 presents individual level regression results for the same set of dependent variables using observations from the matched StatsSA panel. The estimation sample is the set of matched individuals in the Loser and Keeper households already identified. The exact specification is:

$D_{i,g,j,t+1}$ - $D_{i,g,j,t} = \beta_0 + \beta_1 losepen_{j,t} + \beta_2 X_{j,t} + \epsilon_{i,g,j,t}$

where i, g, j, t denotes individual i in demographic group g in household j in $Wave_t$, and D is the dependent variable of interest. Additional X variables include an indicator variable

for urban areas, provincial dummy variables, wave dummies and the number of pensioners in the initial $Wave_t$ period. Group status was determined by age in the $Wave_t$ period.

Youth and young adults generally show small and statistically insignificant effects on labour supply. However, there is a 5.7 percentage point decrease in the probability that youth are in the labour force using the broad definition, and a 7.4 percentage point decrease in the probability that young female adults are actively searching for employment. This might be reflecting a need for younger adults to take up some child care roles.

Amongst middle aged adults, there are large and significant employment effects, but insignificant, smaller and negative effects for labour supply. Middle aged females and males in Loser households are 9.3 and 8.1 percentage points more likely to be employed respectively, relative to the comparison group in Keeper households.

Again, the largest results are obtained for the group of older adults. Older adults of either gender are 10.3 percentage points more likely to be employed, and the increase is statistically significant at the 5% level for the group with both genders combined. These are approximately matched by increases in labour supply, which are also large in magnitude. The increases are larger for women than for men, 12.2 percentage points vs. 10.6 percentage points using the broad definition, and 13.1 percentage points vs. 8.3 percentage points using the narrow definition. The coefficients for women are also statistically significant at the 5% level.

In summation then, for these selected observations, I find significant and large increases in employment rates for middle aged and older adult groups. This is accompanied by increases in labour supply for the older adults, who tend to have relatively low LFP rates to begin with. For the middle aged adults, there is no corresponding increase in labour supply on average. These employment effects are especially striking given that South Africa's unemployment rate is about 30% amongst adults, that most of the unemployed seem to experience chronic long term unemployment[29] and that this occurs within the relatively short space of time

[29]See Kingdon and Knight (2002), Banerjee *et al* (2006)

between waves.

10 Caveats and Robustness checks

The coefficients presented are not necessarily causal estimates. For example, the out-migration of a pensioner might be a consequence of the in-migration of other household members, or a change in other members' employment status. Alternatively, there may be other factors that simultaneously change the pensioner's choice of residence as well as that of other household members. There are other limitations, partly due to the data available. I cannot observe why the pensioner left or where she went to. I also have no information regarding where the new household members came from, where the out-migrating children go to, nor the activities of any non-resident members. There is also only limited information about remittances and resource sharing within families but across physical households. A complete analysis would be able to observe all of these in order to gain a full understanding of the effects of the pensioner's departure.

There is one case, however, where the departure of the pensioner is plausibly exogenous, namely the death of the pensioner. This is still not a panacea, for the family may anticipate the death of the pensioner and start rearranging the family prior to his death. In this case, I would be biased away from finding any results, which implies that the 'death' estimates are biased towards zero, and should thus be interpreted as lower bounds of the true effect.

10.1 Identification using deaths

Since one might be concerned about endogenous out-migration of the pensioner, I further restricted the sample to include only Losers who experience the plausibly exogenous event

of the death of the pensioner.[30] In the wave 5 (March 2003) module, respondents were asked about recent deaths in the household. I use this data to generate an indicator variable for whether an elderly member died recently in a Loser household.[31] This variable is called '$Death1$', which includes 47 Loser observations.

I also infer deaths indirectly and probabilistically, using a combination of marital status and spousal identifiers. To do this, I used the 'good' individual level matches from the StatsSA panel, and identified who was married to a pensioner. This is only possible for the subset that were married in $Wave_t$ and lived with their spouse at the time. I then infer death by identifying those who transitioned to become a widow or widower in $Wave_{t+1}$. I classify the variable '$Death2$'=1 if the above criteria are satisfied in a Loser household. The '$Death2$' sub-sample has 60 Loser observations.

