Determinants of Wealth in South Africa

Submitted in partial fulfillment of the degree of Masters of Commerce in Economics, by coursework and dissertation

Taryn Augustine
AGSTAR001

UNIVERSITY OF CAPE TOWN
SCHOOL OF ECONOMICS

ECO5023W

Title: Determinants of Wealth in South Africa

Student Name & Surname: Taryn Augustine

Student Number: AGSTAR001

Declaration

1. I know that plagiarism is wrong. Plagiarism is to use another's work and pretend that it is one's own.
2. This paper is my own work.
3. I have used the Harvard convention for citation and referencing. Each significant contribution to, and quotation in, this paper from the work(s) of other people has been attributed, and has been cited and referenced.
4. I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work.

Signature: Signed by candidate ……………………………

Date: ……15/08/2015………..…..

Title: Determinants of Wealth in South Africa

Supervisor: Reza Daniels

Abstract:

This paper investigates the determinants of household wealth in South Africa, using the National Income Dynamics Study (NIDS) Wave 2. In particular, we look at the effects of the wealth-age profile and other household demographic variables. The hump-shaped profile of the wealth and age relationship suggested by the life-cycle hypothesis is not present in the data, although there are indications of its presence in the upper quantiles of the wealth distribution. The South African wealth distribution does not conform to the Lifecycle Hypothesis at this point in time. The LCH model appears to apply only to particular quantiles of the population, that is, the wealthier households and the particularly indebted households. In particular, the results found these to be households with younger heads, which align with LCH predictions. Poorer households, or those whose assets and liabilities are approximately equal do not appear to accumulate wealth in the same manner as their upper and lower quantile counterparts. However, we cannot formally identify the LCH econometrically at a particular quantile. We found evidence of different wealth accumulation behaviour in Tribal Authority Areas, where a dual land tenure ownership structure is in place. This has important implications for researchers interested in components of wealth, such as income, saving, assets and liabilities.

Acknowledgements

I am extremely grateful to Reza Daniels, for his tireless efforts and guidance with my dissertation.

I. Introduction

Modern societies tend to exhibit distinct patterns of capital accumulation, with varying levels of persistence according to the characteristics of their economies. However, a significant proportion of research has been focused on income, rather than wealth, which is remarkable given the importance of wealth as a determinant of the consumer's consumption possibilities. The emphasis on income is partly due to the fact that it is easier to measure and more frequently measured in most countries. Furthermore, the availability of wealth data is not as widespread as income data owing to its social sensitivity. Measurements of wealth require accurate valuations of all assets and liabilities, most of which are not easily obtained. For example, the price of assets, such as jewellery or share prices may fluctuate considerably during a period of time and in some cases, the value of the asset can only be determined upon the sale of said asset (Aron & Muellbauer, 2006). Property and assets may also be held in trusts, which further exacerbate the issue of measurement. It must be noted that wealth is significantly different from income, in that it is possible for an individual/ household to earn large income, but exhibit low net worth due to credit - dependent lifestyles (Aron & Muellbauer, 2006).

Wealth is especially important during phases of economic insecurity and in terms of intergenerational transfers. It allows for consumption smoothing over time, particularly in periods when consumption is expected to be high or income is expected to be low. Consumption smoothing is also important when households face capital market imperfections or borrowing constraints (Davies & Shorrocks, 1999). The behaviour of households has important implications for global growth prospects. More importantly, with the growing importance of capital and savings in economies, wealth is crucial to the well being of consumers. There are substantial literature on wealth and the macroeconomic modeling of its distribution in the developed world (Di Matteo, 2010). However, not much research has been done on this in South Africa (SA), especially in the microeconomic context, due to the unavailability of household data on components of wealth.

South Africa's capital accumulation landscape is particularly unique, in that it is one of the most inequitable societies in the world, with a large informal sector, alongside an advanced formal sector (Aron & Muellbauer, 2006). During the Apartheid era and post democracy, SA has undergone international trade and financial sanctions, foreign disinvestment, capital flight and emigration (Keegan, 1986). Financial liberalization in the markets has brought about more competitive markets, rather than bringing financial services to poor African households (Aron & Muellbauer, 2006). Previously, the white minority was afforded education and economic opportunities, while the black population was dispossessed (Keegan, 1986). The effects of this discrimination are still present throughout the economy and the lack of access to credit and assets for a large proportion of the population severely stunts the households sector's ability to amass wealth (Aron & Muellbauer, 2006).

The Apartheid regime systematically and purposefully restricted the majority of South Africans (Africans, Coloureds and Asians) from any meaningful participation in the economy (DTI, nd). The assets of those discriminated against were both directly and indirectly destroyed. This resulted in a biased accumulation process, which confined the creation of wealth to a racial minority (van der Berg, 2011). Furthermore it imposed underdevelopment on the disadvantaged populace, ensured the supply of cheap labour and undermined self-employment and entrepreneurship opportunities (DTI, nd). This systematic disempowerment confined the majority of African people to homeland areas. These areas were the most deprived in terms of living standards, infrastructure and business environment (McDonald & Piesse, 1999). These racially enforced migrations uprooted miliions from their places of residence and lead to large capital forfeitures. Consequently, the drastic restriction of property ownership rights of black individuals rendered it impossible to acquire assets that could serve as collateral for loan financing (van der Berg, 2011). The result is an economic structure, which still excludes the vast majority of South Africans from accumulating wealth.

Investigations into the wealth distribution reveal the pattern of wealth accumulating behaviour, as well as the type of economy in which people and households operate and the type of society in which they reside (Davies & Shorrocks, 1999; Boskin 1991). This paper seeks to investigate South African wealth data found in the National Income Dynamics Study (NIDS) Wave 2, as it is the only household survey to hold data of this nature in South Africa. In particular, the pattern of wealth accumulation in tribal authority areas (TAA) is of interest and whether we can identify significantly different behaviour for households in these areas, in comparison to other South African households throughout the wealth distribution. We are particularly interested in TAAs, as we believe that there are unidentified specialized wealth patterns occurring in these areas. These patterns are as a result of a dual land tenure system that was originally associated with the Bantustans of the Apartheid era, but which subsequent to the transition to democracy, has remained in place through special legislation guaranteeing traditional tribal chieftainships control over all access rights to land in these TAAs. In particular, the land in the area is communally owned and as such, the reports of wealth holdings are believed to be inflated.

We find that the data does not appear to conform to the lifecycle hypothesis (LCH) and we identify sign similarities in the 5^{th} and 95^{th} quantile. Within the estimation, the tribal authority area coefficient is significant and indicates the presence of different wealth accumulation behaviour. The paper is structured as follows, the first being a literature review on previous work on wealth. Thereafter follows the theoretical approach to the investigation and the methodology. Lastly, we discuss the empirical results and summarise the findings.

II. Background

The focus of wealth research has shifted over the years from overall distribution characteristics to the causes of differences in wealth holdings. Pension savings and the growth in availability and sophistication of micro datasets have been instrumental to this shift. There is a distinction between models of wealth accumulation owing to lifecycle factors and those that concentrate on the intergenerational connections between households. Previous work on the subject matter has rarely covered both models simultaneously (Davies & Shorrocks, 1999). Wealth, in this context, is defined as the value of all the assets owned by a household less the value of the all its liabilities at a particular point in time, also known as net worth.

Davies & Shorrocks (1999) consider five stylised facts regarding wealth, namely:

1) Wealth is distributed more unequally than labour income, total income or consumption expenditure. This is true regardless of the inequality indicator used.

2) Financial assets are less equally distributed than non-financial assets, but only when real estate is the major component of the non-financial assets. The reverse is true in countries where land value is especially important.

3) Similarly, the distribution of inherited wealth is more unequal than wealth in general.

4) Across all age groups, there is generally a set of individuals and households with extremely low net worth. Furthermore, the level of financial assets appears to be low in a number of countries, including the US.

5) Wealth inequality has trended downwards in the 20th century, but there have been some interruptions to this trend and reversals.

There have been various theories formulated to account for the attributes of the wealth distribution; in particular, we have the Lifecycle Accumulation theory and Intergenerational Connections (inheritance or bequest motives) (Skopek et al, 2014). Inheritance plays an important role throughout the wealth distribution; it causes major differences in wealth levels and is an important transmission channel of wealth status between generations. However, it plays no role in the intra-generational models. There are also noneconomic reasons for accumulating wealth and these are factors such as power or status, which may be linked to certain types of assets (Skopek et al, 2014).

For many middle-income households, wealth accumulation follows a pattern over the life cycle; it is low among young adults and increases during middle age. It finally reaches a peak just before retirement and then diminishes in the post-retirement years (Modigliani & Brumberg, 1954). The Lifecycle Hypothesis (LCH) model is one of intra-generational accumulation theorised by Modigliani and Brumberg (1954) (Modigliani & Brumberg, 1954). The model posits that households accumulate savings during their lifetime as a result of labour market participation and once they retire, they begin to draw from that wealth/savings. Furthermore consumers derive utility from the consumption stream over the life cycle and the limit to the consumers' choice is essentially his/her lifetime budget constraint (Modigliani & Brumberg, 1954).

The basic assumptions of the model are as follows:

- Consumers are forward looking
- Preferences are defined over the present and future consumption
- A period of retirement is expected at the end of one's lifetime
- No uncertainty
- All consumers face the same constant rate of return r
- Have same length of life T
- No bequest motive

The extensions to this model account for household interests in the consumption of offspring and the size of planned bequests. Further models include capital market imperfections and borrowing constraints or uncertainty in earnings. However, these are not detailed here. The consumer's problem is to maximize

$$Max\ U = U(C_1, \ldots\ldots, C_t)$$

Subject to

$$C^L = \sum_{t=1}^{T} \frac{C_t}{(1+r)^{t-1}} \leq \sum_{t=1}^{R} \frac{E_t}{(1+r)^{t-1}} = E^L$$

where **C**^{L} is lifetime consumption, **E**^{L} is lifetime earnings, **R** is Retirement date and for simplification purposes, we ignore leisure. Assuming a non-working period at the end of life, restrictions on the functional form of utility will ensure that saving occurs. The household smooths consumption by equalising their discounted marginal utility to expenditure in both the present and future periods. In general, labour income is expected to rise at a swift pace in the early years and thus much negative net worth or substantial net borrowing is expected to be prevalent among younger households (where the household head is young) (Modigliani & Brumberg, 1954). The LCH theory is able to deal with systematic variations in income occurring over the lifecycle, as a result of the assumption of finitely lived households (Modigliani, 1986).

Furthermore, the propensity to consume increases with age, while it is lowers for households whose preferences generate a higher desired growth of consumption (Modigliani, 1986). The two foremost implications of this model are that a humped shaped age-profile of wealth is expected, with the peak at or close to retirement age of the household head and that substantial wealth inequality can arise between households, even if all other household characteristics remain the same, other than age (Modigliani, 1986). In other words, age differences alone are expected to have a large impact on the distribution of wealth. It must be noted that observed wealth inequality could arise from the fact that households are sampled at different points in their lifetimes (Modigliani, 1986).

In contrast to the LCH model, the intergenerational connection model analyses the evolution of the wealth distribution over a number of generations and accounts for demographic factors, such as patterns of fertility, marriage, and other economic factors (estate division, taxation etc.). The simplest case is one of a society in which all individuals marry, produce offspring and bestow bequests upon their offspring equal to the amount, which they themselves inherited in the previous period. Thus, the current distribution of wealth depends on the wealth distribution of the previous generation, as well as the patterns of marriage, fertility, estate division and taxation (Davies & Shorrocks, 1999).

There are also models of intergenerational transfer, which are based on a particular specification of parental preferences, in particular altruism. The model utilises the Beckerian approach, which assumes that parental preferences are dependent on the lifetime consumption of parents and the lifetime consumption of each child. Under altruism, parents plan to completely equalise their children's incomes (net of transfers) (Davies & Shorrocks, 1999).

More recently, wealth is enjoying a renewal in research interest, especially with the release of Thomas Picketty's book 'Capital in the 21st Century'. In Picketty's work on capital, he argues that the rate of return on capital tends to be higher than the economy's rate of growth. This is a contradiction of capitalism and previous literature on wealth. This indicates that wealth accumulated in the past grows faster than output and wages (Milanovic, 2014). Since capital ownership tends to be concentrated globally, we have growing income and wealth inequality (Milanovic, 2014). This is the pernicious effect of capital on the distribution of income and wealth. In particular, this effect stems from inherited wealth. As fortunes grow with each generation, the gap between the wealthy and the poor continues to grow. The inequality in labour income also adds to this bequest effect and is at odds with the capital accumulation argument (Milanovic, 2014).

In order to combat the widening inequality, Picketty argues for a wealth tax in each country, as well as a global wealth tax. This is meant to bolster income inequality, as he argues that wealth is the source of income inequality. The global tax is required in order to eliminate the possibility of the wealthy moving their assets to countries with more lenient wealth tax schemes (Aspromourgos, 2014). The implications of this analysis for South Africa are far-reaching. Contrary to Picketty's work, SA's income inequality had little variation, notwithstanding the social transfers and policies enacted to turn the tide. The wealth tax would have little effect on SA's inequality levels, as research has shown that of the 14 million registered tax payers, only 8.4% earn more then R500 000 per annum and contribute 54.4% of all tax revenue (National Treasury, 2014).

The theories regarding wealth distribution have largely been as a result of the empirical work at the time. Prior to the 60s, data on wealth was obtained predominantly from estate tax and wealth tax records, with other evidence coming from small unrepresentative surveys and other sources (Davies & Shorrocks, 1999). However, in South Africa, the lack of wealth data has remained an issue well into the 20th century. Previously, most empirical work has been based on data from the South African Reserve Bank (SARB), who collect macroeconomic data on income, expenditure, savings and debt (Daniels et al, 2014).

The South African legacy is one filled with conquest, settlement, colonisation and, more recently, segregation under Apartheid rule. The Natives Land Act, Act no. 27 of 1913 prevented Africans from freely buying land in certain provinces in South Africa. Colonialism, along with Apartheid, allowed for the development of pluralistic forms of property rights and, in turn, has resulted in challenges for land tenure reform (Home, 2013). The problem was not only the unequal land distributions, but also the insecurity of land tenure, that is, the inability to impose land rights against the claims of others and the inability to profit from the capital and labour invested in the land (Ramutsindela, 2012). The customary land tenure did not confer ownership or the creation of private property rights upon the 'natives' who occupied the land and as such created two parallel land tenure systems. These dual systems were reinforced by the implementation of separate administrative arrangements and separate land development policies (Home, 2013).

