The Migrant Network Effect: An empirical analysis of rural-to-urban migration in South Africa

Caroline Stapleton

ERSA working paper 504

March 2015

Economic Research Southern Africa (ERSA) is a research programme funded by the National Treasury of South Africa.

The Migrant Network Effect: An empirical analysis of rural-to-urban migration in South Africa

Caroline Stapleton

March 2, 2015

Abstract

Recent empirical migration literature in South Africa suggests that access to physical and human capital, in the way of finance and education respectively, are key factors in increasing one's probability of migrating. This paper attempts to extend this literature by directly measuring the extent to which social capital, broadly defined as one's access to a migrant network, affects the probability of rural-to-urban migration. Using the first nationally representative panel dataset in South Africa, the National Income Dynamics Study, I estimate a standard model of migration choice with the inclusion of one's connection to a migrant network. This connection is measured by being part of a household in the baseline wave that contains somebody with current or recent experience as a labour migrant. In line with international migration literature, the empirical results suggest that access to a migrant network increases the likelihood of becoming a migrant (by between 2-3 percentage points). These findings are robust to the inclusion of various controls and therefore suggest that social capital does indeed play a role along with physical and human capital in determining who migrates in South Africa.

Acknowledgements:

I would like to thank Prof. Murray Leibbrandt for his guidance, support and feedback. I would further like to thank the National Research Foundation (NRF) and Economic Research Southern Africa (ERSA) for their financial support of my MSocSc in Economics

1 Introduction

Internal migration of individuals from rural to urban areas is a common occurrence in developing countries as many attempt to escape the poverty and unemployment that often plague rural communities. South Africa is one such example, where for over a hundred years individuals have migrated back and forth between their rural homes and the urban centres in search of employment and higher wages. The decision to migrate, however, is often constrained by the costs and risks involved. This paper will therefore explore the factors that facilitate the migration decision in South Africa, focusing specifically on the role of migrant networks.

Migration in South Africa has its roots in racially discriminatory policies espoused by the pre-apartheid and apartheid governments that restricted the movement and settlement of non-White individuals. As a result, internal migration in South Africa took on an 'oscillating' (or 'circular') pattern whereby individuals migrated back and forth between their rural homes and urban places of employment (Wilson, 2001). Despite the fact that these restrictions have since been lifted, this pattern appears to persist (Posel and Casale, 2003).

Given this fact, the lack of empirical research into migration in South Africa is conspicuous; however, this has largely been due to data limitations and especially the fact that empirical migration studies necessarily require longitudinal data that can track individuals across time. With the development of such datasets in recent years, this has therefore been a growing field of research and much of the recent literature has focused on understanding what factors facilitate the migration decision. These include the state pension (Posel et al., 2006; Ardington et al., 2009; Ardington et al., 2013), housing subsidies (Clarke and Eyal, 2014), age and education (Schiel, 2014; Posel et al., 2006; Clarke and Eyal, 2014).

The factors cited above can be divided into two categories: physical and human capital. Since migration is costly, physical capital such as money or access to credit can relieve one's credit constraints, thereby facilitating migration. Furthermore, human capital can aid the migration decision because if someone is better educated they are more likely to be better informed of, and qualified for, the job opportunities available in the urban centres. In addition to these, a third resource that individuals may draw upon to facilitate movement between areas is *social* capital, which can be broadly understood as one's access to a migrant network (De Haas, 2010; Massey et al., 1993). Migrant networks act as a means through which information can pass from the urban centres to the rural communities, thereby reducing the uncertainty and risk associated with migrating. Furthermore, migrant networks also offer direct assistance to new migrants thereby reducing the financial and psychological costs one might incur when moving across country. While social capital is likely to play a role in facilitating migration within South Africa, it has as yet been ignored in the empirical literature.

This paper will therefore attempt to directly measure the extent to which one's access to a migrant network affects the probability of rural-to-urban migration. Using the first nationally representative panel dataset in South Africa, the National Income Dynamics Study, and defining a rural-to-urban migrant as an individual who is observed moving from a rural area in the baseline wave (2008) to an urban area by Wave 3 (2012), I will estimate a standard model of migration choice. In this analysis, however, I will go further than previous migration studies by also controlling for one's connection to a migrant network.

The paper will proceed as follows: I will begin by discussing the literature around migration, giving a brief overview of the history of migrant labour in

South Africa as well as the relevant theories of migration, focusing particularly on the theoretical role of migrant networks in facilitating the decision to migrate. I will then discuss the data in more detail, followed by an outline of the model and methodology to be used in this analysis. Finally, I will report and discuss the empirical findings and conclude.

2 Literature Review

2.1 Migrant labour in South Africa

Labour migration in South Africa has a long history, stretching all the way back to the discovery of gold and diamonds in the late 19^{th} century. Faced with huge resource reserves which needed to be extracted at the lowest possible cost, the South African Chamber of Mines monopolised the recruitment of African[1] miners, ensuring that they were hired on short-term low wage contracts and housed in single-sex compounds on the mines (Wilson, 2001). This meant that miners could not migrate permanently with their families and were forced to return to their rural homes regularly when their contracts came to an end, laying the foundations of an oscillating system of migration in South Africa whereby migrants moved back and forth between their rural homes and urban places of employment, remitting income back to their families (Wilson, 2001). The Natives Land Act, passed in June 1913, further entrenched this system by limiting the supply of land that African farmers could legally own or rent for independent cultivation, as well as restricting share-cropping arrangements between Africans and Whites on White-owned land. With few agricultural opportunities available to them, many Africans were thus forced to become migrant labourers, moving from their rural homes to urban centres in search of wage work (Walker, 1990; Posel and Casale, 2003).

In 1948 the National Party took power in South Africa and began implementing a legalised system of racial segregation, widely known as apartheid, which restricted the rights and movement of non-White South Africans. A central tenet of the apartheid system was the notion of 'separate development', through which African households were forcibly removed from urban areas and made to live in their designated rural 'homelands' (Posel, 1991). Furthermore, movement out of these homelands was strictly regulated through the use of pass books which contained proof (of employment for example) that the holder of the pass book was legally allowed to be in an urban (White) area for longer than 72 hours (Gelderblom and Kok, 1994; Wilson, 2001). Failure to provide such proof would be grounds for arrest and deportation back to the homelands.

The oscillating system of labour migration that arose out of the growth of the mining sector was therefore firmly entrenched through the apartheid institutions of separate development and restricted movement. Africans were prevented by law from settling permanently in urban areas and those who did

[1] South African racial classifications are as follows: African (black), Coloured (mixed race), Indian/Asian, and White

find employment in the cities were unable to bring their families with them to their places of work (Posel, 2001; Posel and Casale, 2003; Posel, 2004). Migrants would thus leave their rural communities in search of employment while retaining household membership in those communities and would typically support their households financially by remitting income back to them. Additionally, due to the fact that many labour migrants were employed on short term contracts, they would then return to their home towns once their contracts came to an end (Posel, 2001).

After the restrictions surrounding African movement and urbanisation were lifted toward the end of the 1980s, many predicted that the characteristic oscillating system of migration would be replaced by permanent migration to places of employment, since it would finally be possible for entire families to move, rather than just individuals (Posel and Casale, 2003; Posel, 2004). However, in an analysis of the available post-apartheid migration data, Posel and Casale (2003) found no evidence of this trend taking place and in fact found that internal labour migration rose between 1993 and 1999, with labour migrants retaining close economic ties to their original households. This apparent persistence of circular migration is interesting and in unpacking the theory behind the migration decision as well as the available empirical evidence for South Africa, the following section will attempt to provide a more holistic picture of why and how people in South Africa decide to migrate in the present day.

2.2 The migration decision: theoretical motivations and empirical findings

The theory behind the decision to migrate, both internally and internationally, has evolved greatly over time, with each new theory adding its own dimension to this multifaceted issue. The most traditional economic view of migration is that postulated by neoclassical economics, which holds that the migration decision is made at the individual level as a standard cost-benefit calculation – an individual will migrate if the discounted net future earnings (returns to skills) in the destination area outweigh those in the area of origin (Borjas, 1987; Borjas et al., 1992). The decision to migrate is thus purely self-interested and determined by the macro-level supply and demand for labour in the destination and origin labour markets respectively.

Neoclassical economics assumes that all markets are complete and equally accessible by all individuals. These assumptions, however, are largely unrealistic, especially in developing countries, and are challenged in the migration literature by the 'new economics of migration' outlined by Stark and Bloom (1985). This theory assumes that markets – excluding the labour market – such as capital and insurance markets are in fact imperfect and inaccessible; hence the migration decision for a particular individual is instead taken at the household level as a means to spread risk and access additional capital that they are unable to access in their area of origin. The new economics of migration thus incorporates uncertainty and market failure, and rather than being an individual cost-benefit calculation, the migration decision is simply a part of the household's broader strategy for income generation and risk management (Massey et al., 1993).

With reference to the patterns of internal migration observed in South Africa, neoclassical economics would imply that rural-to-urban migration is purely a result of higher returns to skills in urban areas relative to rural areas; however, while this may be true, this theory does not adequately explain why we still observe oscillating patterns of migration in South Africa with migrants retaining strong economic ties to their rural communities. The new economics of migration, on the other hand, goes further than the neoclassical approach by allowing for interaction between the individual's decision to migrate and the interests of the household from which the individual comes. This theory was in fact originally used as a means for understanding an individual's motivation to remit income post-migration, which, as mentioned above, is indeed a prevalent occurrence in South Africa (Lucas and Stark, 1985; Stark and Lucas, 1988). It is plausible that due to the inability for many to generate an adequate income in rural parts of South Africa, sending a member of the household to find work in the cities could constitute a form of insurance against income uncertainty (Posel, 2004).

In addition to the theoretical motivations for migrating, much of the empirical literature, especially the South African literature, is focused on understanding which factors facilitate the actual decision to migrate. As mentioned above, present day South Africa has inherited an ingrained system of oscillating and male-dominated migration; hence it is interesting to explore how the patterns of migration are changing and what factors play a role in assisting the current migration choice.

The standard model of migration choice in the international and local literature is a latent variable model as formulated below:

$$m^* = \gamma X + \varepsilon, \tag{1}$$

$$where\ m = \{ \begin{array}{l} 1\ if\ m^* > 0 \\ 0 \text{ otherwise} \end{array} \tag{2}$$

The above specification implies that the indicator variable m is equal to one if the individual decides to migrate; however, this only happens when some unobserved variable, m^* (some measure of migration feasibility), is greater than zero and this measure is determined by the variables contained in X. While there is some variation in the literature as to what variables should be included here, they can generally be divided up into the following categories: one's individual and household characteristics (gender, age, marital status, and household composition), physical capital (grant receipt, land size and income) and human capital (years of completed schooling) (Posel et al., 2006; Ardington et al., 2009; Ardington et al., 2013; Clarke and Eyal, 2014; Schiel, 2014).