Regression results for the $Death1$ and $Death2$ samples are presented in Tables 9 and 10 respectively. All Keeper households were included in the regressions. The samples are necessarily smaller, with fewer statistically significant effects. The 'treatment' is also different. In the case of $Death2$, I am also selecting on marital status and co-residency of spouses, which probably has some bearing on the coefficient estimates.

From Table 9, there is some evidence that there is an inflow of middle aged adults, but this is not significant. We do observe a statistically significant increase in the number of older adults, of 0.151 persons. In general, all the LFP coefficients for the number of middle and older adults are positive, but are usually not significant. There are marginally significant and positive LFP coefficients for the numbers of young female adults, middle aged adults and middle aged male adults. The category which clearly experiences some change are the older adults, where there is a marginally significant increase in the number of employed older adults of 0.065, and a significant increase in the number in the labour force of about 0.09.

[30] Ardington *et al* note that in the dataset that they used, 77% of Loser households lost their pensioner due to death.

[31] The 'age at death' variable in the 'deaths' file in LFS 5 is corrupted, in that the last digit of the variable is missing. This implies that I only observe the age at death in 10 year intervals. I included all deaths where the age at death was non-missing and greater than or equal to 60.

Table 10 presents the same regressions results using the spouse-widow identification of deaths. In this case, very few of the coefficients are significant. We do still observe that the coefficients for middle aged adults and older adults seem broadly consistent with the previous estimates. There are more middle aged adults in the household and more in each category of LFP. For the middle aged women, there is a marginally significant increase in the number in the labour force using the broad definition of 0.121. The strongest results are once again observed for the older adults. The coefficients are all positive, and there is a marginally significant increase in the number of older adults of both genders combined. There is also a significant increase in the number of older adults either working or actively searching for employment.

On aggregate, the death results lend support to a causal interpretation of our aggregate results. The results, broadly speaking, were similar to those observed for the full sample in Table 6. Despite the small number of observations, we observed positive and significant effects on both residency patterns and the numbers in the labour force for the older adult group in particular.

11 Discussion

At this point, it is worthwhile to place these findings in the context of the broader OAP literature. In terms of compositional changes, the results seem consistent with the papers by Edmonds *et al* (2006), who only analyze pension receipt, and Ardington *et al* (2007), who analyze both pension receipt and loss. There is a reduction in the number of resident school aged children, an inflow of middle aged females, and an inflow of older adults into the household.

It is more difficult to reconcile the various findings regarding labour supply. Bertrand *et al* (2003) found that the pension reduces the labour supply of prime aged individuals in three generational households, using cross-sectional data. Ardington *et al* (2007) find the opposite using longitudinal data on both resident and non-resident family members. By

considering the effects on both labour migrants as well as residents, and controlling for person specific unobservable characteristics, they find that prime aged household members are significantly more likely to be employed following pension gain. When considering the loss of a pensioner[32], and restricting to members of either gender who were resident in both waves of their study, they find that prime aged residents were 1.1 percentage points less likely to be employed, although the estimate was not statistically significantly different from zero.

This study finds that pension loss increases employment probabilities for middle aged adults and older adults who maintain household residency. These findings are more consistent with those of Bertrand *et al.*

Several features of this paper are important for appropriate comparative interpretations. First, due to data constraints, this paper is limited to analyzing the effects on people who are resident in the household. I can make no inference regarding changes for those who out-migrate, in-migrate or are somehow attached to the household but not resident in either wave.

Second, the age groups considered differ in a way that is likely to be important. Both Bertrand *et al* and Ardington *et al* focus on 'prime-aged' adults, which they classify as individuals aged 16 to 50 and 18 to 50 respectively. This effectively ignores the 'older adult' category, which in this paper seems to be the most sensitive in terms of both residency and labour supply. In addition, pooling together the groups of youths, young adults and middle aged adults would lead to an estimate that would be some convex combination of the estimates for each group separately. To investigate this possibility, I use the individual level panel data and the age-group of 18 to 50 year olds. The coefficient for employment in this model reduces considerably to 0.024 with a standard error of 0.0139.[33]

[32]There are likely to be asymmetrical responses to pension receipt and pension loss, as documented by Ardington *et al.*

[33]The coefficients for labour supply were negative, smaller in absolute value, and not statistically significant at the 10% level.