The nationalist government instituted racially based land reforms and consolidated the native reserves into ten self-governing homelands. Each homeland had its own assemblies, government departments and rights to confer citizenship (Hoeks et al, 2014). Prior to 1948, thousands of Africans owned land in South Africa. Forced removals during apartheid, however, lead to severe land losses in favour of Whites (Feinberg, 1995). The government instituted policies, which systematically discriminated on the basis of race. These policies in turn restricted the income generating opportunities open to non-whites and created a legacy of inequality that has been passed down through generations (Nattrass & Seekings, 2001). After the abolishment of Apartheid, the constitution required the government to undertake legislative measures to enable access to land on an equitable basis and restore tenure security to those previously disadvantaged citizens.

Since that time, Apartheid laws governing land ownership have been repealed. However, land in the former Bantustans is still subject to the chiefs and tribal authorities and continues to be registered in the name of the State (Adams et al, 1999). As such, there had been longstanding disputes between government and local leaders with regards to the ownership and control of the land. The traditional leaders argue that the state initiatives undermine the pre-existing land rights, whilst the local government argues that the tribal authorities hinder development (Cousins, 2007). In traditional authority areas, all members of the tribe jointly own the land. The leader of the tribe can allocate access to the land for both tribe members and non-tribe members, although tribe members enjoy preference with regards to non-tribe members (Adams et al, 1999).

The 1994 Land Rights Act (amended in 2008), along with the 1996 Land Reform (Labour Tenants) Act and the 1997 Extension of Security Tenure Act were aimed at restoring the property rights of persons and communities dispossessed of property owing to past racial discriminatory laws. These Acts allowed for the establishment of the Commission on Restitution of Land Rights and a Lands Claims Court, in order to process restitution claims, as well as the protection of rural groups with insecure tenure arrangements (Hoeks et al, 2014). The 1997 Settlement Land Acquisition Grant created a grant that allowed for the purchase of land, tenure rights enhancement, investment in infrastructure, home improvement and farm capital. The 1997 White paper on the topic was aimed at creating a land reform strategy, with three main considerations, namely, land restitution, land redistribution and tenure reform (Hoeks et al, 2014).

Following the recommendations laid out in the aforementioned White paper, the 2001 Land Redistribution and Agricultural Development Programme instituted a sliding scale basis grant, which matched applicants' own contributions and concentrated on the transfer of agricultural land to distinct individuals and the improvement of access to municipal and tribal land for grazing purposes. Furthermore, the 2004 communal Land Rights Act attempted to combine customary land tenure practices and titling, by entrusting ownership of the land in Tribal Authority Areas to Traditional Councils. However, this Act was declared unconstitutional in 2010, as it seemed to yield more power to the Tribal Authorities, rather than individuals in the community (Ramutsindela, 2012; Hoeks et al, 2014). Finally, the 2008 Provision of Land Assistance Amendment Act sought to control the provision of financial assistance for the acquisition or improvement of land and tenure rights (Hoeks et al, 2014).

Land reform progress has been limited, as most of the government's efforts had been directed at land restitution and redistribution. This inadequate progress has been attributed to the complexity of the issues, the dispute between Traditional councils and government on both administrative and physical jurisdiction, as well as the problem of 'elite' individuals, who are able to manipulate the tilting process to their advantage (Hoeks et al, 2014). In Fay's (2009) work on land tenure in SA, he argues the importance

of recognising community driven processes of land reform, rather than state interventions, as the tribal communities appear to have resisted and largely ignored these efforts (Fay, 2009).
Aron & Muellbauer (2006), estimated household sector wealth for South Africa, utilising published data on financial flows and other capital market data (Aron & Muellbauer, 2006). They created the main elements of the household sector balance sheet in order to understand the evolution of aggregate consumer spending and saving (Aron & Muellbauer, 2006). Their work revealed that pension wealth had increased significantly over the years, especially in comparison to housing wealth and increasing debt. They associated part of the increase to financial liberalization in the economy and the influence of changing tax policies (Aron & Muellbauer, 2006). Directly held securities were shown to have a downward trend and cyclical behaviour driven by fluxes in real share prices. Lastly, the household debt ratio had risen between the early 1980s and the late 1990s (Aron & Muellbauer, 2006). Furthermore, the paper argued that the estimates were limited by the lack of data in two important areas, namely household sector ownership of foreign assets and the assets of unincorporated businesses (Aron & Muellbauer, 2006).

In 2013 in 'Wealth, Credit Conditions and Consumption: Evidence from South Africa', the same authors constructed a data intensive exploration of wealth and collateral effects (Aron & Muellbauer, 2013). They argued that SA's expenditure was largely generated by households employed in the formal sector, which is striking considering the large size of the informal sector in the economy (Aron & Muellbauer, 2013). Utilising their constructed measures of marginal propensities to consume they found credit liberalisaton to be the important factor. The main findings of the paper were that a conventional lifecycle view of consumption for SA was not applicable due to the structure of the credit market. Furthermore, credit market liberalization increases the average propensity to consume and interaction effects are present between credit market liberalization, expected income growth and housing wealth (Aron & Muellbauer, 2013). Finally, the findings indicated that housing wealth effects significantly surpasses stock market wealth effects in South Africa (Aron & Muellbauer, 2013).

Zimmer & Das (2014) studied how household wealth varies across household composition in Sub-Saharan Africa (Zimmer & Das, 2014). The authors hypothesized that household composition plays an essential role in household wealth. However, the unique characteristics of Sub-Saharan African countries have rendered traditional systems of within household wealth distribution less dependable. These characteristics are issues such as increasing poverty, labour-based migration and the HIV/AIDS epidemic (Zimmer & Das, 2014). Contrasting all household compositions with those containing an older person, the results indicated that households with older persons present have less wealth than other household types (Zimmer & Das, 2014). Furthermore, larger households, households with younger heads, those with male heads and those living in urban areas tend to accumulate less wealth in comparison to other household types (Zimmer & Das, 2014).

In Ngwenya & Paas (2012) the authors attempt to assess ownership of 16 financial products by households in different lifecycle stages amongst four ethnic groups (Africans, Coloureds, Asians, and Whites) in South Africa (Ngwenya & Paas, 2012). The paper discusses the applicability of the LCH to South Africa and found that the relevance of the LCH was unclear, particularly within racial groups. They hypothesized that the LCH predictions would not hold for previously disadvantaged South Africans and results indicated that younger households did not incur more debts and that intermediate cycle households did not own more assets (Ngwenya & Paas, 2012). However, the LCH predictions do appear to be consistent for White households across the age cohorts. Furthermore, among the African, Coloured and Asian ethnic groups, the results showed that younger, relatively wealth and educated households tended to own more financial products (Ngwenya & Paas, 2012).

The expectation is that the lifecycle hypothesis will not hold within quantiles and should not do so for any particular reason. Quantile regression seeks to identify effects at particular points, quantiles, as defined by the researcher. Each quantile represents a different proportion of the sample population and these do not necessarily correspond in any way to particular ages or age cohorts. We cannot test the LCH predictions at a particular quantile point, only across the entire distribution. Thus, a LCH prediction 'violation' within a quantile does not nullify the use of the LCH theory in our estimation.

For our purposes, the 5th quantile represents those households with negative net worth. We argue that they may be high-income households whom are extremely indebted in relation to their assets. The 50th quantile represents those with approximately zero net worth, that is, the poor and the 95th quantile represent those households with high net worth values. We anticipate that violations of the theory presented will occur in the 50th quantile, where households are unable to accumulate wealth in the same manner as households with similar aged heads in different quantiles.

III. Methodology

i. Data

The National Income Dynamics Study is the only nationally representative household survey to include questions on household wealth; it follows households over time and can be used in both the cross sectional context (single wave) and panel context (merged waves). Wealth questions, however, were only examined in Wave 2, and more recently Wave 4, which has yet to be released to the public. Thus, the data is utilised in its cross sectional context.

The survey examines a large number of detailed questions about different assets and liabilities, which allowed for the construction of each component of the total net worth of the household and individuals. Observations on inheritance are available in the dataset, but they are too few to feasibly utilise. For a detailed description of the wealth construct in NIDS Wave 2, see (Daniels et al, 2014). The survey itself asks detailed information on all members of the household, encompassing residents and non-residents at the time. Data was collected from three main categories, namely, households, adults and children, on the topics of household characteristics, mortality history, living standards, expenditure, consumption, demographics, education, labour force participation, income, health, wealth, education and family support (NIDS, 2012).

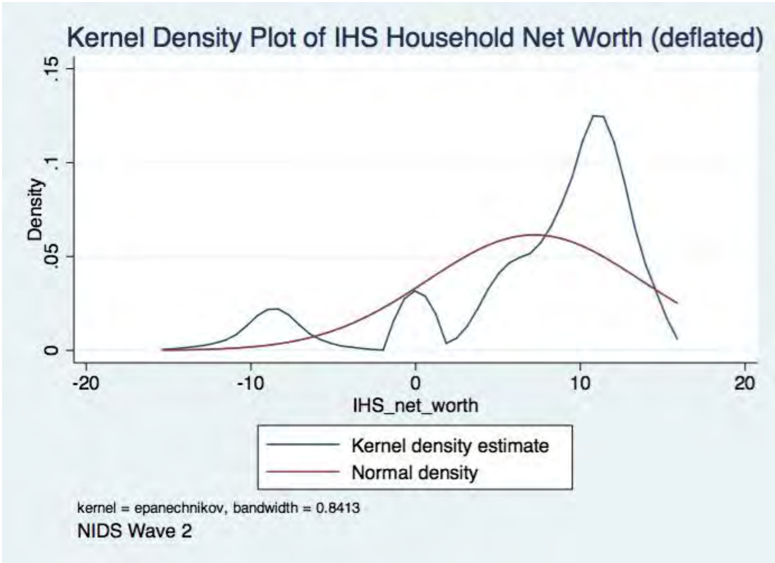
Utilising Wave 2 of the NIDS dataset, specifically looking at the household wealth variables and its components, we consider assets, liabilities and household demographic variables. We use weighted quantile regressions to examine the distribution of wealth, which corrects for the potential bias arising from non-random sample attrition (Maitra & Vahid, 2005). The financial variables were deflated using the CPI value for the base month, September 2010. All incomplete observations were set to missing and duplicate household observations were removed. We excluded the extreme observations of the derived net worth variable only, using the *bacon* algorithm of outlier detection method in STATA. For a detailed discussion of the *bacon* algorithm, see Appendix One. Outliers are a major concern when dealing with wealth data, as they skew the distribution far more than outliers in income data. The sheer distance of the outliers from the mean in our dataset and the high weights associated with each outlier household, assert a level of influence that skews our analysis to a great extent (Daniels et al, 2014). All plots, graphs and estimations are weighted to account for attrition as a result of nonresponse from households between waves.

The net worth variable has been transformed, utilising the inverse hyperbolic sine transformation, which is defined as:

$$\log(y_i+(y_i^2+1)^{1/2})$$

where y_i is the variable of interest. Except for very small values of y, the inverse sine is approximately equal to $\log(2y_i)$ or $\log(2)+\log(y_i)$. Thus it is interpreted in the same manner as a standard logarithmic dependent variable. However, the inverse hyperbolic sine is defined at zero, it is an alternative transformation for correcting the problems of excluding zero and negative wealth values, disproportionately misrepresenting wealth for some households (Friedline et al, 2012). Looking at the kernel density plots of the IHS transformed net worth variable. The distribution graph Figure 1 shows that net worth is non-normally distributed in comparison to the normal distribution.

Figure 1: Distributional Plot of IHS Household Net Worth



ii. Estimation Methods

Research into the socioeconomic determinants of households requires knowledge of the socioeconomic position of said households. Researchers, especially in cases where wealth data is unavailable or unreliable, rely on income or expenditure as a measure of household standard of living (Hargreaves et al, 2007). Furthermore, there are a number of important psychological variables, such as preference parameters like the degree of risk version, self-control, taste for saving and voting patterns (Skopek et al, 2014). Thus, we have the problem of omitted variable bias, as we do not have these variables. Consequently, we are looking at the static implications of a dynamic model, at a single time point.

Models of households with heterogeneous preferences amongst its members have become the standard framework for analysing household behavior (Matilla-Wiro, 1999). Most economic and financial decisions are made at the household level and it is often assumed that households behave as if they were single decision-making unit maximizing a well-behaved utility function (Matilla-Wiro, 1999). Examining household behavior and individual behaviour are not identical investigations, however, in most surveys respondents are usually the representative head of the household (Matilla-Wiro, 1999). Thus, most household investigations require the adaptation of individual level variables into those of the representative household head and then into household variables.

Wealth accumulation models have tended to be macroeconomic general equilibrium, quantitative models, with heterogenous agents. Generally, there are three main types of models used, namely, overlapping generations (OLG) models, infinitely lived models and a model that involves a mixture of both. This investigation is different, in that it is micro-founded, although it does entail utilising the LCH in a similar manner to the OLG model (Cagetti & De Nardi, 2008). There is a common theme found in the explanatory variables utilised in the micro-founded models. The literature found the key determinant variables to be gender, educational attainment, household composition, occupation, ethnic origin, birthplace, urbanisation, and age (Meng, 2007; Di Matteo, 2010; van Rooij et al, 2012; Skopek et al, 2014; Heady et al, 2008).

Turning to estimation, the wealth equation is based on Heady et al (2008) (Heady et al, 2008). Heady et al (2008) was chosen as the Australian economy has had similar discrimination issues and colonization as South Africa (Heady et al, 2008). Furthermore, the wealth distribution has similar characteristics to the South African wealth distribution; namely assets are heavily concentrated in older households, pension wealth is increasing rapidly and the main assets are real estate and financial. (Heady et al, 2008; Aron & Muellbauer 2006). The authors utilized characteristics that household occupants were born with, education, household type, health, hours worked, income, and attitudinal variables as determinants of wealth (Heady et al, 2008). We estimate,

$$\mathbf{W = \beta_0 + \beta_1 gender + \beta_2\, education + \beta_3 race + \beta_4 marital\ status + \beta_5 geotype + \beta_6 health + \beta_7 work\ hours + \beta_8 age + \beta_9 age^2 + \beta_{10} children}$$

Where W is the inverse hyperbolic sine transformation of net worth, β_0 is the expected median net worth level of a single white female, who has fair health, no children and lives in an rural formal area. Race, geotype, marital status and health are dummy variables split into various categories shown in the regression tables. The covariates are individual respondent variables, thus we restrict them to be the observations of the household head and subsequently created household variables from those observations. Variables such as income and work hours are not included. Income is omitted as it is too closely related to wealth and will skew any estimates. Furthermore, work hours are not included as it limits the sample to only the employed households. For a clearer understanding of all variables utilised, Table One below details the variables, the rationale for inclusion in the model and the expected effect on the dependent variable.