Physical capital is important in two respects. Firstly, migration is costly, so those who come from wealthier households or households with access to capital may find their credit constraints relaxed and therefore be more likely to migrate (Lucas, 1997). On the other hand, individuals in possession of physical capital such as land may be less likely to migrate as they may have commitments to tend to the land or may face losing their rights to the land if they were to leave (Lucas, 1997).

Analysing the impact of the Old Age Pension (OAP) on labour supply in South Africa, Bertrand et al. (2003) find that individuals who share a household with pension recipients are less likely to participate in the labour force. Posel et al. (2006), however, challenge this finding by extending the household unit to include those individuals who are non-resident at the time of the survey. The argument is that these individuals are non-resident because they have migrated for employment reasons and should therefore have been included in the analysis conducted by Bertrand et al. (2003). The authors find that the OAP is indeed positively associated with the probability of being a labour migrant, especially for females, and posit that residing with a pension-eligible individual facilitates migration through the alleviation of credit constraints as well as childcare responsibilities. Using more recent longitudinal data from rural KwaZulu-Natal, Ardington et al. (2009; 2013) expand on Posel et al.'s (2006) analysis, using two waves of data to tease out the causal impact of the OAP on labour supply. The authors similarly find that the income boost provided by the OAP leads to an increase in the probability of migration, most likely due to the alleviation of credit constraints. Furthermore, in line with the fact that owning physical capital may also reduce the likelihood of migrating, Clarke and Eyal (2014) find that individuals from households in receipt of a government housing subsidy are less likely to migrate. These individuals are likely tied down to their physical properties through the housing subsidy and are therefore unable to migrate.

With respect to human capital, the literature suggests that being better educated increases the probability of migration (Todaro, 1980). This may be due to the fact that more educated individuals are better informed of employment opportunities in the urban areas or perhaps that the returns to more educated individuals are higher in urban areas relative to rural areas. According to the neoclassical approach, this would certainly make these individuals more likely to make the choice to move. Alternatively, according to the new economics of migration, if a household is going to send an individual to the city in search of wage work to spread the household risk, it would make sense that they should choose the individual most likely to find a job to be the one to go. This would in many cases be the individual with the most education. Various empirical studies of the migration decision in South Africa find that additional human capital, measured by higher completed years of schooling, is indeed associated with an increased probability of migration (Posel et al., 2006; Clarke and Eyal, 2014; Schiel, 2014).

The above theories explain migration with reference to micro-level decision making (the individual or the household) and macro-level structural determinants (expected wages, labour demand and supply, market failure), but they provide little explanation of how information regarding these structural determinants is disseminated from urban to rural areas in order to influence the migration decision. This gap can be bridged by the notion of 'social capital', which in addition to physical and human capital is a third resource that individuals can draw upon to facilitate movement between areas (De Haas, 2010; Massey et al., 1993). Social capital adds a further layer of complexity to our understanding of the migration decision and it is one that is not well-documented in the empirical South African migration literature (Kok et al., 2003).

2.3 Migrant networks

Social capital can be understood as one's access to migrant networks – "sets of interpersonal ties that connect migrants, former migrants, and non-migrants in origin and destination areas through ties of kinship, friendship, and shared community origin." (Massey et al., 1993: 448). Network theory posits that having network connections in the destination area serves to facilitate migration via two broad mechanisms. Firstly, the network acts as a means by which potential migrants can access information regarding the returns to migrating (i.e. information concerning employment opportunities or expected wages). By virtue of this fact, migrant networks can both stimulate and discourage migration depending on the nature of the information being disseminated. For example, if individuals in rural areas hear via the migrant network that job opportunities in the cities are rife, they may be encouraged to move; however, if the news is less positive, individuals may choose to remain in their rural communities (Gelderblom and Adams, 2006). Having this information therefore allows potential migrants to update their existing beliefs regarding the returns to migration and therefore reduces the risk associated with the decision to migrate (Winters et al., 2001).

Secondly, members of the migrant network might offer direct assistance to the potential migrant in the way of food, transport, accommodation or access to employment opportunities, which directly reduces the financial cost of moving, thereby further alleviating the credit constraints associated with the migration decision (Winters et al., 2001). Furthermore, they may also offer 'social assistance' by introducing migrants into their social circles, offering emotional support and 'showing them the ropes', thereby also reducing the psychological costs associated with migrating (Gelderblom and Adams, 2006).

Theoretically, migrant networks therefore reduce both the cost and risk associated with movement, increasing the net returns, and thus the probability of migrating. In fact, an implication of network theory is that migrant networks not only facilitate the initial migration decision but also act as a force for perpetuating migration (Kok et al., 2003). Once the first migrants have left their home towns and established migrant networks of their own, it becomes easier for potential migrants to move given the reasons stated above. Each new migrant then establishes a new link in the migrant network with their own set of social ties to the area of origin as well as the destination area. This ongoing cycle linking potential migrants and the migrant network through kinship and friendship was defined by Massey (1990) as the 'cumulative causation of migration' – migration creates more migration through migrant networks.

In light of the theoretical importance of migrant networks in facilitating as well as perpetuating migration, many have sought to empirically estimate their impact on the probability of migrating. Building on the standard model of migration choice outlined above, the studies cited below all take the following form, where NET is the set of variables measuring one's access to social capital and X contains the standard control variables detailed above:

$$m^* = \beta NET + \gamma X + \varepsilon, \tag{3}$$

$$where\ m = \{ \begin{array}{l} 1\ if\ m^* > 0 \\ 0 \text{ otherwise} \end{array} \tag{4}$$

Analysing the patterns of Mexico-U.S. migration and empirically testing various theoretical predictions, Massey and Espinosa (1997) find that access to social capital does indeed significantly increase the probability of initial migration. The authors employ four different migrant network variables: an indicator variable for whether a respondent's parents had begun migrating to the U.S. at the time of the survey; the number of the respondent's siblings who had begun migrating to the U.S.; the proportion of community members older than 15 who had been to the U.S.; and an indicator variable for whether or not a member of the respondent's household had been legalised under the U.S. Immigration Reform and Control Act. All variables are found to be positive and statistically significant.

In another study of Mexico-U.S. migration, Winters et al. (2001) measure access to migrant networks by controlling for the number of current migrants (at the time of the survey) and historical migrants (who subsequently returned home prior to the survey) in the respondent's household and community respectively. The authors find that the *current* migrant network variables positively and significantly influence the probability of migration. This stands in contrast to having members of one's household or community classified as *historical* migrants, which does not have a significant impact on the decision to migrate, suggesting that the information and assistance that migrant networks provide is specific to the time at which an individual is deciding whether or not to move. This makes intuitive sense as the information provided by historical migrants would likely be out of date and therefore not very useful to a potential migrant.

Given these findings, recent literature concerning Mexico-U.S. migration has delved even further into the role of migrant networks, exploring for instance the actual mechanisms by which migrant networks assist potential migrants (Dolfin and Genicot, 2010) and the extent to which migrant networks affect the selection of Mexican migrants into the U.S. (McKenzie and Rapoport, 2010).

The same type of analysis conducted by Massey and Espinosa (1997) and Winters et al. (2001) can also be used to understand *internal* rural-to-urban migration. Similar to the patterns of migration in South Africa, Zhao (2003) highlights the circular system of migration in China, which by nature serves as a means to maintain and strengthen migrant networks. Zhao (2003) estimates the impact of migrant networks on the probability of being a labour migrant, using the number of 'experienced migrants' in the respondent's village (individuals that had at least 48 cumulative months of migration at the time of the survey) and the number of 'return migrants' (individuals that had some migration experience but returned to the village prior to the survey) to measure one's access to a migrant network. The author finds a positive and significant effect of migrant networks on the probability of migrating internally. This empirical finding is in line with various descriptive studies that also highlight the importance of migrant networks in facilitating the decision to migrate internally. These studies have been conducted using data from India (Banerjee, 1983), Germany (Bauer and Zimmerman, 1997), and the Philippines (Caces, 1986).

Given the oscillating nature of migration and the history of migrants retaining strong economic and social ties to their rural households and communities, it seems that migrant networks could indeed play a fundamental role in facilitating rural-to-urban migration in South Africa. Further to that, it is interesting to consider the role of migrant networks in South Africa as means through which individuals classified officially as 'discouraged unemployed' rather than 'searching unemployed' could in fact be searching for work.[2] Due to the fact that many individuals live in remote rural areas, the cost of travelling to the nearest town or city to look for work may simply be too high. Instead, it is plausible that these individuals, who technically do not fall into the official definition of being unemployed, are indeed searching for work via their migrant networks (Schöer and Leibbrandt, 2006; Posel et al., 2014).

While there has been some loose discussion surrounding the importance of migrant networks in facilitating migration within South Africa (Kok et al., 2003; Gelderblom and Adams, 2006; Gubhaju and De Jong, 2009; Schiel, 2014) there has been no empirical migration study (to my knowledge) that has attempted to directly measure the impact of such networks on the probability of migration. This paper will therefore serve as an attempt to fill this void, directly exploring the role of migrant networks in facilitating rural-to-urban migration in South Africa. I will draw on the literature reviewed above to outline a standard model of the individual's migration decision, paying specific attention to the individual's access to migrant networks, where the key network variable is an indicator variable for the membership of a 'labour migrant' in the respondent's household prior to migration. It is important to note that this is a first attempt at assessing the direct impact of migrant networks in South Africa and will focus on the individual's broad household membership relationship to labour migrants

[2] 'Searching unemployed' are those individuals who are unemployed and have actively looked for work within the previous four weeks, while 'discouraged unemployed' have not actively looked for work within the previous four weeks despite the fact that they still wish to find a job (Ranchhod, 2009). Discouraged unemployed are often excluded from the official measure of unemployment (the 'narrow' measure of unemployment as opposed to the 'broad' measure where the discouraged unemployed are included), and therefore excluded from the labour force. This, however, is problematic and can lead to an underestimate of unemployment in South Africa. Kingdon and Knight (2006) find no distinction between the searching unemployed and the discouraged unemployed that warrants excluding the latter from the labour force. Building on this work, Lloyd and Leibbrandt (2014) find that discouraged unemployed are significantly less happy than the searching unemployed, but conclude that they should still be included in the official measure of unemployment. The discouraged unemployed have merely stopped directly searching, not because they have any less desire to find a job, but because the costs of searching are too high.

as opposed to, for instance, genetic (or kin) relatedness, which has proved to be somewhat impactful in predicting the amount of remittance income sent by migrant workers (Bowles and Posel, 2005). While unpacking the network effect to discern kin from other household members is worthy of discussion (and will indeed be touched upon later), it will not be the central focus of this study.