A third point of departure from Ardington *et al* is in the geographic scope of the study. Whereas this study is based on limited but nationally representative data, their study is based on incredibly detailed data obtained from a rural and extremely poor part of the country. Regional differences in local economic conditions may result in different responses. To investigate this final possibility, I replicated the individual level analysis for the subset of Keepers and Losers in the rural parts of the KwaZulu Natal province.[34]

Table 11 presents these estimates. In contrast to the estimates for the nation, these estimates indicate that youth are significantly *more* likely to be employed, by 3.4 percentage points. With the exception of female young adults, none of the LFP categories for young and middle aged adults are significant at the 10% level.[35] The coefficient estimates for employment are also much smaller here than for the national sample for middle aged adults, at 0.028 vs. 0.088 for middle aged adults of either gender, 0.031 vs. 0.093 for middle aged females, and -0.032 vs. 0.081 for middle aged males. The geographic restriction thus results in estimates much closer to those of Ardington *et al.* Despite the very small sample sizes, I do still find large and statistically significant increases in labour supply and employment amongst the older adult females.

12 Conclusion

How do poorer households adapt in response to the loss of a valuable economic member? The results presented were consistent with most of the prior empirical literature. Household composition and household labour supply both adjust, with an outflow of dependents and an increase in the number of potentially valuable economic contributors. There is some evidence that the relative importance of remittances increase as well.

Conditional on maintained residency within demographic groups, I find large and significant increases in labour force participation and employment amongst older adults of either gender.

[34]The LFS data does not easily allow for more specific areas of analysis.

[35]Although this lack of statistical significance might be due to considerably smaller sample sizes.

I also find large increases in employment rates amongst middle aged men and women, but no corresponding increase in their labour supply. While the proportionate increases are large, the base population within these households is relatively small, and so the average number of people within households that find employment is also small.

On the other hand, there are over 2.1 million Old Age Pension recipients in South Africa, 1.3 million Disability Grant recipients and almost 6.9 million Child Support Grant recipients.[36] Relatively small labour supply and employment elasticities may have a considerable bearing on national employment levels. Policy makers need to consider the effects of cash grants on the labour supply of non-recipients, while simultaneously being aware that there are other significant positive outcomes that arise from such grants.

[36] See Pauw and Mncube (2007) for details.

13 Tables

Table 1: **Sample Sizes - Cross Sections**

Wave	# HHID	+ No Age missing	+ African 'headed'	+ only 1 HH on property
4	27,356	27,253	21,138	17,549
5	29,011	28,931	22,093	17,775
6	26,474	26,393	20,073	18,378
7	26,702	26,653	20,282	18,301
8	26,825	26,792	20,373	18,347
9	26,829	26,791	20,397	18,312
Total	163,197	162,813	124,356	108,662

Table 2: **Sample Sizes - Matched HH Data**

Panel	# Matched HHs	& $\geq$ 1 'good' indiv. Match	& $\geq$ 1 $pens_t$	Gainers	Keepers	Losers	Losers (strict)
4-5	12,634	8,947	2,395	91	1,967	337	236
5-6	14,143	9,560	2,672	81	2,239	352	268
6-7	14,380	9,282	2,524	94	2,097	333	254
7-8	13,554	8,887	2,385	87	1,973	325	242
8-9	13,702	8,906	2,366	91	1,990	285	220
Total	68,413	45,582	12,342	444	10,266	1,632	1,220

Notes:

1. Sample restricted to African headed households with only 1 dwelling per property, in both periods.

2. 'Strict' definition of a Loser household:

(# residents aged $\geq$ ((pension-age)+2)) at $Time_t$ > (# residents aged $\geq$ pension-age) at $Time_{t+1}$