Table One: Explanatory Variables

Variable	Details	Rationale for Inclusion	Expected Effect
Gender	Dummy variable for whether the respondent is male or not	Personal characteristics of respondents	Positive; In the case of gender, one would expect men to accumulate more wealth than women.
Race	Dummy variables for 4 race categories Categories: African, Coloured, Asian Base Category: White	Personal characteristics of respondents	White; positive Asian; positive Coloured; negative African; negative This is due to the legacy of discrimination
Marital Status	Dummies for 5 categories of marital status Categories: Married, Divorced/Separated, Widowed, Living wit Partner Base Category: Never Married	Personal characteristics of respondents	Positive; Marital status for men would be expected to have a positive effect on the level of terminal wealth. Married women provided unpaid labour, such as child care, which allows men to work.
Geotype	Dummies for 4 categories of geographic location Categories: Urban Formal, Tribal Authority Area, Urban Informal Base Category: Rural Formal	Testing the effect of location, particularly with reference to TAAs which may have unidentified effects.	Positive; Urban areas were concentrations of industry and commerce, where one would expect, on average, to see higher levels of wealth. Negtive; Tribal Authority Areas tend to have agricultural based economies and are less integrated with the economy.
Health	Dummies for 5 categories of well-being; Respondent's perceived health status Categories: Excellent Health, Very Good Health, Good Health, Poor Health Base Category: Fair	Personal characteristics of respondents	Positive; For those who are well, they are expected to spend less money on health care Negative; For those weho are ill, they are expected to spend more money on health care
Education	Education in years	Testing the effect of education	Positive; One would expect to see wealth increase as the respondent's years of education increases.
Age	Age in years	The test of the life-cycle motive for saving is the inclusion of age and age squared as variables to see whether a hump-shaped wealth-age profile with statistically significant coefficients emerges.	Positive; Net worth increases with age but because we are using cross-sectional data, we cannot disentangle whether this is attributable to age or cohort effects.

Tables Three and Four in the analytical section indicate the results of the OLS estimation and the quantile estimation of the wealth equation. The regression diagnostics in Appendix Two and Appendix Three reveal that the use of OLS leads to a misspecified model. Thus, quantile regression is apt for the model, as the covariates affect our dependent variable differently across its distribution and we want to investigation different points within our sample. Quantile regression extends the idea of quantiles to the estimation of conditional quantile functions (Koenker & Portnoy, 1999). It models quantiles of the conditional distribution of the response variable as functions of observed covariates (Koenker & Portnoy, 1999).

The quantile is defined through an optimisation problem, where the median is defined as the solution to the problem of minimising a sum of absolute residuals. The symmetry of the absolute value function implies that the minimisation of the sum of absolute residuals must equate the number of positive and negative residuals (Koenker & Portnoy, 1999). This ensures that there is the same number of residuals, both above and below the median. In order to specify the other quantiles, we minimise a sum of asymmetrically weighted absolute residuals, by giving differing weights to the negative and positive residuals (Koenker & Hallock, 2001). Solving

$$\mathbf{Min}_{\xi \varepsilon R} \ \Sigma\, \rho_\tau (Y_i - \xi)$$

Where ρ_τ is the tilted absolute value function that yields the τth sample quantile as its solution. In order to specify the conditional quantiles, we solve the minimisation problem for the estimate of the conditional median function (Koenker & Hallock, 2001). We replace the scalar ξ with the parameter ξ (χ ,β) and set ρ_τ=0.5. To obtain the estimates of the other conditional quantile functions, we replace the absolute values by ρ_τ(.) and solve

$$\mathbf{Min}_{\beta\, \varepsilon R^{\rho}} \ \Sigma\, \rho_\tau (Y_i - \xi (\chi ,\beta))$$

utilising linear programming methods (Koenker & Hallock, 2001). The quantile regression estimator is found to be asymptotically normally distributed, more robust to outliers and non-normal errors than OLS. The estimation process allows for a more detailed characterisation of the data, in order to consider the impact of a particular explanatory variable on the entire distribution of the response variable, and not simply its impact on the conditional mean. Furthermore it is invariant to monotonic transformations of the estimation equation and is semi-parametric, as it avoids any assumptions about the parametric distribution of the error process (Koenker & Hallock, 2001).

The conditional quantile functions of the response variable y, given covariates, x, are assumed to all be parallel to each other, according to the classical theory of linear regression. This implies that the slope coefficients of individual quantile regressions will be identical, however, in practice they tend to vary across quantiles. The Wald test involves testing for equality of slope parameters across quantiles, in order to allow for correct inference with regards to the estimation (Koenker & Portnoy, 1999).

The final estimated wealth equation

$$\mathbf{W = \beta_0 + \beta_1 gender + \beta_2 number\ of\ children + \beta_3 race + \beta_4 marital\ status + \beta_5 geotype + \beta_6 health + \beta_7 work\ hours + \beta_8 age + \beta_9 age^2 + \beta_{10} age^3 + \beta_{11} age^4}$$

The final model is exactly the one estimated in the previous regressions, except that we have improved the model by changing to an age quartic to an age quadratic format. The age quadratic was originally chosen to test the life cycle hypothesis and as seen in Figure 6, this format is not representative of the data's distribution. An investigation into a specification involving age cohorts revealed the depth of non-linearity in derived ne worth and the quartic structure was imposed as it best fit the distribution. The investigation into the appropriate age specification is detailed in the descriptive section below.

IV. Empirical Results

i. Descriptive

Table Two below presents the summary statistics of the components of net worth, as well as variables from the econometric model, barring the binary variables. The one shot variable is in answer to the survey question: 'Suppose you (and your household members living here) were to sell off all your major possessions (including your home), turn all of your investments into cash and pay all your debts –would you have something left over, breakeven or be in debt?' (Daniels et al, 2014). The results indicate that the two measures of household wealth are vastly different, with the derived net worth showing much larger variance than the one shot measure (Daniels et al, 2014). Furthermore, the 10^{th}, 25^{th} and 50^{th} percentile have zero as their net worth observations, signifying that households in these percentiles do not perceive themselves to have a positive or negative net worth.

As mentioned in Daniels et al (2014), the minimum values for several of the assets and liabilities variables are shown to be approximately 1 (Daniels et al, 2014). This indicates the presence of outliers, particularly at the bottom end of the distribution (Daniels et al, 2014). This finding further reiterates the need for the use of the *bacon* algorithm of outlier detection on the net worth variable. The oldest household head is shown to be 104 years old, indicating the possible presence of intergenerational effects present in the distribution, although not explicitly modeled. There are significantly more assets reported than liabilities. The financial and real estate assets are the largest contributors to total assets, whereas financial liabilities are the largest component of total debts. The Coefficient of Variation values are relatively small, except for total assets, real estate assets, vehicle assets and financial assets. This indicates that these variables have a relatively high level of variability in its frequency distributions.

Table Two: Summary Statistics (2010 Rand, weighted)

Variable	Min	P10	P25	P50	Mean	P75	P95	Max	Coefficient of Variation	N
Derived Net Worth	-1000000	-673.170	147.166	7203.905	118026.4	67763.41	840557.6	1692280	2.391	2909
One-shot Net Worth	-1000000	0	0	0	188551.3	20022.57	1001129	2.50e+07	4.396	1865
Total Assets	.950	148.493	1047.638	18993.58	509160.5	116582.7	1846680	3.11e+08	8.523	2571
Real Estate	.950	1007.955	5005.644	40318.18	487184.9	300338.6	1200000	3.11e+08	10.235	1740
Business	100.796	1966.741	5039.773	74831.27	296773.8	200225.7	2015909	2950111	1.919	59
Vehicles	30.034	15016.93	29145.67	55983.27	199927.7	114128.7	388062.5	6.81e+07	10.107	373
Financial	.998	49.49777	110.666	482.0485	161751.1	3074.261	504518	1.50e+07	5.522	1604
Retirement Annuity	149.663	503.977	1088.591	14966.25	487143.7	500564.3	1209546	6228513	2.726	73
Livestock	39.335	1097.525	2950.111	11879.46	39004.94	44538.08	163639.6	512895.8	1.570	95
Total Debts	1.008	484.170	1609.912	7055.682	73540.48	28872.5	496796.1	1181853	2.268	812
Real Estate	35.040	400	50056.43	181431.8	246186.7	400451.5	712259.1	1100000	.953	125
Business	1.008	1001.129	3003.386	5039.773	9976.715	16268.34	35000	104464.7	1.201	13
Vehicles	160.181	9000	10079.55	29230.68	68848.09	90101.58	302386.4	399299.7	1.278	61
Financial	4.983	399.100	1209.546	3573.76	9786.152	10000	33041.24	409762	2.118	735
Age	15	26	32	40	43.003	53	71	104	.340	3516
Years of Education	1	5	8	10	9.775	12	15	18	.332	2910
Children	1	1	1	2	1.994	3	4	7	.589	1275

Turning to the composition of wealth across age profiles and geographical areas, calculated as per Daniels et al, (2014) (Daniels et al, 2014). Figures 2 through 4 show the assets and liabilities across specific age cohorts of the household head and across the four geographical sub-divisions, namely: Rural Formal (RF), Urban Formal (UF), Urban Informal (UI) and Tribal Authority Area (TAA). Figure 2 below characterizes the liabilities across age cohorts; it shows that real estate debt dominates from ages 25 to 74. It increases from age cohorts 15 to 24 to 25 to 34, remains constant from ages 25 to 44 and then starts declining from ages 45 to 74. It appears to be nonexistent in the cohort of household heads aged 75 and over. Financial debts are most prevalent in the very oldest cohort (75 & over) and the very youngest cohort (15 to 24). Vehicle debt tends to be a minor component from ages 25 to 74, whereas business debts are infinitesimal from ages 25 to 75.

Figure 2: Portfolio of Liabilities by Age Cohort

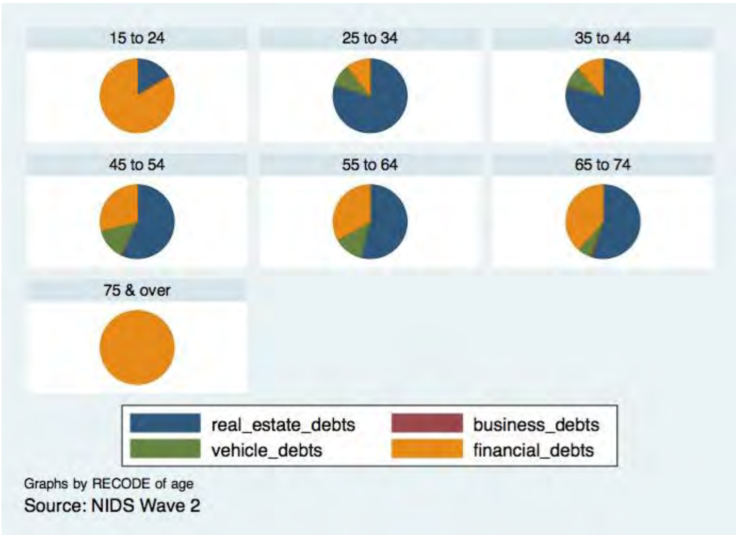
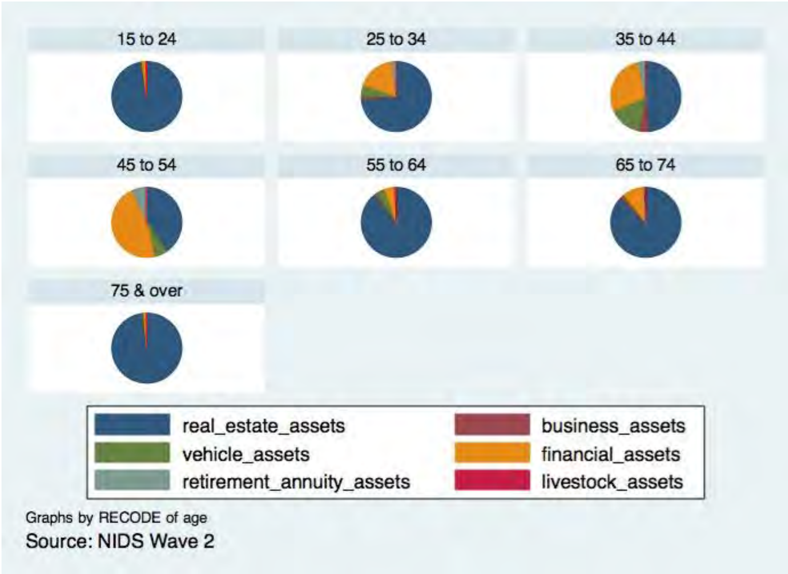


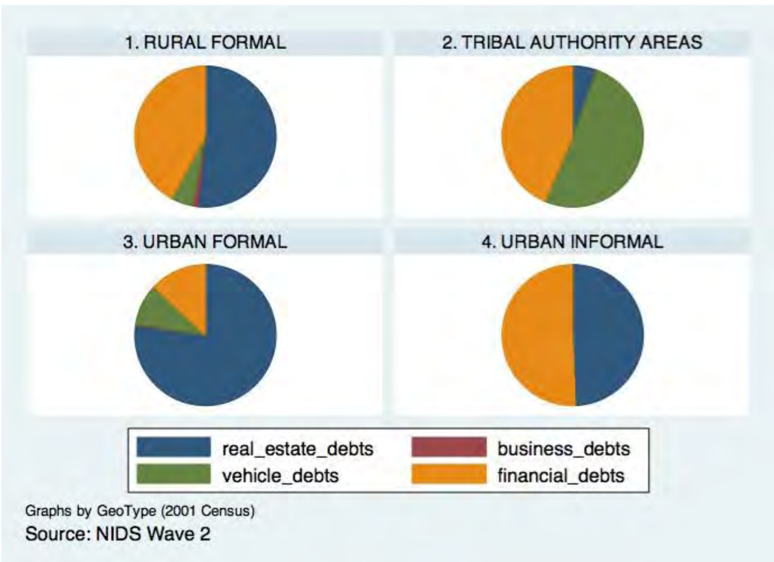
Figure 3 below indicates assets over the age cohorts and similarly to Figure 2, real estate assets dominate throughout the age cohorts. This is an expected result for the older age cohorts, as they accumulate real estate debt when they are younger and as they earn more income over the years, they are able to pay off this debt and turn it into an asset. However, the prevalence of real estate assets in the youngest age cohort (15 to 24) is puzzling, as there is no corresponding real estate debt in the same age cohort in Figure 2. This indicates the presence of an unknown real estate effect in the younger cohort. Livestock, retirement and business assets are minor components throughout the age cohorts, whereas financial assets increase from the youngest age cohort to age cohort 45 to 54 and then begin decreasing from ages 55 to the oldest cohort. This is as expected as we hypothesise that households accumulate financial assets as they get older and then divest themselves of these assets when they retire.