In order to conduct the above-mentioned analysis, I will use the first nationally representative panel dataset in South Africa, the National Income Dynamics Study (NIDS). The following section will describe this dataset in more detail and set up the analysis by examining the descriptive statistics of the sample.

3 Data and Descriptive Statistics

3.1 Data

The data for this analysis will be taken from the NIDS, Waves 1 and 3 (2008 and 2012), a nationally representative household survey conducted by the Southern Africa Labour and Development Research Unit (SALDRU) at the University of Cape Town. The baseline wave consists of 28226 individuals (from approximately 7300 households), young and old, from across all spectrums of South African society (Southern Africa Labour and Development Research Unit, 2014).[3] The survey collects detailed information at the individual and household level through separate questionnaires for individuals aged 15 and over (completed by the individual in question), children younger than 15 (completed by the mother or caregiver of the child) and the household (completed by the oldest woman in the household). Since most migrants in South Africa have historically been, and still are, prime-age African adults, I will limit the sample to only these individuals. This is in line with other empirical studies concerning the migration decision in South Africa (Posel et al., 2006; Ardington et al., 2009; Ardington et al., 2013; Clarke and Eyal, 2014). In this paper, 'prime-age' will refer to individuals aged 18 to 55 in Wave 3.

As noted by Posel (2010), NIDS is different to most previous South African surveys as it employs a broad household residency requirement, which recognises the fact that in South Africa it is possible for an individual to be a member of more than one household or to be a member of a household in which they do not physically reside for much of the year, leading to an important distinction between resident and non-resident household members (Posel et al., 2006; Posel, 2010).[4] Non-resident household members are individuals who are listed on the

[3] SALDRU employed a stratified, two-stage cluster sample design in the sampling of households for the baseline wave of NIDS. 400 Primary Sampling Units (PSUs) were randomly selected within the assigned strata (53 district councils) from StatsSA's 2003 Master Sample consisting of 3000 PSUs. NIDS' target population was private households (as well as residents in workers' hostels, convents and monasteries) in all nine provinces of South Africa (Leibbrandt et al., 2009).

[4] To be listed on the household roster in NIDS, individuals should have lived under the same roof (at the same homestead) for at least 15 days during the previous year, should share food from a 'common pot' and share resources from a common resource pool. To be further classified as a *resident* household member, an individual should spend at least four nights per

household roster but are (or have been) absent from the household for a period of time. Those non-resident household members who have been absent due to employment reasons (working or looking for work) are identified in the literature as 'labour migrants' and have been the subject of a number of recent migration studies in South Africa (Posel et al., 2006; Posel, 2010; Ardington et al., 2009; Ardington et al., 2013; Clarke and Eyal, 2014).

Instead of using this definition of a labour migrant as the dependent variable, as was the case in the studies cited above, I have used it to create an indicator variable for the respondent's connection to a migrant network. This indicator variable is equal to one if the respondent is part of a household in Wave 1 (2008) that contains a labour migrant, defined by Posel (2010) as a member of the household who is absent for at least a month during the year for employment reasons. This measure is in line with the international migration literature reviewed in the previous section.[5] As a dependent variable, I will define a migrant as someone who moves from a rural area to an urban area between Waves 1 and 3.

An important implication of the way in which a migrant has been defined is that the individuals in the sample under analysis all have to have been observed in both Wave 1 and Wave 3 in order to establish which individuals migrate between 2008 and 2012 and which do not. This gives rise to a potential problem encountered in all analyses using panel data – attrition bias. If it is the case that sample attrition appears to be random, then simply analysing those who are observed in both Wave 1 and Wave 3 will not bias the analysis; however, if there is selective attrition in that those who attrite are somehow different based on observable characteristics to those who remain in the sample, then there is a chance that the statistical results could be biased. More specifically, if an individual who was observed and interviewed in the baseline wave of NIDS is for some reason not located or tracked down for re-interview in Wave 3, it is not possible for anyone to know whether or not that individual would have migrated.

Attrition bias in this instance is even more worrying as it is closely related to migrant self-selection which is a common concern in any migration study – if those individuals who attrite fail to be located because they have migrated and are somehow different from those who do not attrite, any analysis of the migration decision based only on those individuals who remain in the sample

week in the household (Southern Africa Labour and Development Research Unit, 2013).

[5] Massey and Espinosa (1997) employ four migrant network variables: an indicator variable for whether a respondent's parents had begun migrating to the U.S. at the time of the survey; the number of the respondent's siblings who had begun migrating to the U.S.; the proportion of community members older than 15 who had been to the U.S.; and an indicator variable for whether or not a member of the respondent's household had been legalised under the Immigration Reform and Control Act. Winters et al. (2001) measure access to migrant networks by controlling for the number of current migrants (at the time of the survey) and historical migrants (who subsequently returned home prior to the survey) in the respondent's household and community respectively. Zhao (2003) similarly controls for the number of 'experienced migrants' in the respondent's village (individuals that had at least 48 cumulative months of migration at the time of the survey) and the number of 'return migrants' (individuals that had some migration experience but returned to the village prior to the survey).

could easily be skewed one way or another. This type of bias has been greatly discussed in the international migration literature with reference to the earnings assimilation of migrants in their host countries and the potential impact of selective out-migration on the estimation of immigrant earnings profiles (Borjas, 1985; Borjas, 1989; Constant and Massey, 2003).

In order to establish whether or not attrition is likely to have an impact on the current analysis, it is important to compare all those observed in the baseline wave to those observed in both Wave 1 and Wave 3. Table 1 depicts the sample means of Wave 1 observable characteristics for the full sample of all adults in Wave 1 as well as the restricted sample of Africans aged 18-55 in Wave 3, the specific sample under analysis in this paper.

Upon examination of Table 1 it seems evident that attrition bias in unlikely to be a problem in this analysis. Of all adults observed in Wave 1, 1388 (roughly 7%) attrite between Wave 1 and Wave 3. Furthermore, when the sample is narrowed down to just prime-aged African adults (aged 18-55 in Wave 3) this number falls to 593, implying an attrition rate for the sample under analysis of only 5% between Waves 1 and 3. Taking a closer look at the table, it is indeed evident that there are very few differences in the sample means between those observed at baseline and those observed in both Wave 1 and Wave 3. Columns 1 and 2 compare the sample means of all adults in Wave 1 and it appears that, barring the odd percentage point difference here and there, the means are almost identical. Perhaps the only noticeable (although still marginal) difference between the two samples is that those observed in both Wave 1 and Wave 3 are slightly better educated. The same sort of pattern is observed for the prime-aged African adults; however, the differences appear even less pronounced than those in the full sample.

All of the above suggests that attrition is unlikely to bias the results in this instance; however, as an extra precaution to ensure that the results are indeed unbiased, I will use the panel weights supplied by the NIDS staff in all multivariate regressions. These weights are designed to adjust the observed sample for subsequent non-response of those observed in Wave 1 and should dampen attrition bias based on observable characteristics (Wittenberg, 2009; De Villiers et al., 2013).

3.2 Descriptive statistics

Table 2 examines the restricted sample (Africans aged 18-55 in Wave 3) in more detail, dividing the sample up into migrants (individuals who move from an urban area to a rural area between Waves 1 and 3) and non-migrants in order to get a broad understanding of what makes them different from one another.[6] This is important so as to control adequately for any confounding factors that may influence the results in the multivariate regressions to follow. In terms of the individual characteristics, it appears that the rural-to-urban migrants tend

[6] See Appendix for Table 10 where the sample is further broken down into males and females – the pattern of differences between migrants and non-migrants is the same across gender.

to be better educated than non-migrants where a lower percentage of migrants have no schooling or just primary school education relative to non-migrants and there are proportionately more migrants with some secondary education (excluding matric). It appears that there is no significant difference between the proportion of migrants and non-migrants with matric (Grade 12) as well as some form of tertiary education, suggesting perhaps that individuals who migrate from a rural area to an urban area (possibly in search of work) drop out of secondary school in order to so, or perhaps are forced to migrate in search of work because they have failed to complete secondary school. The bivariate relationship between migrant status and education is depicted graphically in Figure 1. From this it is evident that the relationship is highly non-linear with the proportion of migrants increasing sharply for individuals with some secondary education and then decreasing sharply after matric.

Furthermore, in line with the above theory that individuals migrate in search of work, it is evident from Table 2 that a much larger percentage of non-migrants (43%) are employed in Wave 1 (prior to migration) relative to migrants (25%). A final point on the individual characteristics, which again supports the above theory, is that migrants are on average a few years younger than non-migrants. Taking a closer look at the migrant-age relationship in Figure 2, it appears that the percentage of migrants increases sharply from the early teenage years until about 18 years of age (in Wave 1) and then decreases. This links back to the fact that migrants are more likely to have fewer completed years of schooling and perhaps migrate when they should in fact still be in school.

The household characteristics of migrants also appear to be somewhat different to non-migrant households (prior to migration taking place). First of all, it appears that individuals who subsequently migrate come from larger households and, secondly, they are more likely to come from households in receipt of state pension income. Of the migration studies previously conducted in South Africa, a number of them have focused on the role of the state pension (OAP) in facilitating migration, especially for women (Posel et al., 2006; Ardington et al., 2009; Ardington et al., 2013). Given this evidence, it is thus particularly important that household pension receipt be controlled for in the multivariate analysis. Table 2 also highlights the fact that 100% of migrants come from households in rural areas. This is a direct result of the fact that a migrant has been defined as an individual who moves from a rural area to an urban area between Waves 1 and 3.