Table 3: **Selection (on observables): Means in** $Wave_t$

Col	I	II	III	IV	V	VI
Variable	X-sect only	Panel	Diff (II-I)	Full X-Section	Panel - Reweighted	Diff (V-IV)
urban	0.372	0.379	0.007	0.376	0.373	-0.003
hhsize	4.915	5.647	0.732**	5.350	5.401	0.051
# kids young (0-7)	0.738	0.886	0.148**	0.826	0.830	0.004
# kids school (8-15)	0.971	1.186	0.215**	1.099	1.120	0.021
# youth (16 - 20)	0.525	0.633	0.108**	0.589	0.599	0.010
# young adults (21 - 35)	0.910	1.076	0.166**	1.009	1.022	0.014
# middle adults (36 - 50)	0.419	0.480	0.062**	0.455	0.456	0.000
# older adults (51 - pension age)	0.193	0.208	0.014*	0.202	0.202	0.000
# pension aged	1.177	1.197	0.020**	1.189	1.191	0.002
# young adults work	0.190	0.213	0.023**	0.203	0.205	0.001
# middle adults work	0.137	0.152	0.015**	0.146	0.144	-0.002
# older adults work	0.052	0.056	0.004	0.055	0.056	0.001
# young adults in LF (broad)	0.741	0.885	0.144**	0.827	0.841	0.014
# middle adults in LF (broad)	0.321	0.367	0.045**	0.348	0.347	-0.001
# older adults in LF (broad)	0.088	0.094	0.006	0.092	0.092	0.000
# young adults in LF (narrow)	0.517	0.602	0.085**	0.567	0.575	0.007
# middle adults (narrow)	0.244	0.278	0.033**	0.264	0.263	-0.001
# older adults (narrow)	0.070	0.074	0.005	0.072	0.073	0.000

Notes:

1. Sample is all African headed households, with a single dwelling on the property, with no member's age missing, and at least one 'pensioner' in the household
2. Data corresponds to $Wave_t$ - i.e. from Waves 4 - 8
3. Means are unweighted, except in column V
4. The 'single dwelling' requirement excludes 6.25% of the sample, when all the other constraints are satisfied
5. * denotes statistical significance at the 5% level, ** denotes the same at the 1% level
6. The pension age is 60 or above for women, and 65 or above for men

Table 4: **Summary statistics in $Wave_t$ (Mean, and ratio of means)**

	Composition number		# Work proportion		# in LF (broad) proportion		# in LF (nar) proportion	
	Keep	Lose	Keep	Lose	Keep	Lose	Keep	Lose
HH size	5.36	6.39						
kids young	0.86	0.97						
kids school	1.11	1.32						
youth	0.59	0.77	0.035	0.023	0.223	0.240	0.112	0.103
young adults	1.02	1.20	0.193	0.185	0.822	0.816	0.551	0.548
young adult:F	0.55	0.64	0.171	0.155	0.812	0.794	0.507	0.501
young adult:M	0.47	0.56	0.219	0.219	0.834	0.842	0.603	0.601
mid-aged adults	0.46	0.58	0.312	0.326	0.764	0.779	0.578	0.588
mid-aged adult:F	0.25	0.35	0.326	0.291	0.766	0.715	0.570	0.520
mid-aged adult:M	0.21	0.23	0.296	0.378	0.761	0.871	0.588	0.688
older adults	0.17	0.25	0.274	0.317	0.485	0.512	0.372	0.397
older adult:F	0.09	0.13	0.264	0.310	0.445	0.454	0.339	0.349
older adult:M	0.08	0.13	0.286	0.325	0.530	0.571	0.410	0.445
# pens age	1.17	1.33						
N	10,266	1,220						

Notes:

1. Means are weighted by [pweight x IPWeight]

2. The 'proportions' are the ratio of the mean number in a particular labour market category and demographic group, to the mean number in that demographic group.

3. Age-groups: kids young (0-7), kids school (8-15), youth (16-20), young adults (21-35), middle adults (36-50), older adults (51 - pension-age)

Table 5: **Summary statistics: Main Income Source**

Distribution of Main Income Source in Household (%)

	T_0		T_1	
	Keeper	**Loser**	**Keeper**	**Loser**
Salaries and/or wages	15.6	26.3	15.4	31.0
Remittances	4.5	7.3	4.2	26.0
Pensions and grants	77.8	61.7	77.9	35.0
Sales of farm product	0.4	0.5	0.4	1.3
Other non-farm income	1.5	2.8	1.8	5.4
no income	0.2	1.3	0.2	1.4
Unspecified	0.0	0.1	0.1	0.0
N	6,054	710	4212	510

Notes:

1. The T_0 data relates to observations in Panels 4-5, 6-7 and 8-9, the T_1 data relates to observations in Panels 5-6 & 7-8