Figure 3: Portfolio of Assets by Age Cohort



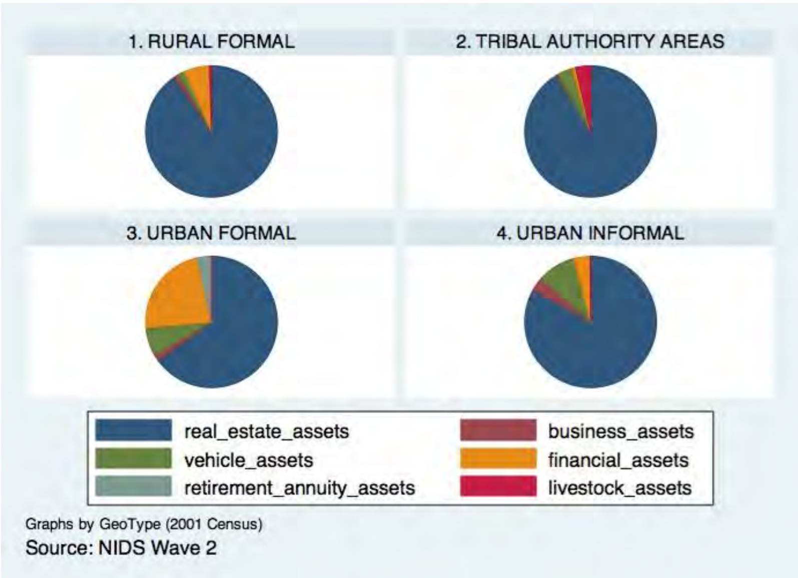
Turning to the composition of liabilities across geographic locations, Figure 4 indicates that Rural Formal areas consist almost entirely of real estate and financial debt, with minor slivers of vehicle and business debt present as well. The real estate assets owned in the TAA would chiefly be the land upon which homes are built and the dwelling itself. The TAA displays a relatively minor component of real estate debt, with extensive vehicle and financial debt. Both Urban Formal and Urban Informal exhibit a large real estate debt component, with UF showing minor components of vehicle and financial debt and the remaining space in UI dominated by financial debts. The composition of liabilities is as expected for each geographical area, since property was given to communities in the tribal authority areas, it is expected that financial debt would be the most prominent liability. Similarly for rural formal and urban formal, real estate debt is expected to be the largest component, as the largest liability would be the acquisition of land.

Figure 4: Portfolio of Liabilities by Geotype



Finally, in Figure 5, we look at the asset composition over the geographic areas. Similar to the age profile of assets, the geographical composition of assets consists mainly of real estate assets. Rural Formal (RF) and Tribal Authority Areas appear to have little variation in asset composition, with RF showing minor livestock, vehicle, business and financial assets. Urban informal has the largest variation, although real estate debt is still the overriding factor. As noted in both Figure 2 and 3, there appears to be some unknown effect occurring within younger age cohorts regarding real estate. Similarly, we have extensive real estate assets present in the TAAs, however; the corresponding TAA real estate debt component in Figure 4 is relatively minor. The property asset allocation prominence in TAAs is precisely the conundrum at hand. The land in the area is communally owned, so we cannot separate household/individuals/ communal land ownership and ascertain the true level of wealth in these areas. We hypothesise that the real estate assets values for TAAs are inflated, as these households report that they are the owners of their land/holdings, when this is in fact not the case.

Figure 5: Portfolio of Assets by Geotype



Following from the descriptive analysis, we expect to see a similarity between the end points of the distribution, namely the 5th and 95th quantiles. This is as a result of similar behaviour in these quantiles. As previously mentioned, those households in the lower quantiles are not necessarily living in poverty, but rather utilise a large amount of credit. If these tail similarities were present, it would be evidence of similar behaviour of households in these quantiles. However, it would not be enough to conclusively state the presence of high-income households in the 5th quantile.

Turning now to the estimation of the original economic model with the quadratic age variables, to allow for the testing of the Lifecycle hypothesis. In Table Three below, the results of the OLS estimation of the original econometric model are contrasted with results of the quantile regression estimation. At first glance, we see that the gender coefficient is insignificant in both forms of estimation and across every quantile tested. The coefficients of the gender variable are negative in the OLS estimation and in the 5th, 10th, 50th and 95th quantile estimation. Furthermore, there is no discernible pattern in its magnitude across the quantiles. Thus, aligned with the literature by Zimmer & Das (2014), the results indicate that more often than not, male- headed households are less wealthy than female-headed households (Zimmer & Das, 2014). Similarly, the dummy variables for the various health categories are also all insignificant in both OLS and quantile regression, except for the Poor Health dummy variable, which is significant at the 5% significance level in the 95th quantile. The Poor Health coefficients decrease in size, although not by much, as we move from the 10th quantile to the 95th quantile. The coefficient on education is only significant in the 25th quantile at the 10% significance level and at the 50th, 75th and 95th quantile at the 1% significance level. The OLS education coefficient is significant at the 5% significance level.

The race coefficients appear to be significant in almost every quantile and the OLS estimation at various significance levels. Their coefficients also appear to increase in magnitude from the 5th to the 95th quantile. The African and Coloured coefficients are all negative, whereas the Asian coefficients are only positive in the 25th, 50th and 95th quantiles. The geotype variables also have significant coefficients across every quantile and the OLS estimation, although less so in the 5th and 10th quantile. The TAA coefficient seems to be most significant around the 50th quantile, which is expected as this quantile represents the poor with approximately zero net worth. It must be noted that all coefficients of the geotype variables in every quantile are positive in comparison to rural formal households. This result seems plausible for UF and possibly even UI households in comparison to RF households. However, this is not an expected result between TAA and RF. Again, there seems to be some hidden effect driving wealth accumulation inside Tribal Authority Areas. This hidden effect is believed to be a direct result of the dual land tenure system in South Africa.

The coefficients on the variables relating to marital status have more varied effects. The significance of the positive married coefficients is at the 1% significance level in the 50th, 75th and 95th quantile and at the 10% significance level in the 5th quantile, with a negative coefficient. This seems to indicate that being married only aids in wealth accumulation for the wealthier proportion of the sample. Living with Partner appears to have negative coefficients throughout the quantiles, except for the 10th quantile and none of these coefficients are significant at any level. The quadratic specification of the age variable is significant only at the 5th quantile at the 10% significance level, the 50th quantile at the 5% significance level and at the 75th quantile at the 1% significance level. Similarly, the quadratic age specification remains insignificant, except for the 5th and 75th quantile. This does not indicate a violation of the lifecycle hypothesis, as the theory is not expected to hold within each quantile. Overall, the OLS model was declared insufficient to model the wealth distribution, as the model failed the diagnostic testing of OLS assumptions shown in Appendix Two. The quantile model is inadequate with regards to the specification of the age variables.

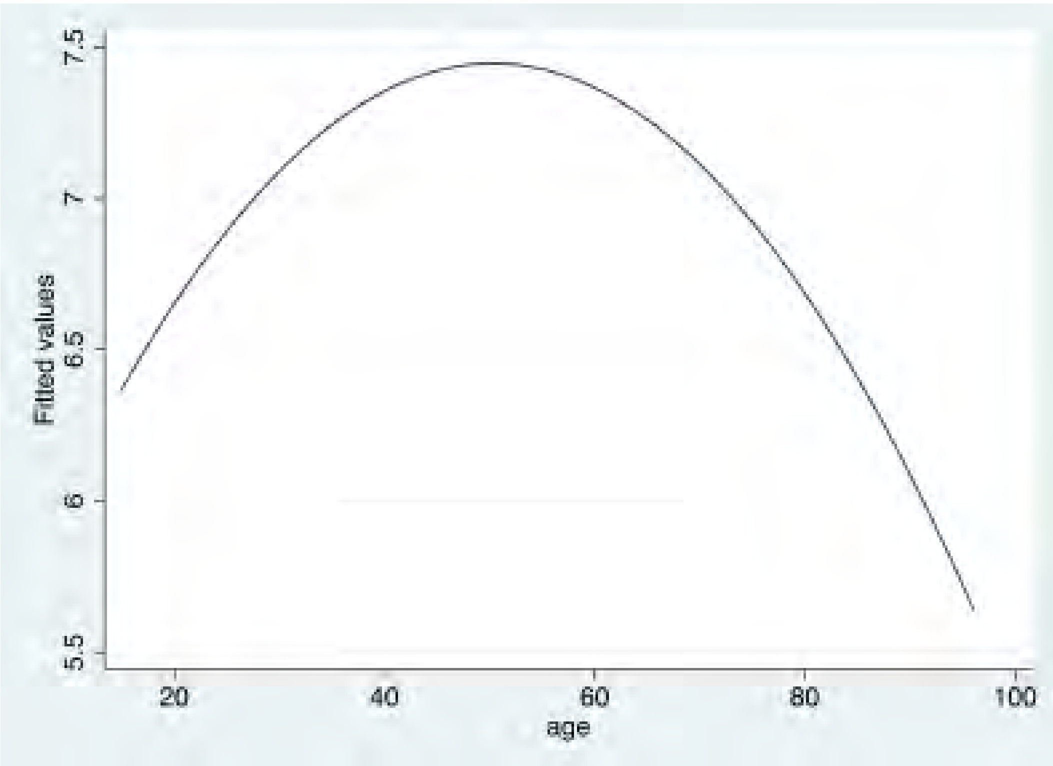
Table Three: OLS and Quantile Regression Results

	Dependent Variable: IHS Net Worth						
	OLS	Quantile					
		0.05	0.10	0.25	0.50	0.75	0.95
Gender	-0.239	-0.939	-2.157	0.130	-0.0375	0.194	-0.00623
	(0.442)	(0.931)	(1.450)	(0.906)	(0.298)	(0.174)	(0.198)
Age	0.127*	0.0467	0.0384	0.199	0.137**	0.0910***	0.0643
	(0.0672)	(0.156)	(0.201)	(0.138)	(0.0593)	(0.0284)	(0.0393)
Age^2	-0.00109*	-0.000812	-0.000277	-0.00164	-0.00112*	-0.000722**	-0.000484
	(0.000634)	(0.00163)	(0.00184)	(0.00131)	(0.000598)	(0.000297)	(0.000346)
Education	0.122**	-0.142	0.00637	0.194*	0.197***	0.129***	0.0920***
	(0.0522)	(0.157)	(0.189)	(0.0989)	(0.0362)	(0.0207)	(0.0283)
African	-5.284***	-20.47*	-20.01***	-5.856***	-2.947***	-1.644***	-1.474***
	(1.317)	(11.15)	(6.410)	(1.453)	(0.514)	(0.424)	(0.540)
Coloured	-3.965***	-20.59*	-18.42***	-3.438**	-1.936***	-1.240***	-1.209**
	(1.353)	(11.24)	(7.005)	(1.494)	(0.420)	(0.445)	(0.549)
Asian	-3.343	-24.33*	-24.62**	1.016	0.0791	-0.489	0.0835
	(2.201)	(14.01)	(12.29)	(8.068)	(0.910)	(0.623)	(0.835)
Children	-0.167	-0.183	-0.112	-0.162	-0.0975	-0.0665	-0.0663
	(0.168)	(0.387)	(0.709)	(0.233)	(0.111)	(0.0719)	(0.0880)
Married	0.885*	-1.541*	-0.281	1.170	1.211***	0.839***	0.795***
	(0.481)	(0.896)	(1.242)	(1.013)	(0.403)	(0.182)	(0.251)
Living with Partner	-0.251	-0.696	1.399	-0.434	-0.175	-0.118	-0.103
	(0.667)	(1.211)	(1.871)	(1.855)	(0.464)	(0.229)	(0.489)
Widowed	1.129**	0.577	0.571	1.336	0.891*	0.568**	0.292
	(0.568)	(1.567)	(1.600)	(1.008)	(0.470)	(0.240)	(0.386)
Divorced/Separated	-1.150	-1.036	-1.256	-4.328**	-0.775	0.552	0.457
	(0.925)	(1.302)	(1.956)	(1.947)	(1.042)	(0.517)	(0.409)
Urban Informal	2.844***	1.574	1.790	7.109***	3.411***	0.534	0.608
	(0.853)	(1.294)	(2.493)	(2.201)	(1.082)	(0.452)	(0.452)
Tribal Authority Area	3.722***	2.730*	8.213***	6.520***	3.674***	0.802**	0.766**
	(0.659)	(1.565)	(1.523)	(1.517)	(1.010)	(0.376)	(0.316)
Urban Formal	2.689***	1.059	1.581	5.453***	3.407***	1.125***	1.353***
	(0.636)	(0.855)	(1.307)	(1.556)	(1.040)	(0.394)	(0.358)
Excellent Health	0.878	0.147	1.199	1.104	0.619	0.361	-0.170
	(0.587)	(1.217)	(1.439)	(1.834)	(0.510)	(0.251)	(0.339)
Very Good Health	0.745	-0.0940	0.896	1.106	0.611	0.114	-0.480
	(0.581)	(1.190)	(1.634)	(1.811)	(0.447)	(0.274)	(0.323)
Good Health	0.210	-0.0450	0.223	0.332	-0.00310	-0.182	-0.340
	(0.580)	(1.203)	(1.322)	(1.910)	(0.443)	(0.282)	(0.332)
Poor Health	0.242	-0.0146	3.258	0.438	-0.595	-0.264	-0.903**
	(0.825)	(3.504)	(2.397)	(2.044)	(0.690)	(0.405)	(0.428)
Constant	4.433*	12.47	10.01	-2.455	3.169*	8.316***	11.01***
	(2.394)	(12.11)	(8.986)	(5.020)	(1.629)	(0.885)	(1.184)
Pseudo R^2	0.055	0.0384	0.0889	0.0505	0.0554	0.0590	0.1011
Observations	1490	1490	1490	1490	1490	1490	1490

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; Dummy variable base categories are as follows, Health: Fair, Geotype: Rural Formal, Race: White, Marital Status: Never Married, Gender: Female

Focusing on the appropriate age specification for the estimation, a hump shaped outline is anticipated for derived household net worth over the lifecycle. The age effect is associated with varying abilities or preferences of households over the lifecycle (Modigliani & Brumberg, 1954). The expectation is that households with younger heads will hold more debts than assets, resulting in either negative or very low levels of net worth (Modigliani & Brumberg, 1954). Figure 6 below shows the calculated turning point of the OLS quadratic specification of age utilised in Table Three. It was expected to be at or close to 60 years of age, retirement age. However, we see that it appears to be closer to 50 years of age, which is ten years before females in South Africa are eligible for pension and 15 years before South African males are eligible for pension. However, the regression in Table Three indicates that the age and age^2 variables are rarely significant in either the OLS or quantile regression models. Thus we turn to an age cohort specification.