Of particular importance in this study is the difference between migrants and non-migrants with respect to the migrant network variables. Before running a multivariate analysis, the data is suggesting that those who subsequently migrate are more likely than non-migrants to come from households containing a labour migrant, pointing to the fact that access to migrant networks is indeed likely to have an impact on an individual's probability of migration. Furthermore, if those defined as labour migrants are merely members of the household who are employed, then it might be the case that a relationship with a labour migrant simply represents one's connection to the labour market and not to a migrant network. Table 2, however, suggests that this is not the case. Individuals who subsequently migrate are 14 percentage points less likely to come from a household containing an employed individual than non-migrants. This in turn provides further impetus for the theory espoused above that those who migrate from a rural area to an urban area do so in search of employment, seeing as those who do not migrate possibly have less need to do so due to the fact that they are already employed or they already reside with an employed individual who is most probably earning an income.

As a final note on the sample in question, it is important to determine how many individuals classified as labour migrants in 2008 subsequently migrate again between survey waves. This is because if all rural-to-urban migrants in 2012 were also labour migrants in 2008, then any positive relationship between having a labour migrant as a member of one's household prior to migration and the probability of migrating would not represent a network effect but rather a behavioural effect of previous migration on future migration. Fortunately, according to Table 3, it appears that only 70 (10%) of rural-to-urban migrants were themselves classified as labour migrants in Wave 1, allowing enough variation to tease out any network effect if there is one.

4 Model and Methodology

The empirical strategy for this analysis is drawn from that used by Zhao (2003) in her study of internal migration in China. Zhao (2003) first estimates the standard model of migration choice, outlined in the literature review above, which excludes migrant network variables. This is to ensure that her initial results are in line with previous migration studies conducted in China.

Following that, she then introduces the migrant network variables to isolate the impact of access to migrant networks on the migration decision. Zhao (2003) uses two migrant network variables: one for the number of 'experienced migrants' in the respondent's village (individuals that had at least 48 cumulative months of migration at the time of the survey) and the other for the number of 'return migrants' (individuals that had some migration experience but returned to the village prior to the survey). She finds a positive migrant network effect for the number of experienced migrants in the village, but no effect related to the number of return migrants, implying that current migrants provide more direct help to potential migrants than do individuals who have migrated in the past and returned home.

I will employ a similar strategy to Zhao (2003) by first estimating a standard migration choice model based on those used in local migration studies by Posel et al. (2006) and Ardington et al. (2009). Subsequent to that I will extend the analysis by including a set of migrant network variables, which has not yet been done in the South African migration literature. The final migration choice model will look as follows:

$$Migrant_{ih,t+2} = \beta NET_{h,t} + \gamma X_{ih,t} + \varepsilon_{ih,t} \tag{5}$$

where for individual i in household h observed in survey wave t, the dependent variable $Migrant$ is an indicator variable equal to one if the respondent is observed to have moved from a rural area in Wave 1 to an urban area in Wave 3, and zero otherwise. This binary choice is modelled as a function of a set of migrant network variables, NET, the key variable of which is an indicator variable equal to one if the respondent originates from a household in Wave 1 containing a labour migrant (defined as an individual who is absent from the household for at least a month during the year for employment reasons – this is derived from the dependent variable used by both Posel et al. (2006) and Ardington et al. (2009; 2013)). Also included in this set is an indicator variable equal to one if the origin household receives monthly remittance income, and an indicator variable equal to one if the origin household contains an employed individual. Finally, X is a set of control variables for the individual and household characteristics that distinguish migrants from non-migrants prior to migration. This includes a gender dummy, a full set of indicators for the respondent's completed years of schooling, a quartic in age (in order to capture the non-linearity observed in Figure 2), an indicator variable equal to one if the respondent is married, an indicator variable equal to one if the origin household contains a pension-eligible individual,[7] and the number of household members in the following age categories: 0-5, 6-17, 18-55, 56 and over. All of the above control variables are taken from the baseline wave of NIDS.

The migration choice model will be estimated using a linear probability model (LPM) as well as a probit model to serve as a comparison. This is to account for the fact that the LPM can lead to predictions outside of the [0,1] interval, particularly at the tails of the distribution (Wooldridge, 2002). As I am only interested in the average partial effects, the LPM should be sufficient, but the probit model will act as a useful robustness check (Wooldridge, 2002). I will first estimate the model for the entire sample and then by gender so as to tease out any gender-specific effects. All regressions will be weighted by the Wave 3 panel weights to account for sample attrition between waves.

5 Empirical Analysis

The results of the multivariate regressions outlined above are reported in Tables 4, 5 and 6. Table 4 contains the results for the full sample under analysis – Africans aged 18-55 in Wave 3 (2012). Column (1) is the standard migration choice model based on those estimated by both Posel et al. (2006) and Ardington et al. (2009); column (2) introduces the first (and key) migrant network variable, while column (3) includes two additional migrant network control variables. All of the above are LPMs while column (4) reports the average partial effects (APE) for the final probit model.

[7] In 2008 (Wave 1) the age eligibility criteria for the state pension was 60 for females and 65 for males – this has subsequently been changed to 60 for all. Upon reaching the required age, individuals are eligible for the pension provided they meet certain means test criteria.

5.1 Main results

According to Table 4, the decision to migrate does not appear to be gender biased as the coefficient on the female dummy is very small and statistically insignificant across all models. As discussed previously, however, age and education do appear to play a role. The quartic in age is highly significant across all LPMs, capturing the non-linear relationship observed in Figure 2. The first and second polynomials capture the distinct inverted U-shape observed in the graph, and while the coefficients on the third and fourth polynomial are very small, these are likely capturing the flattening out and the upturn at the tail observed in Figure 2. In terms of education, it seems that individuals with only primary school education are approximately 3 percentage points less likely than those with matric (the omitted category) to migrate, while those with some form of tertiary education are about 5 percentage points more likely to migrate, all else equal. Interestingly, while the descriptive statistics discussed above seemed to imply that individuals with incomplete secondary education were more likely to be migrants, this does not appear evident from the regression results – after controlling for other individual characteristics, those with incomplete secondary education are no more or less likely than those with matric to be migrants. The above all seems to imply that the probability of being a migrant is higher the younger and better educated one is.

In line with both Posel et al. (2006) and Ardington et al. (2009; 2013) individuals from pension-eligible households are about 2 percentage points more likely to be migrants. This finding supports the theory that individuals will be more likely to migrate if their credit constraints are relaxed. People living with pension-eligible individuals have a chance of benefitting from the cash injection provided by the state pension which might then facilitate the move across country. Interestingly, the probability of migrating increases slightly for each additional child in the origin household and decreases for each additional prime-age adult. This may signify that a household with relatively more children contains fewer individuals that can work for a wage, thus the need to migrate in search of work would be greater, while those who live in households with relatively more adults would face less need to find work as there are multiple individuals who could theoretically bring in an income. The above findings are all largely in line with previous migration studies in South Africa.

The principal question this analysis is striving to answer is whether and to what extent migrant networks influence the migration decision in South Africa. Columns (2)-(4) thereby extend the existing migration choice model by including controls for an individual's access to migrant networks. From column (2) it is evident that access to a migrant network increases the probability of migrating by approximately 2 to 3 percentage points. This finding is statistically significant at the 1% level and robust to the inclusion of further migrant network controls (columns (3) and (4)). As mentioned previously, it could be the case that one's connection to a labour migrant is merely capturing one's connection to the labour market as opposed to a migrant network; however, after controlling for one's connection to an employed individual, the migrant network variable remains highly significant. Furthermore, the control variable is itself negative and significant indicating that sharing a household with an employed individual reduces the likelihood of migration by approximately 2 percentage points, all else equal. Finally, it could also be the case that one's connection to a labour migrant is capturing the effect of remittance income which might facilitate migration in the same way as pension income; however, once again, including a control for the receipt of remittance income does not change the coefficient on the migrant network variable and is itself statistically insignificant.

While there appears to be no gender bias in that migration rates are fairly similar between males and females, Tables 5 and 6 yield some interesting results which suggest that the factors facilitating migration tend to differ slightly across gender. One of the most noticeable differences between the male and female results is that for males (Table 5), age appears to be a completely insignificant factor statistically speaking, while for females (Table 6) the results for age are in line with those discussed above. This may be due to the fact that males are generally considered to be the breadwinners and therefore may be likely to migrate at any stage of their adulthood in search of work, while for females, who might be expected to care for children in their latter adulthood years, the likelihood of migration is higher in young adulthood. A further distinction is that originating from a pension-eligible household only appears to impact on the migration decision for females. This finding is in line with that of Posel et al. (2006) who posit that females not only benefit from the inevitable cash injection provided by the pension, but also perhaps find it easier to migrate if they are relieved of childcare constraints by the pensioner in receipt of the pension. Taking a closer look at the migrant network effect by gender, it appears that both males and females are more likely to migrate if they have some connection to the migrant network, and the magnitude of this effect is similar for both. Interestingly, however, it appears that it is the men who are less likely to migrate given that they are a member of a household containing an employed individual, while for women this is statistically insignificant. This perhaps suggests that men are more likely to migrate for employment reasons (and may not do so if they or another member of their household is already employed) while perhaps women migrate for other reasons such as marriage.

5.2 Robustness checks

As mentioned in the earlier discussion of the descriptive statistics, it is important to tease out the potentially confounding impact of those classified as labour migrants in 2008 as well as rural-to-urban migrants in 2012. If, for instance, all individuals who had previous migration experience subsequently migrated again between survey waves, then the network variable 'household contains a labour migrant' would simply capture the previous migration experience of the individual in question and not the intended network effect.

While just 10% of the rural-to-urban migrant sample is classified as both migrant types (see Table 3), it is still important to test whether or not these individuals are in fact driving the results. Table 7 reports the regression results from this robustness check, where an additional control has been included in the main model for whether the individual in question was a labour migrant in their original household and the migrant network variable has been amended to zero for those individuals who were the *only* labour migrants in their household of origin. The migrant network variable along with this additional control (both in bold) indicate that, fortunately, this does not seem to be an issue in the analysis. Access to a migrant network remains statistically significant, suggesting an increase in the probability of migration of approximately 2-3 percentage points, all else equal, and the labour migrant control is statistically insignificant with small coefficients. This could suggest that, after controlling for relevant individual and household characteristics, those who have migrated previously are no more or less likely to migrate between waves relative to those with no previous migration experience. Alternatively, since there are just 70 individuals who classify as both labour migrants in 2008 and rural-to-urban migrants in 2012, perhaps there is simply not enough statistical power to tease out the impact of previous migration experience.