2. Means are weighted by [pweight x IPWeight]

Table 6: **Regression Results: Composition and Activity (in # of people)**

	Composition		# Work		# in LF (broad)		# in LF (nar)	
Outcome variable	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
△ HH size	-0.86	[0.079]***						
△ # kids young	0.031	[0.033]						
△ # kids school	-0.063	[0.032]**						
△ # youth	-0.007	[0.026]	0.008	[0.007]	-0.027	[0.020]	0.007	[0.014]
△ # young adults	0.039	[0.034]	0.056	[0.019]***	0.04	[0.032]	0.055	[0.035]
△ # young adult:F	0.025	[0.025]	0.032	[0.013]**	0.025	[0.025]	0.031	[0.024]
△ # young adult:M	0.014	[0.025]	0.024	[0.013]*	0.015	[0.023]	0.023	[0.023]
△ # mid-adults	0.063	[0.026]**	0.05	[0.018]***	0.01	[0.025]	0.03	[0.025]
△ # mid-adult:F	0.052	[0.019]***	0.035	[0.011]***	0.021	[0.018]	0.027	[0.016]
△ # mid-adult:M	0.011	[0.016]	0.015	[0.012]	-0.012	[0.016]	0.003	[0.014]
△ # older adults	0.279	[0.022]***	0.102	[0.015]***	0.118	[0.017]***	0.118	[0.016]***
△ # older adult:F	0.144	[0.015]***	0.051	[0.010]***	0.06	[0.011]***	0.058	[0.011]***
△ # older adult:M	0.135	[0.016]***	0.051	[0.011]***	0.059	[0.013]***	0.06	[0.012]***
△ # pens age	-1.064	[0.007]***						

Notes:

1. Robust Std. Errors, clustered at the 'hhid' level are reported

2. Omitted coefficients on variables included in the regression for variables:

Province dummies, wave dummies, urban dummy and no. of pensioners in the '$Wave_t$' period.

3. With the exception of the regression on household size, all the other

regressions also controlled for initial household size.

4. N = 11486 in each of the regressions

5. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level

6. Each coefficient and corresponding standard error is obtained from a separate regression.

7. Age-groups: kids young (0-7), kids school (8-15), youth (16-20), young adults (21-35),

middle adults (36-50), older adults (51 - pension-age)

Table 7: **Regression Results as changes in mean proportions**

	% △ Composition	△ Proportion Work	△ Proportion in LF (brd)	△ Proportion in LF (nar)
HH size	-13.5			
kids young	3.2			
kids school	-4.8			
youth	-0.9	0.011	-0.033	0.010
young adults	3.2	0.039	0.007	0.027
young adult - F	3.9	0.042	0.008	0.028
young adult - M	2.5	0.036	0.006	0.025
mid-aged adults	10.9	0.046	-0.061	-0.011
mid-aged adult - F	15.1	0.050	-0.041	0.000
mid-aged adult - M	4.7	0.044	-0.088	-0.019
older adults	109.9	0.025	-0.047	0.014
older adult - F	112.5	0.024	-0.020	0.029
older adult - M	107.3	0.027	-0.069	0.000
pens age	-79.7			

Notes:

1. Table 7 is derived entirely from Tables 4 and 6.

2. For example, for youth: Number working in Loser households in $Wave_t$ = 0.023*0.77 = 0.018. The new ratio, using the diff-in-diff coefficients, = (0.018+0.008)/(0.77+(-0.007))= 0.034. The difference in proportions is thus 0.034 - 0.023 = 0.011.

3. The tables only show up to 3 digits, but calculations were done at a much higher level of precision. This leads to small discrepancies if one were to calculate these ratios directly from tables 4 and 6 in the paper.