Figure 6: Turning Point of Age



An alternative method of accounting for the lifecycle hypothesis in the regression model would be to replace the age and age^2 covariates with age cohorts. The purpose would be to try and understand the age, cohort, and period effects at work within the wealth distribution (Glenn, 2005). Households and individuals age in a dynamic society where social and cultural changes impinge on their attitudes and behaviour (Glenn, 2005). Differences by age shown in cross sectional data may not necessarily arise as a result of age effects; rather the differences might be due to different cohort experiences. Furthermore, any estimation of the differences between older and younger households could also be as a result of the correlation between wealth and longevity (Glenn, 2005). These complications are as a result of the identification problem present between age, cohort and period effects. It is statistically impossible to separate these effects (Glenn, 2005). We divide the sample population into three equal sized age cohorts, namely from 15 to 34, 35 to 48, and 49 to 104. We then utilize OLS regression methods to estimate the original estimation model with each age cohort.

Table Four: Regression Model with Age Cohorts

	Dependent Variable: IHS Net Worth		
	OLS		
Gender	-0.0731	0.0314	-0.0440
	(0.490)	(0.491)	(0.496)
Age Cohort 1 (15 to 34)	-0.951*	-	-
	(0.508)		
Age Cohort 2 (35 to 48)	-	0.539	-
		(0.419)	
Age Cohort 3 (49 to 104)	-	-	0.159
			(0.513)
Education	0.116*	0.0767	0.0998
	(0.0684)	(0.0681)	(0.0725)
African	-5.301***	-5.422***	-5.460***
	(1.361)	(1.360)	(1.361)
Coloured	-3.885***	-4.034***	-4.053***
	(1.402)	(1.400)	(1.402)
Asian	-3.237	-3.343	-3.292
	(2.255)	(2.256)	(2.259)
Children	-0.295	-0.310	-0.236
	(0.192)	(0.196)	(0.193)
Married	0.724	0.917*	0.951*
	(0.531)	(0.516)	(0.521)
Living with Partner	-0.301	-0.405	-0.330
	(0.735)	(0.736)	(0.739)
Widowed	1.233*	1.618***	1.492**
	(0.641)	(0.618)	(0.658)
Divorced/Separated	-0.965	-0.618	-0.654
	(0.987)	(0.968)	(0.985)
Urban Informal	2.893***	2.933***	2.873***
	(0.950)	(0.952)	(0.953)
Tribal Authority Area	4.097***	4.130***	4.072***
	(0.754)	(0.755)	(0.756)
Urban Formal	2.803***	2.854***	2.813***
	(0.708)	(0.709)	(0.710)
Excellent Health	1.049	0.862	0.926
	(0.672)	(0.668)	(0.677)
Very Good Health	0.925	0.825	0.861
	(0.678)	(0.677)	(0.681)
Good Health	0.200	0.102	0.129
	(0.697)	(0.696)	(0.698)
Poor Health	0.212	0.205	0.234
	(1.056)	(1.057)	(1.057)
Constant	7.970***	7.919***	7.785***
	(1.800)	(1.801)	(1.865)
R^2	0.060	0.059	0.058
Observations	1212	1212	1,212

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Dummy variable base categories are as follows, Health: Fair, Geotype: Rural Formal, Race: White, Marital Status: Never Married, Gender: Female

Table Four above details the original estimation across the different age cohorts separately. The results are startlingly similar across all three models. In each model the coefficients of African, Coloured, Urban Informal, Tribal Authority Area and Urban Formal are significant at the 1% level. Furthermore, Widowed is significant in each regression, but at different significance levels; 10 % in the first regression, 1% in the second regression and 5% in the last regression. The difference between the regressions lies in the variables age cohort, education and married. The first regression model is the only estimation in which an age cohort (15 to 34) variable is significant at the 10% level.

It appears that age is an important factor, but only for households with young heads. In the same estimation education is also found to be significant at the 10% level. It seems to indicate that education is only a prominent factor for households headed by younger individuals and not for older headed households. Lastly, the married coefficient only becomes significant in the last two regressions. In the second and last regression married is significant at the 10% level. This may be due to the fact that being married is only a critical factor for older headed households where assets and liabilities may have to be shared.

Looking at the distribution of IHS derived net worth across the specified age cohorts and each Figure 7 through 9 indicates at least three modes in each age cohort. In each age cohort, derived net worth showcases multi-modality, in particular three modes. Thus, a quartic specification of age is more relevant for the estimation model. This finding is also reiterated in the work of Murphy & Welch (1990) who found that an age quartic specification was a better fit for testing the earnings-age profile within the lifecyle hypothesis (Murphy & Welch, 1990). We apply the same logic to net worth in this investigation and hypothesis that the signs of the coefficients will be negative for the age variable, positive for age^2, negative for age^3 and lastly positive for age^4.

Figure 7: Distributional Plot of IHS Household Net Worth for Age Cohort 15 to 34

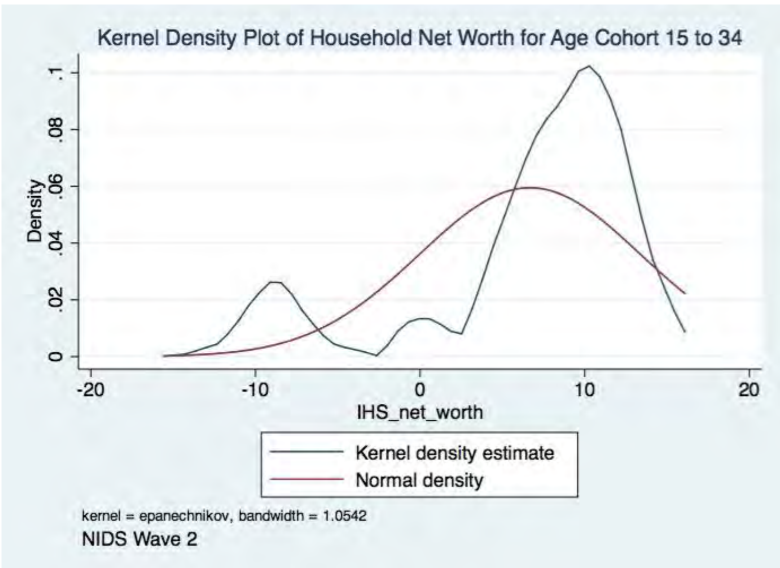


Figure 8: Distributional Plot of IHS Household Net Worth for Age Cohort 35 to 48

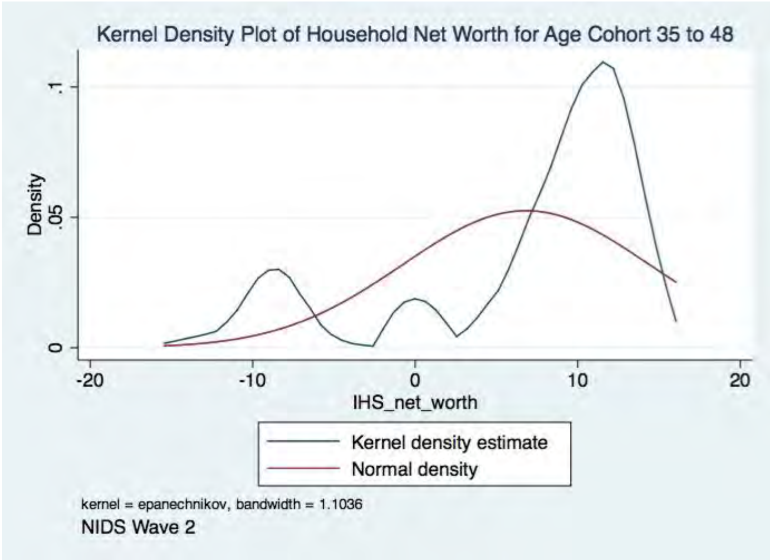
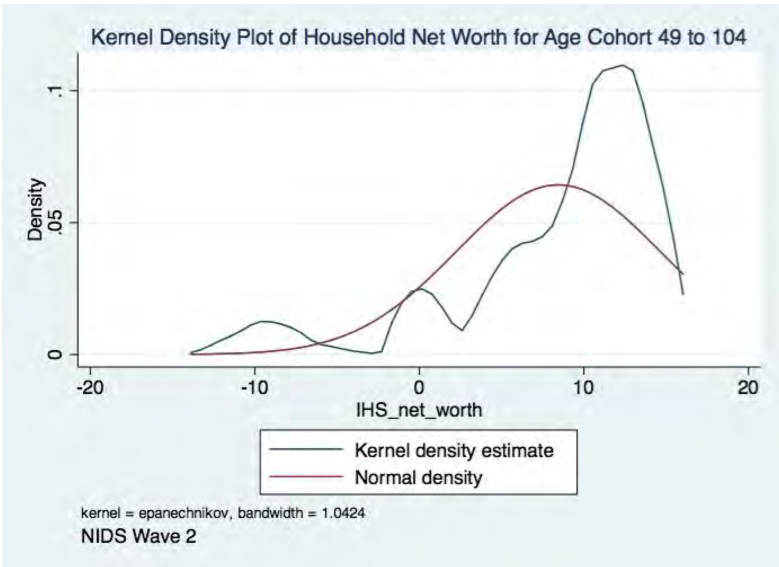


Figure 9: Distributional Plot of IHS Household Net Worth for Age Cohort 49 to 104



The age cohort estimations, while generally informative, do not critically explain or separate the age, cohort, and period effects within the wealth distribution. Thus, we separate the sample population into quantiles, in order to better analyse the distribution at specific points in the sample. The idea is to look at the particularly indebted proportion of the population, that is, the 5th quantile, as well as the poor at the 50th quantile and the richest part of the sample at the 95th quantile.

ii. Analytical

Initially, the investigation sought to explain the results found in the descriptive analysis. The TAAs hold the majority of its assets in real estate, with minor real estate liabilities reported. We attempt to test the hypothesis that the wealth accumulation process is detectably dissimilar in TAAs in comparison to the other geographic locations. The preliminary regression model was estimated controlling for the geographical areas. In all regression results tables, the coefficients are shown to three decimal places, as well as the standard error of each coefficient in brackets. The results in Table Five below showed that the 5th quantile and the 95th quantile did not exhibit any sign similarities and that the TAA binary variable had coefficients that were statistically significant at the 1% level in all quantiles tested, except for the 5th and 95th quantile. Furthermore, the coeffcients on the TAA variable are the largest in every quantile tested. For example, in the 10th quantile, a household living in the TAA increases its net worth by 834.3% in comparison to a rural formal household, ceteris paribus. This result is astronomical and while not necessarily a feasible indication of rural wealth accumulation, does point to the inflated effect hypothesized within TAAs.

This does not indicate that the tribal authority area is more important when estimating the wealth distribution associated with NIDS Wave 2, or that the behaviour and pattern of wealth accumulation in these areas is different from other geographical areas. None of the coefficients of the variables are particularly large relative to each other in any quantile. However, when we interpret these coefficients, it amounts to a much larger change. For example, in the 10th quantile, a one unit change in the TAA variable yields an 83,43% change in IHS net worth, ceteris paribus.

Table Five: Geotype Regression Model

	Dependent Variable: IHS Net Worth					
	0.05	0.10	0.25	0.50	0.75	0.95
Urban Informal	0.699	3.766	6.216***	1.176**	0.422	-1.025*
	(1.019)	(3.117)	(1.781)	(0.462)	(0.344)	(0.550)
Tribal Authority Area	2.639	8.343***	5.607***	1.283***	0.715***	-0.634
	(2.175)	(0.694)	(1.629)	(0.432)	(0.265)	(0.502)
Urban Formal	-0.321	0.965	5.481***	2.375***	1.624***	1.032**
	(0.392)	(2.029)	(1.695)	(0.453)	(0.267)	(0.502)
Constant	-9.070***	-8.343***	-0	7.937***	10.58***	13.25***
	(0.243)	(0.694)	(1.622)	(0.419)	(0.253)	(0.498)
Pseudo R^2	0.0147	0.0533	0.0097	0.0141	0.0170	0.0523
Observations	2,845	2,845	2,845	2,845	2,845	2,845

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1;Dummy variable base category is Rural Formal

In the final estimation model, the transformed derived net worth variable is regressed on a number of explanatory variables as per the estimation equation shown in the methodology section. Table Six showed that there were sign similarities in the 5th and 95th quantiles, except for the variables gender, education, married, living with partner, divorced and poor health. Thus, there appears to be some evidence of similar behaviour patterns in wealth accumulation at the endpoints of the wealth distribution. However, we cannot say anything about what this behaviour in fact is, we are simply attempting to identify its existence.

The coefficients on the gender, children and health variables are all insignificant in all the quantiles. It is quite surprising that children were found to be of no value in the model. Young dependents, in general, tend to stifle the ability of a household to accumulate wealth. A better measure of this effect would be to ask respondents questions regarding the exact nature, age and employment of all household members. The age coefficients are only significant in the 5th and 95th quantile, with sign similarities, all at the 5% significance level. For a one-unit change in the 5th quantile age coefficients, the dependent variable decreases by 314.53 %. This seems to indicate that the presence of the LCH in the 5th and 95th quantiles, implying that LCH effects are prevalent in the rich and indebted proportions of the sample. These would tend to be the high-income households with younger household heads. Lastly, the age coefficients all conform to the initial sign hypothesis. The health coefficients have much the same results as in Table Three; the only significant coefficient is on Poor Health in the 95th quantile at the 10% significance level. It does not appear that the health variable utilised from the survey has had much effect in helping to explain wealth accumulation in the sample. This may be due to the fact that the variable is the respondent's perceived health status, rather than a variable based on medical records.

The coefficients on African and Coloured are significant in every quantile at the 1%, 5% and 10% significance levels. However, the Asian coefficients are only significant at the 10% significance level in the 5th and 10th quantiles. The negative coefficients on the race variables indicate that the marginalization created during apartheid is still very present. The overall picture that emerges from the analysis is that the effects of Apartheid are ever present in the economy, as evidenced by the negative and significant coefficients on the race variables. The education coefficients are significant only in the 25th quantile at the 10% significance level and in the 50th, 75th and 95th quantile at the 1% significance level. Thus it appears that education is important for the relatively poor and relatively rich, but has little significance for those with negative net worth. The education variable has a negative, although insignificant coefficient in the 5th quantile. This is an intriguing result, as it appears to signify that when households are indebted, their education level has no effect on their ability to accumulate wealth. Furthermore, the education coefficients, while mostly significant, do not have particularly large coefficients. It seems that education, while important, is not quintessential for the household's prosperity. This is a somewhat shocking finding and might be as a result of hidden effects present owing to the differing education systems received by different races during apartheid.