An additional issue raised in this paper's earlier discussion is that one's connection to a labour migrant may capture one's connection to the labour market instead of one's access to a migrant network. While this has been refuted by the above analysis, a further robustness check is required to establish whether the negative coefficient on the variable 'household contains an employed individual' is capturing the individual's *own* employment status or their relationship with some *other* employed individual in the household, as both could plausibly lead to a reduction in one's probability of migrating. The results of this second robustness check are reported in Table 8 and show that it is in fact the former effect at play – if individuals are already employed, they are approximately 1.5 percentage points less likely to migrate than those who are unemployed or not economically active, all else equal. While being a member of a household with some other employed individual excluding oneself also has a negative coefficient, this coefficient is very small and statistically insignificant. This finding again supports the theory that those who migrate from rural to urban areas do so in search of employment.

5.3 Genetic relatedness

The above analysis focuses on one's connection to a migrant network via household membership, which, as stated earlier, can be rather loosely defined in South Africa to include individuals that do not physically reside in the household for much of the year. Furthermore, it is not uncommon in South Africa for households to include multiple generations and extended family. A natural extension of the above analysis is therefore to see if it is possible to break down the migrant network effect even further in order to test whether this effect differs by how closely related one is to a labour migrant. This idea is loosely drawn from Bowles and Posel's (2005) work on the role of genetic relatedness in predicting the amount of remittance income sent by migrant workers in South Africa.

In an attempt to tease out the impact of genetic relatedness with respect to the migrant network effect, I have identified a respondent's immediate family or 'kin' (mother, father and siblings) and included in the analysis an additional indicator variable equal to one if the respondent's kin was classified as a labour migrant in 2008. Further to that, I have also included an interaction term between this indicator variable and the original migrant network variable which will capture the effect of those whose kin is classified as a labour migrant *and* share a household with the respondent.

The results of this analysis are reported in Table 9. From column (2) it appears that, while still controlling for having a labour migrant in the household, one's immediate relation to a labour migrant has a positive but insignificant effect. Column (3), however, yields some interesting results. It appears that the above effect can be further broken down to distinguish between simply being related to a labour migrant and being related to a labour migrant *in the same household*. The final results suggest that being related to a labour migrant who shares a household with you produces a very strong network effect, while simply being related to a labour migrant (and not necessarily sharing a household with them) almost completely cancels this effect out with astrong negative impact on the probability of migrating (together yielding the positive but insignificant effect observed in column (2)). The coefficient on the interaction term seems intuitive – for those who share a household with their immediate family and those family members have recent migration experience, it is likely that these individuals would have easy access to a migrant network which would offer direct support in the migration decision. The opposite effect of being immediately related to a labour migrant (regardless of household membership) is puzzling. Perhaps individuals who are immediately related to a labour migrant, regardless of household membership, are less likely to migrate (all else held constant) if familial or caregiving responsibilities have fallen on them due to the absence of their labour migrant relation(s).

These explanations are speculative, but it can be said that given the robustness of the migrant network variable and the statistical significance of the kinship variables, migrant networks definitely affect one's probability of migrating; however, there appears to be some form of household bargaining process at play which determines exactly how one's connection to a labour migrant affects one's migration choice. This analysis can only go so far in unpacking this household bargaining process, but perhaps more can be said on this issue when future waves of the NIDS data become available.

6 Conclusion

Rural-to-urban migration has been a fixture in South African history for a long time and persists to this day; however, due to data limitations and a lack of available longitudinal data, the South African migration literature is still in its infancy. Recent empirical work has therefore largely focused on understanding the basic factors that facilitate the individual's migration decision and the findings of these studies suggest that access to physical and human capital, in the way of finance and education respectively, are key factors in increasing one's probability of migrating (Posel et al., 2006; Ardington et al., 2009; Ardington et al., 2013; Clarke and Eyal, 2014; Schiel, 2014). These studies, however, have not accounted for an individual's access to *social* capital, which can be broadly understood as one's access to a migrant network and has been widely recognised in the international migration literature as an important factor in reducing the costs and risks associated with migrating (Massey, 1990; Massey et al., 1993; Massey and Espinosa, 1997; Winters et al., 2001; Zhao, 2003; McKenzie and Rapoport, 2010; Dolfin and Genicot, 2010). This paper has therefore extended the current South African migration literature by focusing specifically on the role of migrant networks in facilitating rural-to-urban migration.

In order to formulate the empirical strategy, this analysis has actively drawn on the local and international migration literature. Using the definition of a 'labour migrant' provided in the recent South African migration literature as an individual who is absent from the household for at least a month during the year for employment reasons (Posel, 2010), I measured one's access to a migrant network by whether or not an individual was a member of a household containing a 'labour migrant' in the baseline wave of NIDS (2008). This is in line with the strategies employed in the international migrant network literature. I then took advantage of the longitudinal nature of NIDS by defining a rural-to-urban migrant (my dependent variable) as an individual who is observed moving from a rural area in the baseline wave (2008) to an urban area by Wave 3 (2012).

Following the procedure of Zhao (2003) in her study of internal migration in China, I first estimated a migration choice model in line with previous local studies and then introduced controls to capture the impact of migrant networks on the probability of migrating. The findings of the migration choice model affirm established findings. The probability of migrating tends to increase the more educated one is and the age-migration relationship is hump-shaped, initially increasing with age in the late teens to early twenties and then decreasing. Furthermore, in line with the findings of Posel et al. (2006) and Ardington et al. (2009), women appear to be more likely to migrate if they share a household with a pension-eligible individual. This is possibly due to the fact that residing with a pensioner could not only reduce the credit constraints one may face in deciding to migrate, but also the childcare constraints facing young mothers.

In line with the international migration literature, the empirical results also suggest that sharing a household with somebody with recent experience as a labour migrant increases the likelihood of becoming a migrant oneself (by between 2-3 percentage points). These findings are robust to the inclusion of the individual's prior migration experience and employment status. Furthermore, while it was not the main focus of this analysis, there appears to be evidence that being genetically (and immediately) related to someone with previous migration experience induces some form of household bargaining process such that the impact of being closely related to a labour migrant on one's probability of migration is somewhat unclear. These findings therefore suggest that social capital does indeed seem to play a role along with physical and human capital in determining who migrates in South Africa.

The obvious policy implication of this result is that people who migrate from rural communities to urban centres in South Africa do not necessarily do so in some blind attempt to find work – they have probably received information from their migrant networks regarding possible job opportunities and have made the move with the help of their friends and family. As more and more migrants join the migrant network, migration can become a self-sustaining process and Massey's (1990) theory of the cumulative causation of migration will become a reality. Given the growing populations and already limited supply of housing in South African cities, government should bare these findings strongly in mind, and improve its efforts to supply more accommodation in the urban areas or create employment opportunities in the rural areas to keep people from migrating in the first place.

Due to data limitations, I was only able to measure one's connection to a migrant network via household membership and, to some extent, kinship. In terms of further research, it would be interesting to explore whether the migrant network effect extends to the community level, as this has been found to be the case in other countries (Massey and Espinosa, 1997; Winters et al., 2001; Zhao, 2003). It could also be informative to tease out the role of genetic relatedness even further when more NIDS data becomes available.

7 References

References

[1] Ardington, Cally, Anne Case, and Victoria Hosegood. 2009. "Labor Supply Responses to Large Social Transfers: Longitudinal Evidence from South Africa." *American Economic Journal: Applied Economics* 1 (1): 22–48.

[2] Ardington, Cally, Till Bärnighausen, Anne Case, and Alicia Menendez. 2013. "Social protection and labour market outcomes of youth in South Africa." A Southern Africa Labour and Development Research Unit Working Paper Number 96. Cape Town: SALDRU, University of Cape Town

[3] Banerjee, Biswajit. 1983. "Social Networks in the Migration Process: Empirical Evidence on Chain Migration in India." *The Journal of Developing Areas* 17 (2): 185–96.

[4] Bauer, Thomas, and Klaus F. Zimmermann. 1997. "Network Migration of Ethnic Germans." *International Migration Review* 31 (1): 143–49.

[5] Bertrand, Marianne, Sendhil Mullainathan, and Douglas Miller. 2003. "Public Policy and Extended Families: Evidence from Pensions in South Africa." *The World Bank Economic Review* 17 (1): 27–50.

[6] Borjas, George J. 1985. "Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants." *The Journal of Labour Economics* 3 (4): 463–89.

[7] ———. 1987. "Self-Selection and the Earnings of Immigrants". *The American Economic Review* 77 (4): 531–53.

[8] ———. 1989. "Immigrant and Emigrant Earnings: A Longitudinal Study". *Economic Inquiry* 27: 21–37.

[9] Borjas, George J., Stephen G. Bronars, and Stephen J. Trejo. 1992. "Self-Selection and Internal Migration in the United States." *Journal of Urban Economics* 32 (2): 159–85.

[10] Bowles, Samuel, and Dorrit Posel. 2005. "Genetic Relatedness Predicts South African Migrant Workers' Remittances to their Families." *Nature* 434: 380–83.

[11] Caces, Fe. 1986. "Immigrant Recruitment in to the Labour Force: Social Networks Among Filipinos in Hawaii." *Amerasia Journal* 13 (1): 23–28.

[12] Clarke, Rowan, and Katherine Eyal. 2013. "Microeconomic Determinants of Spatial Mobility in Post-Apartheid South Africa: Longitudinal Evidence from the National Income Dynamics Study." *Development Southern Africa* 31 (1): 168–94.

[13] Constant, Amelie F., and Douglas S. Massey. 2003. "Self-Selection, Earnings, and Out-Migration: A Longitudinal Study of Immigrants to Germany." IZA Discussion Paper Series, No. 672.

[14] De Haas, Hein. 2010. "Migration and Development: A Theoretical Perspective." *International Migration Review* 44 (1): 227–64.

[15] De Villiers, Louise, Michael Brown, Ingrid Woolard, Reza C. Daniels and Murray Leibbrandt, eds. 2013. "National Income Dynamics Study Wave 3 User Manual", Cape Town: Southern Africa Labour and Development Research Unit.

[16] Dolfin, Sarah, and Garance Genicot. 2010. "What Do Networks Do? The Role of Networks on Migration and 'Coyote' Use." *Review of Development Economics* 14 (2): 343–59.

[17] Gelderblom, Derik, and Pieter Kok. 1994. *Urbanisation: South Africa's Challenge Volume 1.* Pretoria: HSRC Publishers.

[18] Gelderblom, Derik, and Whynie Adams. 2006. "The limits and possibilities of migrant networks." In P. Kok, D. Gelderblom, J. O. Oucho and J. van Zyl (Eds) *Migration in South and Southern Africa: Dynamics and Determinants*. Cape Town: HSRC Press, pp. 227–48.