Table 8: **Individual level regressions for changes in labour Market status**

		Dependent variable		
Group	N	△ work	△ in LF br	△ in LF nar
youth	4391	0.015	-0.057	0.005
		[0.011]	[0.026]**	[0.021]
young adults	6856	0.002	-0.027	-0.035
		[0.021]	[0.019]	[0.027]
young adult: F	3765	-0.012	-0.021	-0.074
		[0.027]	[0.029]	[0.036]**
young adult: M	3091	0.016	-0.032	0.007
		[0.030]	[0.024]	[0.035]
mid-aged adults	3316	0.088	0.016	0.039
		[0.030]***	[0.031]	[0.031]
mid-aged adult: F	1850	0.093	0.025	0.025
		[0.037]**	[0.042]	[0.043]
mid-aged adult: M	1466	0.081	0.000	0.058
		[0.041]**	[0.041]	[0.040]
older adults	1537	0.103	0.111	0.107
		[0.043]**	[0.043]**	[0.044]**
older adult: F	856	0.103	0.122	0.131
		[0.059]*	[0.062]**	[0.059]**
older adult: M	681	0.103	0.106	0.083
		[0.053]*	[0.055]*	[0.059]

Notes:

1. Outcome variables in units of △ in labour market status, values of -1, 0 & 1 in data.
2. Reported coefficient corresponds to dependent variable *losepen*
3. Robust Std. Errors, clustered at the 'hhid' level are reported
4. Omitted coefficients on variables included in the regression for variables:

Province dummies, wave dummies, urban dummy and no. of pensioners in the $Wave_t$ period.

5. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level
6. Observations are weighted by [pweight x IPWeight]
7. Each coefficient and corresponding standard error is obtained from a separate regression.
8. Age-groups: youth (16-20), young adults (21-35), middle adults (36-50), older adults (51 - pension-age)

Table 9: **Regression results using 'Deaths1' for identification: Composition and Activity (in number of people)**

	Composition		# Work		# in LF (broad)		# in LF (nar)	
Outcome variable	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
△ HH size	-0.91	[0.261]***						
△ # kids young	-0.017	[0.123]						
△ # kids school	-0.02	[0.093]						
△ # youth	0.003	[0.119]	-0.007	[0.015]	-0.026	[0.071]	-0.109	[0.075]
△ # young adults	-0.047	[0.150]	0.073	[0.089]	-0.168	[0.148]	0.001	[0.153]
△ # young adult:F	0.114	[0.093]	0.112	[0.069]	-0.033	[0.085]	0.161	[0.093]*
△ # young adult:M	-0.161	[0.112]	-0.04	[0.065]	-0.135	[0.105]	-0.16	[0.108]
△ # mid-adults	0.132	[0.108]	0.145	[0.111]	0.2	[0.114]*	0.153	[0.113]
△ # mid-adult:F	0.066	[0.081]	0.058	[0.056]	0.089	[0.082]	0.073	[0.088]
△ # mid-adult:M	0.065	[0.078]	0.087	[0.064]	0.111	[0.063]*	0.08	[0.080]
△ # older adults	0.151	[0.066]**	0.065	[0.037]*	0.094	[0.043]**	0.089	[0.044]**
△ # older adult:F	0.092	[0.063]	0.037	[0.028]	0.036	[0.028]	0.044	[0.031]
△ # older adult:M	0.058	[0.036]	0.028	[0.025]	0.058	[0.033]*	0.045	[0.031]
△ # pens age	-1	[0.000]***						

Notes:

1. Death 1 is obtained from the Deaths file in wave 5
2. Robust Std. Errors, clustered at the 'hhid' level are reported
3. Omitted coefficients on variables included in the regression for variables:
Province dummies, wave dummies, urban dummy and no. of pensioners in the '$Wave_t$' period.
4. With the exception of the regression on household size, all the other
regressions also controlled for initial household size.
5. N = 10313 in each of the regressions, 47 Losers, 10266 Keeper
6. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level
7. Each coefficient and corresponding standard error is obtained from a separate regression.
8. Age-groups: kids young (0-7), kids school (8-15), youth (16-20), young adults (21-35),
middle adults (36-50), older adults (51 - pension-age)

Table 10: **Regression results using 'Deaths2' for identification: Composition and Activity (in number of people)**