The TAA variable is statistically significant at the 1% level in the 10th, 25th and 50th quantiles at the 1% significance level and at the 75th and 95th quantiles at the 10% significance level. The Urban Formal coefficient is significant at the 1% level in the 25th, 50th, 75th and 95th quantiles and the Urban Informal coefficient is only significant at the 1% level in the 25th and 50th quantile. The location of the household appears to be vital to modeling the wealth distribution, as it defines the type of capital accumulation occurring in the region. This outcome is especially convincing when we recollect that those households living in rural areas tend to be those who bore the brunt of discrimination. The positive coefficients on every geotype variable in every quantile indicate that all geotypes are wealthier than rural formal areas. This is not an expected result, although rural areas do tend to house the poorest of the poor. This effect could be due to the incorrect reporting of assets and liabilities, particularly in TAAs.

The married coefficient is significant at the 1% level in the 50th, 75th and 95th quantiles. Similarly to the results in Table Three, it appears that being married is only meaningful for the richer proportion of the sample. Divorced and Living with Partner coefficients do not show any significance across the quantiles in comparison to households with heads whom have never been married. Whereas, results indicated that households with a widowed head has a positive effect on said household's wealth accumulation. Furthermore, the significance on the coefficients indicates that being widowed matters particularly for the poorest, average and wealthiest households. The main conclusions from the model are that being a white, educated, healthy household in any geographic location other than rural formal, is the optimal model for wealth accumulation across all quantiles in our sample. The marital status of a household is also particularly important, however, the effects vary across the quantiles.

The model is limited severely by the lack of available wealth data from the survey and from expanded questionnaires to further probe the anomalies of South African wealth accumulation. Most notably, another time period of data would also allow broader scope for the research and more meaningful conclusions. Questions on risk preferences, saving behavior and attitudes tend to be erroneously reported for various reasons, depending on the circumstances and characteristics of a household. Furthermore, household composition data could be illuminating particularly with respect to both inter- and intra-generational effects. The findings of this estimation point us in the direction of further research, rather than answer any particular questions regarding the South African wealth distribution.

Table Six: Final Regression Model

	Dependent Variable: IHS Net Worth					
	0.05	0.10	0.25	0.50	0.75	0.95
Gender	-1.315	-2.031	-0.100	0.0233	0.231	0.0985
	(0.868)	(1.467)	(0.902)	(0.285)	(0.200)	(0.201)
Age	-3.246**	-0.675	-0.940	-0.449	-0.206	-1.373**
	(1.524)	(1.816)	(1.212)	(0.645)	(0.325)	(0.602)
Age^2	0.102**	0.0284	0.0367	0.0192	0.00951	0.0441**
	(0.0455)	(0.0523)	(0.0353)	(0.0177)	(0.00945)	(0.0188)
Age^3	-0.00135**	-0.000456	-0.000527	-0.000288	-0.000143	-0.000579**
	(0.000581)	(0.000637)	(0.000431)	(0.000206)	(0.000117)	(0.000252)
Age^4	6.32e-06**	2.46e-06	2.53e-06	1.41e-06	6.94e-07	2.69e-06**
	(2.69e-06)	(2.81e-06)	(1.88e-06)	(8.66e-07)	(5.16e-07)	(1.22e-06)
Education	-0.139	0.00107	0.177*	0.204***	0.134***	0.108***
	(0.119)	(0.157)	(0.0942)	(0.0370)	(0.0244)	(0.0255)
African	-20.24*	-19.50***	-5.910***	-2.967***	-1.621***	-1.811***
	(11.32)	(6.931)	(1.460)	(0.534)	(0.358)	(0.519)
Coloured	-20.36*	-17.75**	-3.542***	-1.935***	-1.204***	-1.386**
	(11.49)	(7.152)	(1.311)	(0.475)	(0.398)	(0.558)
Asian	-24.17*	-24.17*	1.111	0.0683	-0.470	-0.324
	(12.90)	(13.37)	(11.72)	(2.977)	(0.652)	(0.789)
Children	-0.183	-0.447	-0.213	-0.156	-0.115	-0.0171
	(0.412)	(0.713)	(0.291)	(0.111)	(0.0729)	(0.0837)
Married	-1.135	-0.638	1.250	1.210***	0.880***	0.740***
	(0.970)	(1.544)	(1.045)	(0.395)	(0.192)	(0.219)
Living with Partner	-0.0654	1.311	-0.202	-0.214	-0.122	-0.0157
	(1.121)	(2.067)	(1.814)	(0.568)	(0.269)	(0.356)
Widowed	0.516	0.998	1.059	1.157***	0.724***	0.267
	(1.732)	(1.679)	(1.123)	(0.438)	(0.228)	(0.302)
Divorced/Separated	-1.437	-1.297	-4.254*	-0.788	0.497	0.284
	(1.846)	(2.344)	(2.196)	(1.442)	(0.594)	(0.379)
Urban Informal	0.463	1.903	7.069***	3.002***	0.439	0.330
	(1.376)	(2.538)	(2.521)	(1.069)	(0.462)	(0.389)
Tribal Authority Area	1.809	7.929***	6.802***	3.140***	0.724*	0.543*
	(1.703)	(1.868)	(1.837)	(1.014)	(0.407)	(0.303)
Urban Formal	0.272	1.394	5.804***	2.805***	0.987***	1.138***
	(1.067)	(1.731)	(1.580)	(1.055)	(0.378)	(0.337)
Excellent Health	0.0403	1.745	1.211	0.583	0.363	0.0441
	(1.253)	(1.844)	(1.555)	(0.419)	(0.266)	(0.238)
Very Good Health	-0.484	1.342	1.279	0.584	0.119	-0.417
	(1.086)	(1.805)	(1.450)	(0.409)	(0.268)	(0.255)
Good Health	-0.433	0.349	0.460	0.0984	-0.242	-0.434
	(1.069)	(1.571)	(1.411)	(0.388)	(0.280)	(0.306)
Poor Health	0.787	3.505	0.689	-0.728	-0.165	-0.613*
	(3.491)	(2.639)	(1.642)	(0.736)	(0.445)	(0.351)
Constant	50.04**	16.09	8.881	9.495	11.35***	27.27***
	(22.50)	(22.76)	(14.50)	(8.348)	(4.056)	(6.912)
Pseudo R^2	0.0445	0.0907	0.0519	0.0573	0.0601	0.1062
Observations	1490	1490	1490	1490	1490	1490

Standard errors in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$; Dummy variable base categories are as follows, Health: Fair, Geotype: Rural Formal, Race: White, Marital Status: Never Married, Gender: Female

Following from the legacy of economic marginalisation mentioned in the Background section, we explicitly tested the hypothesis that there is a discernible difference in the wealth accumulation of those that reside in the tribal authority areas and those that reside in the other geotype categories. In each quantile, we tested the null hypothesis that the three coefficients of the geotype variable were equal. The two possible null hypotheses being tested are as follows:

$$(1)\ \ -\text{geotype4} + \text{geotype2} = 0$$

$$(2)\ \ \text{geotype2} - \text{geotype3} = 0$$

The Wald test indicates that the coefficients of the geographical variables are statistically different from each other, except in the 25th quantile of the Geotype regression and the 5th, 25th and 50th quantile of the final estimation. This echoes previous findings in all the regression models estimated and the hypotheses stated in the background literature. The Wald test for 25th quantile in both models yields a failure to reject the null hypothesis, possibly indicating some of the overlap found in the pairwise correlation of Tribal Authority Area and Urban Formal (Shown in Appendix Two). The geotype variables are particularly significant for the estimation of the wealth distribution in NIDS Wave 2. Specifically, the Tribal Authority Area has an unidentified effect, particularly with reference to land ownership and the dual land tenure system.

Table Seven: Geotype Regression Post-estimation Test

Wald test of Linear Hypothesis for the Geotype Regression						
	0.05	**0.10**	**0.25**	**0.50**	**0.75**	**0.95**
F(2, 2841)	22.95	2183.40	0.98	54.49	44.87	92.97
Prob>F	0.0000	0.000	0.3752	0.0000	0.0000	0.0000

Table Eight: Final Regression Model Post-estimation Test

Wald test of Linear Hypothesis for the Final Estimation						
	0.05	**0.10**	**0.25**	**0.50**	**0.75**	**0.95**
F(2,1466)	0.52	19.96	0.90	0.67	4.04	5.71
Prob>F	0.5924	0.0000	0.4053	0.5112	0.0177	0.0034

V. Conclusion

This investigation sought to explore the determinants of wealth in South Africa, utilising Wave 2 of NIDS. In particular, we evaluated the significance of the lifecycle hypothesis in the distribution and attempted to identify the existence of different wealth accumulation behaviour in tribal authority areas. The focus on tribal authority areas was as a result of the extensive racially based marginalisation of households in that geographic location. Quantile regression methods were utilised in order to account for different wealth accumulation behavior apparent at different points in the household wealth distribution,

The portfolio (assets and liabilities) composition across the age distribution appeared to indicate some evidence that the lifecycle hypothesis is not present in the data. Furthermore the distribution of assets and liabilities across the 4 geotypes specified indicated the presence of a hidden effect within Tribal Authority Areas. While real estate assets largely dominated most of the geographic locations, TAAs did not exhibit the corresponding real estate debt required to finance such asset accumulation.

It appears that wealth accumulation in South Africa cannot be explained utilizing only the life cycle hypothesis. This is evidenced by the insignificance of the age variables in both the quadratic and quartic specification, except possibly for those households with very high levels of net worth. This is particularly true in rural households where the household structure does not require young and middle age adults to save for retirement. Rather, these adults may depend on the older household head. Furthermore, income and asset accumulation may be uncertain in poor households and as such these households may only save when they have surplus income. The presence of non-linearity in the wealth distribution is not easily explained by the estimation and more in-depth investigation is required. Perhaps the inclusion of time preferences, risk and savings attitudes in future waves of NIDS would allow for a more comprehensive understanding of the effects at work., but for NIDS Wave 2, this can only be conjectured and represents an important form of omitted variable bias in the models estimated in this paper.

It must be noted that these results do not indicate that the LCH predictions are violated in the South African wealth distribution. The LCH model appears to apply to particular quantiles of the population, that is, the wealthier households and the particularly indebted households. In particular, the results found these to be households with younger heads, which align with LCH predictions. Poorer households, or those whose assets and liabilities are approximately equal do not appear to accumulate wealth in the same manner as their upper and lower quantile counterparts. However, we cannot formally identify the LCH econometrically at a particular quantile. Rather, the LCH must be present in the overall distribution of wealth. The multi-modality of the household wealth distribution, the failure of the predicted age effects and the failing of the OLS estimation method suggest that the South African wealth distribution does not conform to the LCH at this point in time.

Tribal Authority Areas coefficients were statistically significant in almost every quantile and estimation in which it appears. Furthermore, both the descriptive and analytical results indicated an anomaly in the distribution of wealth over the geographical areas. The effect cannot be identified within this investigation. However, the evidence for this hidden effect is present throughout the findings. These outcomes are only the stepping-stones to identifying the effect of communal land ownership, the legacy of discriminatory apartheid laws and chieftain rule.

No major conclusions can be made from the investigation into these effects beyond the implication of a different coefficient in tribal authority areas due to the fact that data are observational only, rather than set up specifically to identify different wealth behavior resulting from different land tenure arrangements in TAAs. In order to conclude any strong effects, the wealth questions must continue to be asked in the NIDS household surveys, so that we may have panel data with which to compare households over time and see if the effect persists.

Bibliography

Adams, M. Cousins, B. Manona, S. 1999. Land Tenure and Economic Development in Rural South Africa: Constraints and Opportunities. *Working Paper 125.*

Aron, J & Muellbauer, J 2006 Estimates of Household Sector Wealth for South Africa 1970-2003 *Review of Income and Wealth Series 52 Number 2.*

Aron, J & Muellbauer, J. 2013 Wealth, Credit Conditions, and Consumption: Evidence From South Africa. *Review of Income and Wealth*. Series 59.

Aspromourgos, T. 2014. Thomas Piketty, The Future of Capitalism and the Theory of Distribution: A Review essay. *Metroeconomica*. 66:2.

Billor, N., Hadi, A. S., & Velleman, P. F. 2000. BACON: blocked adaptive computationally efficient outlier nominators. *Computational Statistics & Data Analysis*, *34*(3), pp 279-298.

Boskin, M 1991 The Issue of Measurement and Interpretation of Saving and Wealth *Fifty Years of Economic Measurement: The Jubilee of the Conference on Research in Income and Wealth pp 159-184.*

Cagetti, M & De Nardi, M 2008 Wealth Inequality: Data and Models *Macroeconomic Dynamics* 12 (Supplement 2) 2008 285–313.

Cousins, B. 2007. More Than Socially Embedded: The Distinctive Character of 'Communal Tenure' Regimes in South Africa and its Implications for Land Policy. *Journal of Agrarian Change*. Vol. 7 No. 3. pp. 281–315.

Daniels,. R. Finn, A & Msundwa. S. 2014. Wealth data quality in the National Income Dynamics Study Wave 2. *Development Southern Africa*. 31:1. pp 31-50.

Davies, J. & Shorrocks, A. F. 2000. The distribution of wealth. ch. 11. p. 605-675 in A.B., Atkinson and F., Bourguignon eds. *Handbook of Income Distribution*. vol. 1. Elsevier.

Department of Trade and Industry. n.d. *South Africa's Economic Transformation: A Strategy for Broad-Based Black Economic Empowerment department of trade and industry*. Pretoria: Department of Trade and Industry.

Di Matteo, L. 2010. Age Wealth and Portfolio Composition: A Quantile Approach. Paper Prepared for the Meetings of the Canadian Economics Association, Quebec City. [Draft].

Fay, D. 2009. Land Tenure Land Use and Land Reform at Dwesa-Cwebe South Africa: Local Transformations and the Limits of the State. *World Development.* Vol. 37. No. 8. pp. 1424–1433.

Feinberg, H. 1995. South Africa and Land Ownership: What's in a Deed?. *History in Africa.* Vol 22. pp 439-443.

Friedline, T., Masa, R., & Chowa, G. (2015). Transforming wealth: Using the inverse hyperbolic sine (IHS) and splines to predict youth's math achievement. *Social Science Research*, 49, 264-287.

Hadi, A. S. 1992. Identifying multiple outliers in multivariate data. *J. R. Statist. Soc*. B, 54, 761-771.

Hargreaves, JR; Morison, LA; Gear, JS; Kim, JC; Makhubele, MB; Porter, JD; Watts, C; Pronyk, PM; (2007) Assessing household wealth in health studies in developing countries: a comparison of participatory wealth ranking and survey techniques from rural South Africa. *Emerg Themes Epidemiol*, 4. p4.