[19] Gubhaju, Bina, and Gordon F. De Jong. 2009. "Individual versus Household Migration Decision Rules: Gender and Marital Status Differences in Intentions to Migrate in South Africa." *International Migration* 47 (1): 31–61.

[20] Kingdon, Geeta and John Knight. 2006. "The measurement of unemployment when unemployment is high." *Labour Economics* 13 (3): 291-315.

[21] Kok, Pieter, Michael O'Donovan, Oumar Bouare, and Johan van Zyl. 2003. *Post-Apartheid Patterns of Internal Migration in South Africa*. Cape Town: HSRC Publishers.

[22] Leibbrandt, Murray, Ingrid Woolard and Louise de Villiers. 2009. "Methodology: Report on NIDS Wave 1: Technical Paper no. 1", Cape Town: Southern Africa Labour and Development Research Unit.

[23] Lloyd, Neil, and Murray Leibbrandt. 2014. "New evidence on subjective well-being and the definition of unemployment in South Africa" *Development Southern Africa* 31 (1): 85–105.

[24] Lucas, Robert EB. 1997. "Internal Migration in Developing Countries." *Handbook of Population and Family Economics* 1: 721–98.

[25] Lucas, Robert EB, and Oded Stark. 1985. "Motivations to Remit: Evidence from Botswana." *The Journal of Political Economy* 93 (5): 901–18.

[26] Massey, Douglas S. 1990. "Social Structure, Household Strategies, and the Cumulative Causation of Migration." *Population Index* 56 (1): 3–26.

[27] Massey, Douglas S., Joaquin Arango, Graeme Hugo, Ali Kouaouci, Adela Pellegrino, and J. Edward Taylor. 1993. "Theories of International Migration: A Review and Appraisal." *Population and Development Review* 19 (3): 431–66.

[28] Massey, Douglas S., and Kristin E. Espinosa. 1997. "What's Driving Mexico-U.S. Migration? A Theoretical, Empirical, and Policy Analysis." *American Journal of Sociology* 102 (4): 939–99.

[29] McKenzie, David, and Hillel Rapoport. 2010. "Self-Selection Patterns in Mexico-U.S. Migration: The Role of Migration Networks." *The Review of Economics and Statistics* 92 (4): 811–21.

[30] Posel, Deborah. 1991. "Curbing African Urbanisation in the 1950s and 1960s." In M. Swilling, R. Humphries, and K. Shubane (Eds) *Apartheid City in Transition*. Cape Town: Oxford University Press, pp. 19–32.

[31] Posel, Dorrit. 2001. "How Do Households Work? Migration, the Household and Remittance Behaviour in South Africa." *Social Dynamics* 27 (1): 165–89.

[32] ———. 2004. "Have Migration Patterns in Post-Apartheid South Africa Changed?" *Journal of Interdisciplinary Economics* 15 (3/4): 277–92.

[33] ———. 2010. "Households and Labour Migration in Post-Apartheid South Africa." *Studies in Economics and Econometrics* 34 (3): 129–41.

[34] Posel, Dorrit, and Daniela Casale. 2003. "What Has Been Happening to Internal Labour Migration in South Africa, 1993–1999?" *South African Journal of Economics* 71 (3): 455–79.

[35] Posel, Dorrit, Daniela Casale, and Claire Vermaak. 2014. "Job Search and the Measurement of Unemployment in South Africa" *South African Journal of Economics* 82 (1): 66–80.

[36] Posel, Dorrit, James A. Fairburn, and Frances Lund. 2006. "Labour Migration and Households: A Reconsideration of the Effects of the Social Pension on Labour Supply in South Africa." *Economic Modelling* 23 (5): 836–53.

[37] Ranchhod, Vimal. 2009. "Labour Market: Analysis of the NIDS Wave 1 Dataset: Discussion Paper no. 12", Cape Town: Southern Africa Labour and Development Research Unit Southern Africa Labour and Development Research Unit. 2013. "Wave 1, Update: September 2013: Introduction to NIDS data", Cape Town: Southern Africa Labour and Development Research Unit.

[38] Schiel, Reinhard. 2014. "Migrant Labour in Contemporary South Africa", University of Cape Town Master's Thesis, Unpublished.

[39] Schöer, Volker, and Murray Leibbrandt. 2006. "Determinants of Job Search Strategies: Evidence from the Khayelitsha/Mitchell's Plain Survey" *South African Journal of Economics* 74 (4): 702–724.

[40] Stark, Oded, and David E. Bloom. 1985. "The New Economics of Labor Migration." *The American Economic Review* 75 (2): 173–78.

[41] Stark, Oded, and Robert EB Lucas. 1988. "Migration, Remittances, and the Family." *Economic Development and Cultural Change* 36 (3): 465–81.

[42] Todaro, Michael. 1980. "Internal Migration in Developing Countries: A Survey." In R.A. Easterlin (Ed.) *Population and Economic Change in Developing Countries*. Chicago: University of Chicago Press, pp. 361–402.

[43] Walker, Cherryl. 1990. "Gender and the Development of the Migrant Labour System c.1850-1930: An Overview" In C. Walker (Ed.) *Women and Gender in Southern Africa to 1945.* Cape Town: David Philip Publishers, pp. 168–96.

[44] Wilson, Francis. 2001. "Minerals and Migrants: How the Mining Industry Has Shaped South Africa." *Daedalus* 130 (1): 99–121.

[45] Winters, Paul, Alain De Janvry, and Elisabeth Sadoulet. 2001. "Family and Community Networks in Mexico-US Migration." *Journal of Human Resources* 36 (1): 159–84.

[46] Wittenberg, Martin. 2009. "Weights: Report on NIDS Wave 1: Technical Paper no. 2", Cape Town: Southern Africa Labour and Development Research Unit.

[47] Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

[48] Zhao, Yaohui. 2003. "The Role of Migrant Networks in Labour Migration: The Case of China" *Contemporary Economic Policy* 14 (4): 500–11.

Data:

- Southern Africa Labour and Development Research Unit. National Income Dynamics Study 2008, Wave 1 [dataset]. Version 5.2. Cape Town: Southern Africa Labour and Development Research Unit [producer], 2009. Cape Town: DataFirst [distributor], 2014.
- Southern Africa Labour and Development Research Unit. National Income Dynamics Study 2012, Wave 3 [dataset]. Version 1.2. Cape Town: Southern Africa Labour and Development Research Unit [producer], 2014. Cape Town: DataFirst [distributor], 2014.

Table 1: Baseline wave individual and household characteristics of all adults in Wave 1 (aged 15+) and prime-aged African adults in Wave 3 (aged 18-55) – analysing the impact of attrition between waves

	All adults aged 15+ in Wave 1		Africans aged 18-55 in Wave 3	
Respondent observed in:	Wave 1	Wave 1 and Wave 3	Wave 1	Wave 1 and Wave 3
Individual characteristics:				
Gender				
Female	0.52	0.53	0.51	0.51
Male	0.48	0.47	0.49	0.49
Race				
African	0.77	0.76	-	-
Coloured	0.09	0.09	-	-
Asian/Indian	0.03	0.03	-	-
White	0.11	0.12	-	-
Education				
Years of education	8.87	9.03	9.04	9.10
No schooling	0.09	0.08	0.05	0.05
Primary	0.19	0.19	0.22	0.22
Incomplete secondary	0.41	0.41	0.46	0.46
Matric	0.23	0.24	0.23	0.23
Some tertiary	0.08	0.09	0.05	0.05
Employment status				
Employed	0.44	0.45	0.43	0.43
Unemployed (narrow)	0.14	0.14	0.18	0.17
Unemployed (broad)	0.19	0.19	0.24	0.23
Age	36.24	35.55	28.86	28.66
Household characteristics:				
Household size	4.76	4.75	5.05	5.06
Receives pension income	0.21	0.20	0.17	0.17
Rural	0.38	0.37	0.45	0.45
Monthly household income (rands)				
Income quartile 1	803.66	802.99	783.30	785.53
Income quartile 2	1821.16	1820.08	1810.47	1810.81
Income quartile 3	3296.72	3295.86	3295.80	3293.11
Income quartile 4	15597.50	15716.60	11104.87	11035.28
Migrant network variables:				
Household contains a labour migrant	0.11	0.11	0.13	0.13
Household receives remittance income	0.15	0.15	0.16	0.16
Household contains an employed individual	0.67	0.67	0.66	0.66
Observations:	18603	17215	11854	11261

Notes: Wave 1 (2008) survey weights applied. Own calculations using the NIDS data.

Table 2: Baseline wave individual and household characteristics of migrants and non-migrants in the sample of Africans aged 18-55 in Wave 3

	Migrant[1]	Non-migrant[2]
Individual characteristics:		
Gender		
Female	0.53	0.52
(observations)	(362)	(5461)
Male	0.47	0.48
(observations)	(337)	(4342)
Education		
Years of education	10.05***	9.08
No schooling	0.02	0.05***
Primary	0.11	0.22***
Incomplete secondary	0.55***	0.45
Matric	0.25	0.23
Some tertiary	0.07	0.05
Employment status		
Employed	0.25	0.43***
Unemployed (narrow)	0.17	0.18
Unemployed (broad)	0.24	0.23
Age	23.91	28.98***
Household characteristics:		
Household size	5.84***	5.13
Receives pension income	0.28***	0.17
Rural	1.00***	0.57
Monthly household income (rands)		
Income quartile 1	783.67	788.33
Income quartile 2	1816.58	1814.20
Income quartile 3	3340.46	3290.08
Income quartile 4	18964.14*	10551.05
Migrant network variables:		
Household contains a labour migrant	0.25***	0.13
Labour migrant is kin (parent or sibling)	0.10***	0.04
Household receives remittance income	0.20	0.16
Household contains an employed individual	0.53	0.67***
Observations:	699	9803

Notes: *** indicates that the mean difference between the migrant sample and non-migrant sample is significant at the 1% level; ** indicates significance at the 5% level; and * indicates significance at the 10% level. Wave 1 (2008) survey weights applied. Own calculations using the NIDS data.

[1] Respondent (observed in 2008 and 2012) moves from a rural area to an urban area between survey waves

[2] Respondent (observed in 2008 and 2012) does not move from a rural area to an urban area between survey waves

Table 3: Rural-to-urban migrants in 2012 vs labour migrants in 2008

Rural-to-urban migrant 2012	Labour migrant 2008			**Total**
	No	Yes	Missing	
No	9212	531	60	9803
Yes	623	70	6	699
Missing	725	33	1	759
Total	10560	634	67	11261

Note: Own calculations using the NIDS data.