	Composition		# Work		# in LF (broad)		# in LF (nar)	
Outcome variable	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
△ HH size	-0.651	[0.233]***						
△ # kids young	0.016	[0.067]						
△ # kids school	0.048	[0.107]						
△ # youth	0.134	[0.094]	0.041	[0.038]	0.026	[0.048]	0.038	[0.075]
△ # young adults	-0.012	[0.108]	0.097	[0.096]	0.014	[0.124]	-0.052	[0.105]
△ # young adult:F	-0.044	[0.077]	0.081	[0.066]	0.03	[0.074]	-0.058	[0.067]
△ # young adult:M	0.032	[0.075]	0.016	[0.056]	-0.016	[0.093]	0.007	[0.085]
△ # mid-adults	0.12	[0.098]	0.069	[0.069]	0.088	[0.087]	0.058	[0.088]
△ # mid-adult:F	0.114	[0.088]	0.103	[0.063]	0.121	[0.070]*	0.104	[0.079]
△ # mid-adult:M	0.005	[0.058]	-0.034	[0.026]	-0.033	[0.035]	-0.046	[0.050]
△ # older adults	0.077	[0.041]*	0.022	[0.024]	0.025	[0.024]	0.078	[0.036]**
△ # older adult:F	0.025	[0.026]	0.001	[0.005]	0.001	[0.005]	0.04	[0.024]*
△ # older adult:M	0.052	[0.032]	0.021	[0.023]	0.024	[0.023]	0.038	[0.027]
△ # pens age	-1.026	[0.019]***						

Notes:

1. Death2 is obtained using the 'spouse - widow' algorithm described in the paper
2. The sample in $death2$ conditions on marital status and co-residency, so it is not entirely comparable to the other two samples.
3. Robust Std. Errors, clustered at the 'hhid' level are reported
4. Omitted coefficients on variables included in the regression for variables:
Province dummies, wave dummies, urban dummy and no. of pensioners in the '$Wave_t$' period.
5. With the exception of the regression on household size, all the other
regressions also controlled for initial household size .
6. N = 10326 in each of the regressions, 60 Losers, 10266 Keepers
7. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level
8. Each coefficient and corresponding standard error is obtained from a separate regression.
9. Age-groups: kids young (0-7), kids school (8-15), youth (16-20), young adults (21-35), middle adults (36-50), older adults (51 - pension-age)

Table 11: **Individual level regressions for changes in labour Market status: Rural KwaZulu Natal**

		Dependent variable		
Group	N	△ work	△ in LF br	△ in LF nar
youth	636	0.034	-0.031	-0.007
		[0.018]*	[0.062]	[0.049]
young adults	840	0.042	-0.008	-0.075
		[0.049]	[0.048]	[0.064]
young adult: F	471	0.036	0.029	-0.166
		[0.068]	[0.088]	[0.080]**
young adult: M	369	0.043	-0.049	0.003
		[0.074]	[0.037]	[0.090]
mid-aged adults	342	0.028	-0.099	-0.039
		[0.059]	[0.079]	[0.066]
mid-aged adult: F	205	0.031	-0.142	-0.086
		[0.068]	[0.097]	[0.090]
mid-aged adult: M	137	-0.032	-0.073	0.004
		[0.109]	[0.138]	[0.071]
older adults	170	0.19	0.145	0.17
		[0.093]**	[0.100]	[0.110]
older adult: F	86	0.283	0.247	0.304
		[0.115]**	[0.126]*	[0.117]**
older adult: M	84	0.021	-0.059	-0.053
		[0.132]	[0.126]	[0.154]

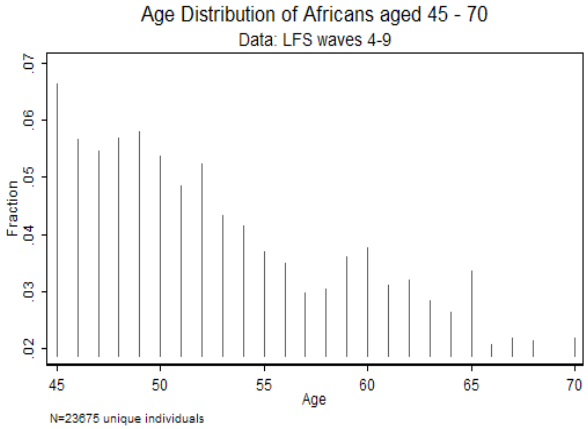
1. Outcome variables in units of △ in labour market status, values of -1, 0 & 1 in data.
2. Reported coefficient corresponds to dependent variable *losepen*
3. Robust Std. Errors, clustered at the 'hhid' level are reported
4. Omitted coefficients on variables included in the regression for variables:

Wave dummies and no. of pensioners in the $Wave_t$ period.

5. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level
6. Observations are weighted by [pweight x IPWeight]
7. Each coefficient and corresponding standard error is obtained from a separate regression.
8. Age-groups: youth (16-20), young adults (21-35), middle adults (36-50), older adults (51 - pension-age)

14 Figures

Figure 1: **Age Heaping in the LFS Data**



Bibliography

Ardington C, Case A & Hosegood V, "Labor supply responses to large social transfers: Longitudinal evidence from South Africa", *American Economic Journal: Applied Economics* 1(1): 2007

Banerjee A, Galiani S, Levinsohn J & Woolard I, "Why Has Unemployment Risen in the New South Africa?", *CID Working Paper No. 134*, October 2006

Becker G, "A Theory of Marriage", *Journal of Political Economy* 81: 813-846, 1973

Beittel M, "The Witwatersrand: Black Households, White Households", in *Creating and Transforming Households: The constraints of the world economy*, edited by Smith J. & Wallerstein I. Cambridge University Press 1992

Bertrand M, Mullainathan S & Miller D, "Public policy and extended families: Evidence from pensions in South Africa", *World Bank Economic Review* 17 (1): 27-50 2003

Case A, "Does money protect health status? Evidence from South African pensions", *NBER working paper 8495*, October 2001

Case A & Deaton A, "Large cash transfers to the elderly in South Africa", *Economic Journal* 108 (450): 1330-1361, Sep 1998

Duflo E, "Grandmothers and granddaughters: Old-age pensions and intrahousehold allocation in South Africa", *World Bank Economic Review* 17 (1): 1-25 2003

Duflo E, "Child health and household resources in South Africa: Evidence from the Old Age Pension program", *American Economic Review* 90 (2): 393-398 MAY 2000

Edmonds EV, Mammen K & Miller DL, "Rearranging the family? Income support and elderly living arrangements in a low-income country", *Journal of Human Resources* 40 (1): 186-207 WIN 2005

Edmonds EV, "Child labor and schooling responses to anticipated income in South Africa", *Journal of Development Economics* 81 (2), 386-414, December 2006

Hamoudi A, & Thomas D, "Pension Income and Well-Being of Children and Grandchildren: New Evidence from South Africa", *California Center for Population Research On-Line Working Paper Series*, CCPR-043-05, October 2005.

Jensen RT, "Do private transfers 'displace' the benefits of public transfers? Evidence from South Africa", *Journal of Public Economics* 88 (1-2): 89-112 JAN 2004

Kingdon G & Knight J, "What have we learnt about unemployment from microdatasets in South Africa?", *Social Dynamics*, 27, No. 1: Spring, 2002.

Keller S, "Household formation, poverty and unemployment - The case of rural households in South Africa", *South African Journal of Economics* 72 (3): 437-483 SEP 2004

Lund F, . "State social benefits in South Africa.", *International Social Security Review* 46, 1993

McLaren Z, "Why is Unemployment so high in South Africa? The Effect of household job loss on labor market outcomes of the unemployed.", Unpublished Manuscript, April 2007

Pauw K. & Mncube L., "Expanding the Social Security Net in South Africa: Opportunities, Challenges and Constraints", *International Poverty Centre, Country Study:Cash Transfer Research Programme, Country Study number 8*, July 2007

Posel D, Fairburn JA & Lund F, "Labour migration and households: A reconsideration of the effects of the social pension on labour supply in South Africa", *Economic Modelling* 23 (5): 836-853, 2006

Ranchhod V, "The effect of the South African old age pension on labour supply of the elderly", *South African Journal of Economics* 74 (4): 725-744 Dec 2006

Ranchhod V, and Wittenberg M, "Income Effects of Time Allocation in Poor South African Households: Evidence from South Africas Old Age Pension", presented at the IZA Topic Week on Nonmarket Time in Economics, Bonn, May 2007

Rosenzweig R. & Stark O., "Consumption Smoothing, Migration and Marriage: Evidence from Rural India.", *Journal of Political Economy*, 97: 905-926, 1989

Sagner A. & Mtati R., "Politics and Pension Sharing in Urban South Africa", *Ageing and Society* 19: 393-416, 1999

Wooldridge J. M.,"Econometric Analysis of Cross Section and Panel Data", MIT Press, 2002