Heady, B. Warren, D. & Wooden, M. 2008. The structure and distribution of household wealth in Australia: cohort differences and retirement issues. *Social Policy Research Paper No. 33*. Melbourne Institute of Applied Economic and Social Research. University of Melbourne.

Hoeks, C. Azadi, H. Khachak, P. Troyo-Dieguez, E. Van Passel, S & Witlox, F. 2014. Reforming land-Tenure Systems in South Africa: Routes to Socio-Economic and Agricultural Sustainability. *Development Policy Review*. 32 (6). pp 647-674.

Home, R. 2013. Culturally Unsuited to Property Rights? Colonial Land Laws and African Societies. *Journal of Law and Society.* Volume 40. No. 3. pp 403 – 419.

Glenn, N. D. 2005. Cohort Analysis. Second Edition. *Sage University Paper Series on Quantitative Applications in Social Sciences*. 07-005. Beverly Hills, CA: Sage.

Juster, F. Smith, J. Stafford, F. 1999. The Measurement and Structure of Household Wealth. *Labour Economics.* Issue 6. pp 253-27.

Keegan, T. 1986. The Dynamics of Rural Accumulation in South Africa: Comparative and Historical Perspectives. *Comparative Studies in Society and History*, Vol. 28, No. 4, pp. 628-650.

Koenker, R & Hallock, K. 2001. Quantile Regression: An Introduction. *Journal of Economic Perspectives*. 15(4). pp 143-156.

Koenker, R & Portnoy, S. 1999. Quantile Regression. Preprint.

Maitra, P. & Vahid. F. 2005. The Effect of Household Characteristics on Living Standards in South Africa 1993 - 98: A Quantile Regression Analysis with Sample Attrition. *Working Papers in Economics and Econometrics*. Working Paper No. 452.

Matilla-wiro, P. 1999. Economic Theories of the Household:A Critical Review. *World Institute for Development Economics Research.* The United Nations University. Working Papers No. 159.

Meng, X. 2007. Wealth Accumulation and Distribution in Urban China. *Economic Development and Cultural Change*. Vol. 55. No. 4. pp. 761-791.

McDonald, S. & Piesse, J. 1999. Legacies of Apartheid: The Distribution of Income in South Africa. *Journal of International Development*. pp 985-1004.

Milanovic, B. 2014. The Return of "Patrimonial Capitalism": A Review of Thomas Picketty's Capital in the Twenty First Century. *Journal of Economic Literature.* 52(2). pp 519–534.

Modigliani, F. & Brumberg, R. 1954. Utility Analysis and the Consumption Function: an interpretation of cross-section data," in Kenneth K. Kurihara, ed., *PostKeynesian Economics*, New Brunswick, NJ. Rutgers University Press. Pp 388–436.

Modigliani, F. 1986. Lifecycle Individual Thrift and the Wealth of Nations. *The American Economic Review*. Vol. 76. No. 3. pp 297-313.

Murphy, K. & Welch, F. 1990. Empirical Age-Earnings Profiles. *Journal of Labor Economics*, Vol. 8, No. 2. pp. 202-229.

Nattrass, N & Seekings, J. 2001. Democracy and Distribution in Highly Unequal Economies: The Case of South Africa. *The Journal of Modern African Studies.* Vol 39. No.3. pp 471-498.

Ngwenya, A. & Paas, L. 2012. Lifecycle effects on consumer financial product portfolios in South Africa: An exploratory analysis of four ethnic groups. *Journal of Economic Psychology*. 33. pp 8–18.

Ramutsindela, M. 2012. Property rights land tenure and the racial discourses. *GeoJournal.* Issue 77. pp 753-763.

Skopek, N. Buchholz, S & Blossfeld, H. 2014. National patterns of income and wealth inequality. *International Journal of Comparative Sociology*. Vol. 55(6). pp 463–488.

South Africa (2014). 2011 Tax Statistics.
http://www.treasury.gov.za/publications/tax%20statistics/2012/2011%20Tax%20Statistics.pdf
Accessed: 16 April 2015

van Rooij, M. Lusardi, A & Alessie, R. 2012. Financial Literacy Retirement Planning and Household Wealth. *The Economic Journal* .122. pp 449-478.

Van der Berg, S. 2011. Current poverty and income distribution in the context of South African history. *Economic History of Developing Regions*. 26:1. pp 120-140.

Zimmer, Z. & Das, S. 2014. The Poorest of the Poor: Composition and Wealth of Older Person Households in Sub-Saharan Africa. *Research on Aging.* Vol 36(3). pp 271-296.

Data Source

Southern Africa Labour and Development Research Unit. National Income Dynamics Study 2010-2011, Wave 2 [dataset]. Version 2.2. Cape Town: Southern Africa Labour and Development Research Unit [producer], 2012. Cape Town: DataFirst [distributor], 2012.

Appendix 1: Outlier Detection Methods - *Bacon* vs*Hadimvo*

Hadimvo is an outlier identification algorithm in STATA, which identifies multiple outliers in multivariate data using the method of Hadi (1992, 1994), creating a new variable equal to 1 if an observation is an "outlier" and 0 otherwise. The process for detecting the outliers is outlined in Hadi (1992) (Hadi, 1992). For an nxp data matrix, the n observations are ordered using a robust measure of outlyingness. The data set is then divided into two subsets; one being the p+1 'good' observations and the other being the n-p-1 'bad' observations (Hadi, 1992). Then the relative distance from each point in the data set to the centre of the former data set is computed, relative to the covariance matrix of that subset (Hadi, 1992). Next, the n observations are rearranged in ascending order and the dataset is once again divided into two subsets; one with p+2 'good' observations and the other with n-p-2 'bad' observations (Hadi, 1992). This process is repeated until the stopping criterion is met and the final subset of 'bad' observations is declared as outliers (Hadi, 1992).

The *bacon* outlier detection method identifies multiple outliers in multivariate data using the blocked adaptive computationally efficient outlier nominators (*BACON*) algorithm proposed by Billor, Hadi, and Velleman (2000) (Billor et al, 2000). It creates a dummy variable, which is equal to 1 if an observation is an outlier and equal to 0 otherwise. Furthermore, a second variable can be created which contains the distances from the basic subset. *Bacon* establishes an improvement of *hadimvo*, which also is used to identify multivariate outliers, but is much less computationally intensive and provides similar results. *Hadimvo* tends to take hours or sometimes days to indicate outliers, particularly in large datasets.

The idea behind *bacon* is that m outlier-free observations must be identified in the dataset of n observations and over p variables. The Mahalanobis distance is used as the criterion, since it is scale-invariant, especially when dealing with variables of different magnitudes and units (Billor et al, 2000). The tables below indicate the results of using *hadimvo* and *bacon* respectively. It appears that *bacon* is stricter in its identification of outliers than hadimvo, although *hadimvo* does not allow for outlier detection of the one shot net worth variable, as a singular matrix is created. Although *bacon* and *hadimvo* are meant to be perfect substitutes, the differences in the variables are quite significant. For example, financial assets have 550 outliers when utilizing, but only 99 outliers when *Bacon* is employed. Conversely, for retirement assets, *hadimvo* indicated 70 outliers, whereas *bacon* indicated 142 outliers.

Table Nine: Outlier Results

	Hadimvo			Bacon	
Variable	**Before**	**After**	**Outliers**	**After**	**Outliers**
Income	3516	3349	167	3482	34
Net Worth	2966	2577	389	2909	57
One-Shot Net Worth	1865	Collinearity issues		1730	135
Total Assets	2571	2220	351	2519	52
Total Debts	812	679	133	765	47
Real Estate Debt	125	125	0	125	0
Business Debt	13	11	2	12	1
Vehicle Debt	61	56	5	61	0
Financial Debt	735	690	45	725	10
Real Estate Assets	1740	1444	296	1703	37
Business Assets	59	30	29	6	53
Vehicle Assets	373	352	21	368	5
Financial Assets	1604	1341	263	1552	52
Retirement Annuity Assets	73	36	37	72	1
Livestock Assets	95	89	6	94	1

Appendix 2: OLS Regression Diagnostics

The 6 OLS Assumptions are presented below in order to validate the use of the OLS estimation model. Regression diagnostic tests were performed to test each assumption of each model. Each test is detailed under the assumptions below. The regression models being tested are the OLS regression with the age quadratic specification and the three OLS regressions with Age Cohort Specifications. The Four OLS models estimated were found to be misspecified as the assumptions are not all met. In particular, the assumptions of homoscedasticity, normality of the residuals, linearity in parameters and the zero conditional mean were violated.

Assumption 1: No Perfect Collinearity
In the sample, none of the independent variables are constant and exhibit no exact linear relationships.

The assumption was tested, by creating a pairwise correlation table of all the explanatory variables in each model. For the assumption to hold, the independent variables must not be closely correlated in a negative or positive manner. The only values of concern are the relationships between African and Coloured, Age and Age^2, and Urban Formal and Tribal Authority Area. African and Coloured are highly negatively correlated at -0.809, however, this correlation is to be expected between the two races. Similarly for Age and Age^2, a high correlation value is expected (0.983) as they are perfectly non-linearly related. Urban Formal and Tribal Authority Area are negatively correlated at -0.682, which is not an unexpected result, as geographical locations have areas of overlap. The results in Tables Ten through Thirteen below showed that there are no exact or close to exact linear relationships between the explanatory variables and thus the assumption is not violated.

Table Ten: Pairwise Correlation of Covariates for Age Quadratic Specification

	Gender	Age	Age²	Education	African	Coloured	Asian	Children	Married	Living with Partner	Widowed	Divorced	Urban Informal	Tribal Authority Area	Urban Formal	Excellent Health	Very Good Health	Good Health	Poor Health
Gender	1																		
Age	-0.115	1																	
Age²	-0.116	0.983	1																
Education	0.105	-0.392	-0.376	1															
African	-0.048	-0.076	-0.062	-0.099	1														
Coloured	-0.018	0.010	-0.000	-0.067	-0.809	1													
Asian	0.043	0.012	0.006	0.059	-0.196	-0.034	1												
Children	0.046	-0.096	-0.137	-0.122	0.023	0.001	0.005	1											
Married	0.196	0.119	0.080	0.005	-0.126	0.042	0.076	0.150	1										
Living with Partner	0.085	-0.122	-0.127	-0.034	-0.015	0.052	-0.012	0.016	-0.178	1									
Widowed	-0.273	0.480	0.491	-0.226	0.027	-0.020	0.003	-0.094	-0.300	-0.137	1								
Divorced	-0.060	0.085	0.066	0.029	-0.114	0.081	-0.020	-0.0257	-0.147	-0.067	-0.113	1							
Urban Informal	0.026	-0.070	-0.071	-0.016	0.096	-0.066	-0.025	0.014	-0.068	0.0523	-0.048	0.034	1						
Tribal Authority Area	-0.166	0.193	0.196	-0.185	0.352	-0.283	-0.069	0.120	0.015	-0.079	0.178	-0.067	-0.211	1					
Urban Formal	0.076	-0.096	-0.096	0.228	-0.322	0.244	0.049	-0.113	0.029	-0.029	-0.089	0.066	-0.247	-0.682	1				
Excellent Health	0.101	-0.282	-0.272	0.205	-0.021	0.008	0.011	0.052	0.025	0.037	-0.171	-0.003	-0.010	-0.091	0.085	1			
Very Good Health	0.022	-0.096	-0.096	0.042	0.011	-0.003	-0.031	-0.009	0.002	0.002	-0.049	-0.011	-0.008	0.010	-0.034	-0.410	1		
Good Health	-0.018	0.135	0.128	-0.075	'-0.003	-0.013	0.024	-0.005	0.006	-0.020	0.048	0.014	0.008	0.033	-0.024	-0.368	-0.353	1	
Poor Health	-0.064	0.171	0.166	-0.172	0.026	-0.017	0.008	-0.025	-0.022	-0.014	0.125	0.004	0.029	0.035	-0.020	-0.163	-0.156	-0.140	1

Table Eleven: Pairwise Correlation of Covariates for Age Cohort 15 to 34 Specification

	Gender	Age Cohort (15 to 34)	Education	African	Coloured	Asian	Children	Married	Living with Partner	Widowed	Divorced	Urban Informal	Tribal Authority Area	Urban Formal	Excellent Health	Very Good Health	Good Health	Poor Health
Gender	1																	
Age Cohort (15 to 34)	0.0268	1																
Education	0.1046	0.0632	1															
African	-0.0478	-0.0311	-0.0987	1														
Coloured	-0.0179	0.0496	-0.0670	-0.8089	1													
Asian	0.0428	0.0221	0.0590	-0.1963	-0.0343	1												
Children	0.0462	0.2748	-0.1222	0.0229	0.0005	0.0053	1											
Married	0.1956	0.1169	0.0049	-0.1262	0.0420	0.0764		1										
Living with Partner	0.0845	0.0761	-0.0338	-0.0146	0.0519	-0.0115	0.0163	-0.1783	1									
Widowed	-0.2726	-0.1958	-0.2259	0.0271	-0.0196	0.0028	-0.0935	-0.2997	-0.1370	1								
Divorced	-0.0600	0.0037	0.0285	-0.1139	0.0805	-0.0202	-0.0257	-0.1466	-0.0670	-0.1126	1							
Urban Informal	0.0263	-0.0189	-0.0175	0.0962	-0.0663	-0.0250	0.0142	-0.0678	0.0527	-0.0482	0.0337	1						
Tribal Authority Area	-0.1656	-0.0744	-0.1850	0.3519	-0.2834	-0.0691	0.1195	0.0154	-0.0791	0.1780	-0.0672	-0.2111	1					
Urban Formal	0.0758	0.0467	0.2281	-0.3224	0.2439	0.0493	-0.1133	0.0288	-0.0288	-0.0893	0.0657	-0.2470	-0.6815	1				
Excellent Health	0.1014	0.0806	0.2052	-0.0210	0.0082	0.0113	0.0521	0.0249	0.0371	-0.1705	-0.0030	-0.0099	-0.0909	0.0845	1			
Very Good Health	0.0222	0.0379	0.0423	0.0107	-0.0031	-0.0310	-0.0086	0.0023	0.0021	-0.0490	-0.0114	-0.0077	0.0100	-0.0343	-0.4101	1		
Good Health	-0.0179	-0.0482	-0.0749	-0.0029	-0.0126	0.0240	-0.0054	0.0060	-0.0198	0.0475	0.0136	0.0075	0.0334	-0.0243	-0.3681	-0.3530	1	
Poor Health	-0.0637	-0.0468	-0.1719	0.0256	-0.0171	0.0079	-0.0252	-0.0220	-0.0141	0.1247	0.0037	0.0287	0.0347	-0.0196	-0.1627	-0.1561	-0.1401	1