Table 4: The effect of migrant networks on the probability of migration for Africans aged 18-55 in Wave 3 (2012)

	Dependent variable: Rural-to-urban migrant			
	(1)	(2)	(3)	(4)
	LPM	LPM	LPM	Probit APE
Female	-0.003	-0.004	-0.005	-0.005
	(0.005)	(0.005)	(0.005)	(0.005)
Age	0.112***	0.108***	0.103***	0.063*
	(0.032)	(0.032)	(0.033)	(0.036)
Age^2	-0.006***	-0.005***	-0.005***	-0.003
	(0.002)	(0.002)	(0.002)	(0.002)
Age^3	0.000***	0.000***	0.000***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Age^4	-0.000***	-0.000**	-0.000**	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
No schooling	-0.016	-0.016	-0.020*	-0.022
	(0.010)	(0.010)	(0.010)	(0.016)
Primary	-0.031***	-0.031***	-0.034***	-0.044***
	(0.007)	(0.007)	(0.007)	(0.010)
Incomplete secondary	0.000	0.000	-0.001	-0.004
	(0.006)	(0.006)	(0.006)	(0.006)
Some tertiary	0.043*	0.045*	0.050**	0.050***
	(0.024)	(0.024)	(0.025)	(0.016)
Married	0.007	0.006	0.006	0.004
	(0.007)	(0.007)	(0.007)	(0.007)
Pension eligible household	0.029***	0.026***	0.024**	0.019***
	(0.010)	(0.010)	(0.010)	(0.007)
Number resident household members aged 0-5	0.006*	0.005*	0.006*	0.006**
	(0.003)	(0.003)	(0.003)	(0.003)
Number resident household members aged 6-17	0.009***	0.009***	0.008***	0.008***
	(0.002)	(0.002)	(0.002)	(0.002)
Number resident household members aged 18-55	-0.012***	-0.011***	-0.010***	-0.009***
	(0.002)	(0.002)	(0.002)	(0.002)
Number resident household members aged 56 and over	0.010	0.011*	0.011	0.010*
	(0.007)	(0.007)	(0.007)	(0.005)
Household contains a labour migrant		0.027***	0.027***	0.021***
		(0.008)	(0.008)	(0.006)
Household contains an employed individual			-0.021***	-0.019***
			(0.005)	(0.005)
Household receives remittance income			-0.007	-0.007
			(0.007)	(0.006)
Observations	9341	9315	9287	9287

Note: Robust standard errors in parentheses. Wave 3 panel weights applied. Column (4) depicts the average partial effects (APE) from the probit regression. Own calculations using the NIDS data.
*** p<0.01, ** p<0.05, * p<0.1

Table 5: The effect of migrant networks on the probability of migration for African males aged 18-55 in Wave 3 (2012)

	Dependent variable: Rural-to-urban migrant			
	(1)	(2)	(3)	(4)
	LPM	LPM	LPM	Probit APE
Age	0.075	0.071	0.064	0.018
	(0.046)	(0.046)	(0.046)	(0.043)
Age^2	-0.003	-0.003	-0.003	-0.000
	(0.003)	(0.003)	(0.002)	(0.002)
Age^3	0.000	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Age^4	-0.000	-0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
No schooling	-0.009	-0.009	-0.014	-0.020
	(0.017)	(0.017)	(0.017)	(0.026)
Primary	-0.028***	-0.028***	-0.031***	-0.040***
	(0.010)	(0.010)	(0.010)	(0.013)
Incomplete secondary	0.002	0.003	0.002	-0.000
	(0.009)	(0.009)	(0.009)	(0.008)
Some tertiary	0.061	0.062	0.067	0.061**
	(0.045)	(0.045)	(0.045)	(0.026)
Married	0.012	0.012	0.013	0.009
	(0.010)	(0.010)	(0.010)	(0.009)
Pension eligible household	0.022	0.019	0.016	0.013
	(0.013)	(0.014)	(0.014)	(0.011)
Number resident household members aged 0-5	0.002	0.002	0.003	0.003
	(0.004)	(0.004)	(0.004)	(0.004)
Number resident household members aged 6-17	0.013***	0.012***	0.012***	0.010***
	(0.003)	(0.003)	(0.003)	(0.003)
Number resident household members aged 18-55	-0.012***	-0.011***	-0.009***	-0.008***
	(0.003)	(0.003)	(0.003)	(0.003)
Number resident household members aged 56 and over	0.015	0.016	0.014	0.011
	(0.010)	(0.010)	(0.010)	(0.008)
Household contains a labour migrant		0.028**	0.028**	0.021***
		(0.013)	(0.013)	(0.008)
Household contains an employed individual			-0.034***	-0.031***
			(0.008)	(0.007)
Household receives remittance income			-0.006	-0.007
			(0.010)	(0.008)
Observations	4064	4052	4040	4040

Note: Robust standard errors in parentheses. Wave 3 panel weights applied. Column (4) depicts the average partial effects (APE) from the probit regression. Own calculations using the NIDS data.
*** p<0.01, ** p<0.05, * p<0.1

Table 6: The effect of migrant networks on the probability of migration for African females aged 18-55 in Wave 3 (2012)

	Dependent variable: Rural-to-urban migrant			
	(1)	(2)	(3)	(4)
	LPM	LPM	LPM	Probit APE
Age	0.148***	0.143***	0.140***	0.109**
	(0.043)	(0.043)	(0.044)	(0.053)
Age^2	-0.008***	-0.007***	-0.007***	-0.006**
	(0.002)	(0.002)	(0.002)	(0.003)
Age^3	0.000***	0.000***	0.000***	0.000*
	(0.000)	(0.000)	(0.000)	(0.000)
Age^4	-0.000***	-0.000***	-0.000***	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)
No schooling	-0.020*	-0.021*	-0.023**	-0.027
	(0.011)	(0.011)	(0.012)	(0.018)
Primary	-0.034***	-0.033***	-0.036***	-0.047***
	(0.010)	(0.010)	(0.011)	(0.014)
Incomplete secondary	-0.002	-0.002	-0.003	-0.006
	(0.009)	(0.009)	(0.009)	(0.008)
Some tertiary	0.031	0.032	0.036	0.039*
	(0.027)	(0.027)	(0.027)	(0.020)
Married	0.001	-0.000	-0.000	-0.003
	(0.009)	(0.009)	(0.009)	(0.009)
Pension eligible household	0.035***	0.031**	0.030**	0.024**
	(0.013)	(0.013)	(0.014)	(0.010)
Number resident household members aged 0-5	0.008*	0.007*	0.008*	0.008**
	(0.004)	(0.004)	(0.004)	(0.004)
Number resident household members aged 6-17	0.006**	0.006*	0.006*	0.006**
	(0.003)	(0.003)	(0.003)	(0.003)
Number resident household members aged 18-55	-0.012***	-0.011***	-0.010***	-0.010***
	(0.002)	(0.002)	(0.003)	(0.003)
Number resident household members aged 56 and over	0.007	0.008	0.009	0.009
	(0.009)	(0.009)	(0.009)	(0.008)
Household contains a labour migrant		0.026**	0.026**	0.020**
		(0.011)	(0.011)	(0.008)
Household contains an employed individual			-0.010	-0.009
			(0.007)	(0.007)
Household receives remittance income			-0.008	-0.006
			(0.009)	(0.008)
Observations	5277	5263	5247	5247

Note: Robust standard errors in parentheses. Wave 3 panel weights applied. Column (4) depicts the average partial effects (APE) from the probit regression. Own calculations using the NIDS data.
*** p<0.01, ** p<0.05, * p<0.1

Table 7: The effect of migrant networks on the probability of migration for Africans aged 18-55 in Wave 3 (2012) – controlling for labour migrant status

	Dependent variable: Rural-to-urban migrant			
	(1)	(2)	(3)	(4)
	LPM	LPM	LPM	Probit APE
Female	-0.003	-0.004	-0.005	-0.005
	(0.005)	(0.005)	(0.005)	(0.005)
Age	0.112***	0.111***	0.106***	0.066*
	(0.032)	(0.032)	(0.033)	(0.036)
Age^2	-0.006***	-0.006***	-0.005***	-0.003
	(0.002)	(0.002)	(0.002)	(0.002)
Age^3	0.000***	0.000***	0.000***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Age^4	-0.000***	-0.000**	-0.000**	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
No schooling	-0.016	-0.016	-0.020*	-0.022
	(0.010)	(0.010)	(0.010)	(0.016)
Primary	-0.031***	-0.031***	-0.034***	-0.045***
	(0.007)	(0.007)	(0.007)	(0.010)
Incomplete secondary	0.000	-0.000	-0.002	-0.004
	(0.006)	(0.006)	(0.006)	(0.006)
Some tertiary	0.043*	0.045*	0.050**	0.050***
	(0.024)	(0.024)	(0.025)	(0.016)
Married	0.007	0.007	0.007	0.004
	(0.007)	(0.007)	(0.007)	(0.007)
Pension eligible household	0.029***	0.025**	0.022**	0.018**
	(0.010)	(0.010)	(0.010)	(0.007)
Number resident household members aged 0-5	0.006*	0.005	0.006*	0.006**
	(0.003)	(0.003)	(0.003)	(0.003)
Number resident household members aged 6-17	0.009***	0.009***	0.008***	0.008***
	(0.002)	(0.002)	(0.002)	(0.002)
Number resident household members aged 18-55	-0.012***	-0.010***	-0.010***	-0.009***
	(0.002)	(0.002)	(0.002)	(0.002)
Number resident household members aged 56 and over	0.010	0.012*	0.012*	0.011**
	(0.007)	(0.007)	(0.007)	(0.005)
Household contains a labour migrant (excl. self)		**0.028*** **	**0.028*** **	**0.021*** **
		(0.009)	(0.009)	(0.006)
Labour migrant		**0.017**	**0.017**	**0.012**
		(0.014)	(0.014)	(0.009)
Household contains an employed individual			-0.021***	-0.019***
			(0.005)	(0.005)
Household receives remittance income			-0.008	-0.008
			(0.007)	(0.006)
Observations	9341	9283	9255	9255

Note: Robust standard errors in parentheses. Wave 3 panel weights applied. Column (4) depicts the average partial effects (APE) from the probit regression. Own calculations using the NIDS data.
*** p<0.01, ** p<0.05, * p<0.1

Table 8: The effect of migrant networks on the probability of migration for Africans aged 18-55 in Wave 3 (2012) – controlling for employment status

	Dependent variable: Rural-to-urban migrant			
	(1)	(2)	(3)	(4)
	LPM	LPM	LPM	Probit APE
Female	-0.003	-0.007	-0.007	-0.007
	(0.005)	(0.006)	(0.006)	(0.006)
Age	0.112***	0.094**	0.094**	0.073
	(0.032)	(0.046)	(0.046)	(0.048)
Age^2	-0.006***	-0.005**	-0.005**	-0.004
	(0.002)	(0.002)	(0.002)	(0.003)
Age^3	0.000***	0.000*	0.000*	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Age^4	-0.000***	-0.000*	-0.000*	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
No schooling	-0.016	-0.015	-0.016	-0.019
	(0.010)	(0.012)	(0.012)	(0.018)
Primary	-0.031***	-0.028***	-0.029***	-0.037***
	(0.007)	(0.008)	(0.008)	(0.011)
Incomplete secondary	0.000	-0.004	-0.005	-0.006
	(0.006)	(0.007)	(0.007)	(0.007)
Some tertiary	0.043*	0.056**	0.056**	0.057***
	(0.024)	(0.027)	(0.027)	(0.018)
Married	0.007	0.010	0.010	0.007
	(0.007)	(0.008)	(0.008)	(0.008)
Pension eligible household	0.029***	0.021*	0.020*	0.016*
	(0.010)	(0.012)	(0.012)	(0.009)
Number resident household members aged 0-5	0.006*	0.007**	0.008**	0.007**
	(0.003)	(0.004)	(0.004)	(0.003)
Number resident household members aged 6-17	0.009***	0.007**	0.007**	0.006***
	(0.002)	(0.003)	(0.003)	(0.002)
Number resident household members aged 18-55	-0.012***	-0.012***	-0.011***	-0.010***
	(0.002)	(0.002)	(0.002)	(0.003)
Number resident household members aged 56 and over	0.010	0.012	0.012	0.011*
	(0.007)	(0.008)	(0.008)	(0.006)
Household contains a labour migrant		0.024**	0.025***	0.019***
		(0.010)	(0.010)	(0.007)
Employed		**-0.015*** **	**-0.015*** **	**-0.017** **
		(0.006)	(0.006)	(0.007)
Household contains an employed individual (excl. self)			**-0.007**	**-0.007**
			(0.007)	(0.007)
Household receives remittance income			-0.009	-0.009
			(0.008)	(0.007)
Observations	9341	7461	7461	7461

Note: Robust standard errors in parentheses. Wave 3 panel weights applied. Column (4) depicts the average partial effects (APE) from the probit regression. Own calculations using the NIDS data.
*** p<0.01, ** p<0.05, * p<0.1

Table 9: The effect of migrant networks on the probability of migration for Africans aged 18-55 in Wave 3 (2012) – controlling for genetic relatedness

	Dependent variable: Rural-to-urban migrant		
	(1)	(2)	(3)
	LPM	LPM	LPM
Female	-0.005	-0.005	-0.005
	(0.005)	(0.005)	(0.005)
Age	0.103***	0.103***	0.103***
	(0.033)	(0.033)	(0.033)
Age^2	-0.005***	-0.005***	-0.005***
	(0.002)	(0.002)	(0.002)
Age^3	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)
Age^4	-0.000**	-0.000**	-0.000**
	(0.000)	(0.000)	(0.000)
No schooling	-0.020*	-0.019*	-0.019*
	(0.010)	(0.010)	(0.010)
Primary	-0.034***	-0.034***	-0.034***
	(0.007)	(0.007)	(0.007)
Incomplete secondary	-0.001	-0.002	-0.002
	(0.006)	(0.006)	(0.006)
Some tertiary	0.050**	0.050**	0.050**
	(0.025)	(0.025)	(0.025)
Married	0.006	0.006	0.006
	(0.007)	(0.007)	(0.007)
Pension eligible household	0.024**	0.023**	0.023**
	(0.010)	(0.010)	(0.010)
Number resident household members aged 0-5	0.006*	0.006*	0.006*
	(0.003)	(0.003)	(0.003)
Number resident household members aged 6-17	0.008***	0.008***	0.008***
	(0.002)	(0.002)	(0.002)
Number resident household members aged 18-55	-0.010***	-0.010***	-0.010***
	(0.002)	(0.002)	(0.002)
Number resident household members aged 56 and over	0.011	0.011	0.011
	(0.007)	(0.007)	(0.007)
Household contains a labour migrant	**0.027*** **	**0.020** **	**0.020** **
	(0.008)	(0.009)	(0.009)
Household contains an employed individual	-0.021***	-0.022***	-0.022***
	(0.005)	(0.005)	(0.005)
Household receives remittance income	-0.007	-0.007	-0.007
	(0.007)	(0.007)	(0.007)
Labour migrant is kin (parent/sibling)		**0.024**	**-0.064*** **
		(0.017)	(0.008)
(Labour migrant is kin) x (Household contains a labour migrant)			**0.088*** **
			(0.018)
Observations	9287	9287	9287

Note: Robust standard errors in parentheses. Wave 3 panel weights applied. Own calculations using the NIDS data.
*** p<0.01, ** p<0.05, * p<0.1

Figure 1: Proportion of migrants by completed years of education in 2008

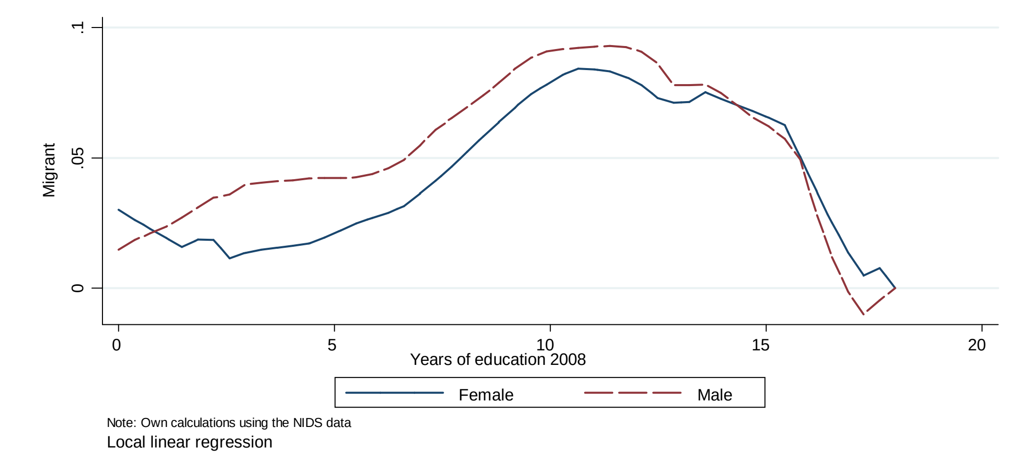
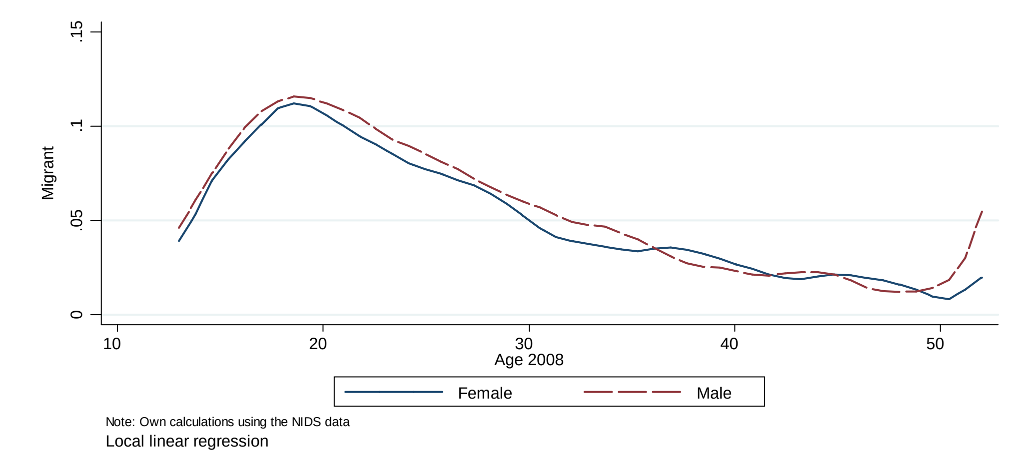


Figure 2: Proportion of migrants by age in 2008



Appendix

Table 10: Baseline wave individual and household characteristics of migrants and non-migrants in the sample of Africans aged 18-55 in Wave 3 – males vs females

	Male		Female	
	Migrant[1]	Non-migrant[2]	Migrant[1]	Non-migrant[2]
Individual characteristics:				
Education				
Years of education	9.75***	8.94	10.32***	9.20
No schooling	0.02	0.05**	0.02	0.05***
Primary	0.14	0.24***	0.08	0.20***
Incomplete secondary	0.53***	0.43	0.57***	0.48
Matric	0.24	0.24	0.26	0.22
Some tertiary	0.07	0.04	0.07	0.05
Employment status				
Employed	0.28	0.54***	0.22	0.35***
Unemployed (narrow)	0.17	0.14	0.17	0.21
Unemployed (broad)	0.20	0.17	0.26	0.28
Age	23.86	28.50***	23.96	29.42***
Household characteristics:				
Household size	5.53***	4.74	6.10***	5.48
Receives pension income	0.27***	0.16	0.29***	0.18
Monthly household income (rands)				
Income quartile 1	777.10	789.67	789.61	787.32
Income quartile 2	1801.19	1821.47	1827.97	1807.50
Income quartile 3	3323.03	3312.07	3359.57	3267.57
Income quartile 4	15172.22	10751.77	22416.99*	10347.23
Migrant network variables:				
Household contains a labour migrant	0.24***	0.11	0.27***	0.15
Labour migrant is kin (parent or sibling)	0.11***	0.04	0.10***	0.04
Household receives remittance income	0.19*	0.15	0.20	0.17
Household contains an employed individual	0.52	0.71***	0.54	0.63***
Observations:	337	4342	362	5461

Notes: *** indicates that the mean difference between the migrant sample and non-migrant sample is significant at the 1% level; ** indicates significance at the 5% level; and * indicates significance at the 10% level. Wave 1 (2008) survey weights applied. Own calculations using the NIDS data.

[1] Respondent (observed in 2008 and 2012) moves from a rural area to an urban area between survey waves

[2] Respondent (observed in 2008 and 2012) does not move from a rural area to an urban area between survey waves