Table Twelve: Pairwise Correlation of Covariates for Age Cohort 35 to 48 Specification

	Gender	**Age Cohort (35 to 48)**	**Education**	**African**	**Coloured**	**Asian**	**Children**	**Married**	**Living with Partner**	**Widowed**	**Divorced /Separated**	**Urban Informal**	**Tribal Authority Area**	**Urban Formal**	**Excellent Health**	**Very Good Health**	**Good Health**	**Poor Health**
Gender	1																	
Age Cohort (35 to 48)	0.0839	1																
Education	0.1046	0.3077	1															
African	-0.0478	0.0998	-0.0987	1														
Coloured	-0.0179	-0.0493	-0.0670	-0.8089	1													
Asian	0.0428	-0.0321	0.0590	-0.1963	-0.0343	1												
Children	0.0462	-0.1289	-0.1222	0.0229	0.0005	0.0053	1											
Married	0.1956	-0.2180	0.0049	-0.1262	0.0420	0.0764	0.1499	1										
Living with Partner	0.0845	0.0672	-0.0338	-0.0146	0.0519	-0.0115	0.0163	-0.1783	1									
Widowed	-0.2726	-0.2706	-0.2259	0.0271	-0.0196	0.0028	-0.0935	-0.2997	-0.1370	1								
Divorced /Separated	-0.0600	-0.1129	0.0285	-0.1139	0.0805	-0.0202	-0.0257	-0.1466	-0.0670	-0.1126	1							
Urban Informal	0.0263	0.0627	-0.0175	0.0962	-0.0663	-0.0250	0.0142	-0.0678	0.0527	-0.0482	0.0337	1						
Tribal Authority Area	-0.1656	-0.1055	-0.1850	0.3519	-0.2834	-0.0691	0.1195	0.0154	-0.0791	0.1780	-0.0672	-0.2111	1					
Urban Formal	0.0758	0.0411	0.2281	-0.3224	0.2439	0.0493	-0.1133	0.0288	-0.0288	-0.0893	0.0657	-0.2470	-0.6815	1				
Excellent Health	0.1014	0.2137	0.2052	-0.0210	0.0082	0.0113	0.0521	0.0249	0.0371	-0.1705	-0.0030	-0.0099	-0.0909	0.0845	1			
Very Good Health	0.0222	0.0539	0.0423	0.0107	-0.0031	-0.0310	-0.0086	0.0023	0.0021	-0.0490	-0.0114	-0.0077	0.0100	-0.0343	-0.4101	1		
Good Health	-0.0179	-0.1019	-0.0749	-0.0029	-0.0126	0.0240	-0.0054	0.0060	-0.0198	0.0475	0.0136	0.0075	0.0334	-0.0243	-0.3681	-0.3530	1	
Poor Health	-0.0637	-0.1257	-0.1719	0.0256	-0.0171	0.0079	-0.0252	-0.0220	-0.0141	0.1247	0.0037	0.0287	0.0347	-0.0196	-0.1627	-0.1561	-0.1401	1

Table Thirteen: Pairwise Correlation of Covariates for Age Cohort 49 to 104 Specification

Column1	Gender	Age Cohort (35 to 48)	Education	African	Coloured	Asian	Children	Married	Living with Partner	Widowed	Divorced	Urban Informal	Tribal Authority Area	Urban Formal	Excellent Health	Very Good Health	Good Health	Poor Health
Gender	1																	
Age Cohort (35 to 48)	-0.0987	1																
Education	0.1046	-0.3565	1															
African	-0.0478	-0.0592	-0.0987	1														
Coloured	-0.0179	-0.0024	-0.0670	-0.8089	1													
Asian	0.0428	0.0079	0.0590	-0.1963	-0.0343	1												
Children	0.0462	-0.1584	-0.1222	0.0229	0.0005	0.0053	1											
Married	0.1956	0.0851	0.0049	-0.1262	0.0420	0.0764	0.1499	1										
Living with Partner	0.0845	-0.1301	-0.0338	-0.0146	0.0519	-0.0115	0.0163	-0.1783	1									
Widowed	-0.2726	0.4212	-0.2259	0.0271	-0.0196	0.0028	-0.0935	-0.2997	-0.1370	1								
Divorced	-0.0600	0.0967	0.0285	-0.1139	0.0805	-0.0202	-0.0257	-0.1466	-0.0670	-0.1126	1							
Urban Informal	0.0263	-0.0371	-0.0175	0.0962	-0.0663	-0.0250	0.0142	-0.0678	0.0527	-0.0482	0.0337	1						
Tribal Authority Area	-0.1656	0.1607	-0.1850	0.3519	-0.2834	-0.0691	0.1195	0.0154	-0.0791	0.1780	-0.0672	-0.2111	1					
Urban Formal	0.0758	-0.0790	0.2281	-0.3224	0.2439	0.0493	-0.1133	0.0288	-0.0288	-0.0893	0.0657	-0.2470	-0.6815	1				
Excellent Health	0.1014	-0.2643	0.2052	-0.0210	0.0082	0.0113	0.0521	0.0249	0.0371	-0.1705	-0.0030	-0.0099	-0.0909	0.0845	1			
Very Good Health	0.0222	-0.0829	0.0423	0.0107	-0.0031	-0.0310	-0.0086	0.0023	0.0021	-0.0490	-0.0114	-0.0077	0.0100	-0.0343	-0.4101	1		
Good Health	-0.0179	0.1351	-0.0749	-0.0029	-0.0126	0.0240	-0.0054	0.0060	-0.0198	0.0475	0.0136	0.0075	0.0334	-0.0243	-0.3681	-0.3530	1	
Poor Health	-0.0637	0.1549	-0.1719	0.0256	-0.0171	0.0079	-0.0252	-0.0220	-0.0141	0.1247	0.0037	0.0287	0.0347	-0.0196	-0.1627	-0.1561	-0.1401	1

Assumption 2: Homoskedasticity of the Residuals

The variance of the residuals is constant.

Employing the Breusch-Pagan test for homoskedasticty in residuals tested the assumption of constant variance. In order for the variance of the residuals to be determined as constant, the test must fail to reject the null hypothesis of constant variance. The p-values of each test are quite small and close to zero in Table Fourteen below indicates that the null hypothesis is rejected and the assumption is violated. The residuals appear to be heteroskedastic in nature, indicating that they exhibit non-constant variance in all the OLS regression models estimated.

Ho: Constant variance

Variables: fitted values of IHS net worth

Table Fourteen: Breusch-Pagan test for Homoskedasticty of Residuals

	Age Quadratic	Age Cohort 15 to 34	Age Cohort 35 to 48	Age Cohort 49 to 104
chi(1)	9.38	9.22	10.45	9.73
Prob>F	0.0022	0.0024	0.0012	0.0018

Assumption 3: Normality of Residuals

The residuals are normally distributed.

The normality of the residuals is tested by looking at a kernel density plot of the residuals in relation to the normal distribution. Furthermore, qnorm and pnorm plots were created to further emphasise the results of the kernel density plot. These graphs showcase the quantiles of the residual against quantiles of normal distribution and the standard normal probability plot of the residuals respectively. All of the plots for each OLS regression model estimated indicate a significant deviation from normality. Thus the assumption is violated for all four OLS models.

Figure 10: Kernel Density Plot of Residuals (Age Quadratic)

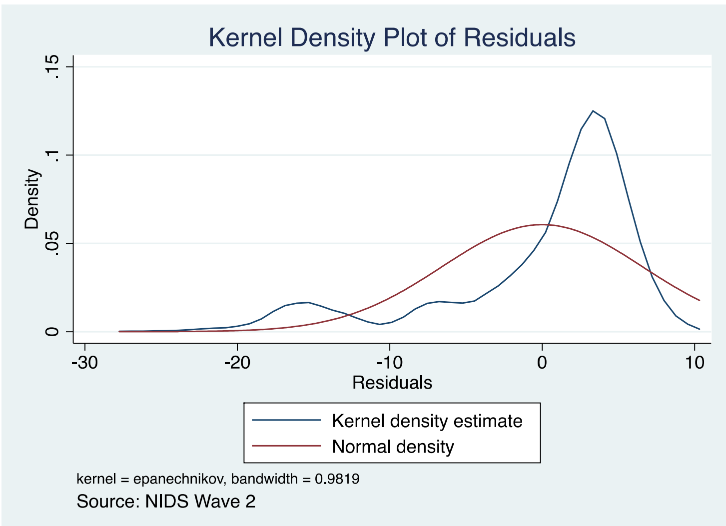


Figure 11: Kernel Density Plot of Residuals (Age Cohort 15 to 34)

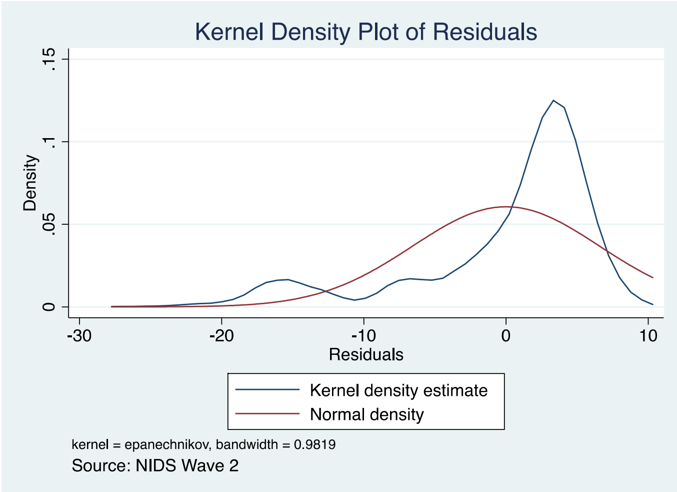


Figure 12: Kernel Density Plot of Residuals (Age Cohort 35 to 48)

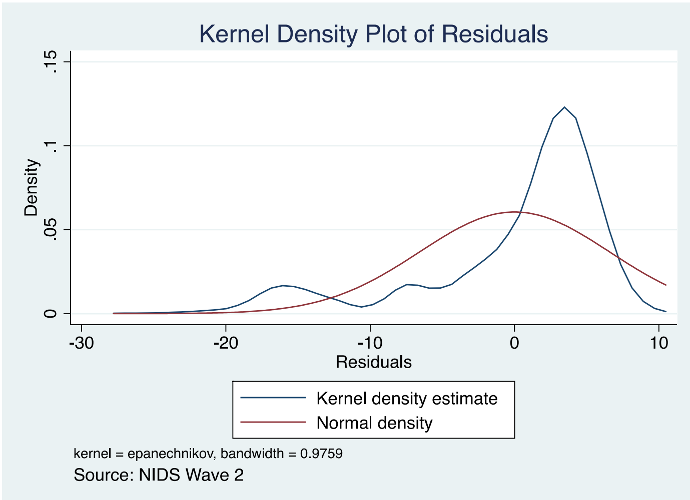


Figure 13: Kernel Density Plot of Residuals (Age Cohort 49 to 104)

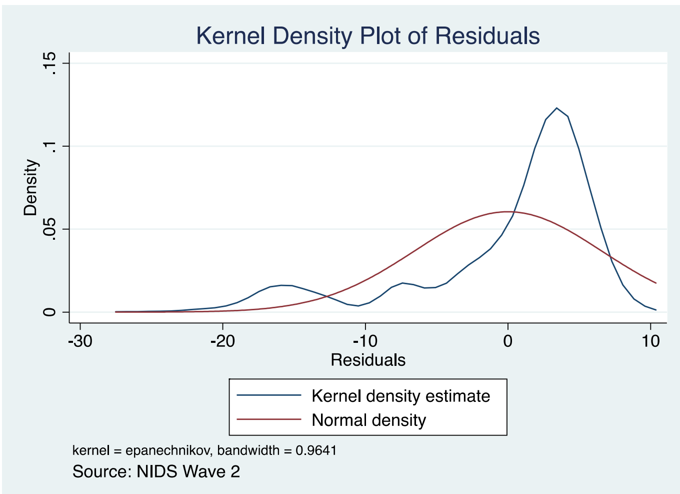


Figure 14: Qnorm Plot of Residuals (Age Quadratic)

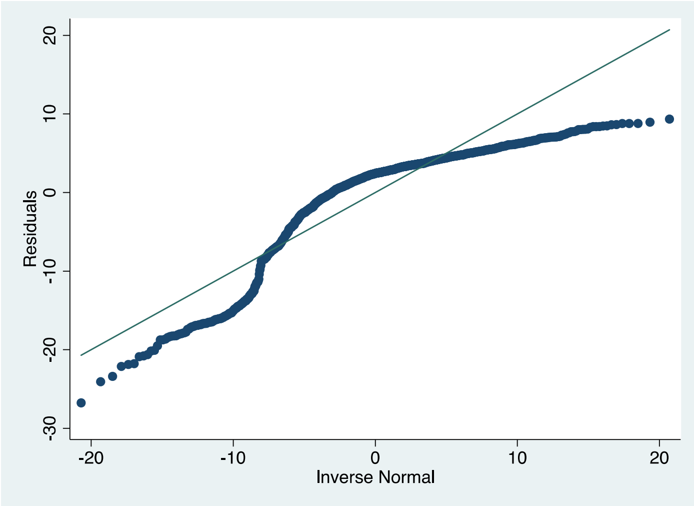


Figure 15: Qnorm Plot of Residuals (Age Cohort 15 to 34)

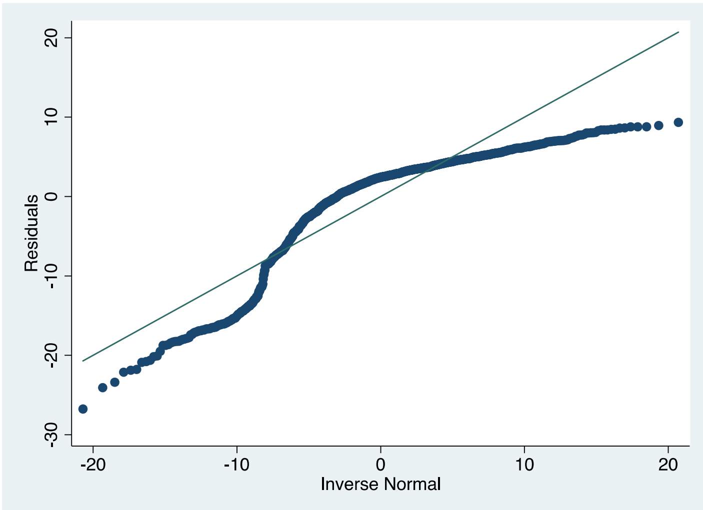


Figure 16: Qnorm Plot of Residuals (Age Cohort 35 to 48)

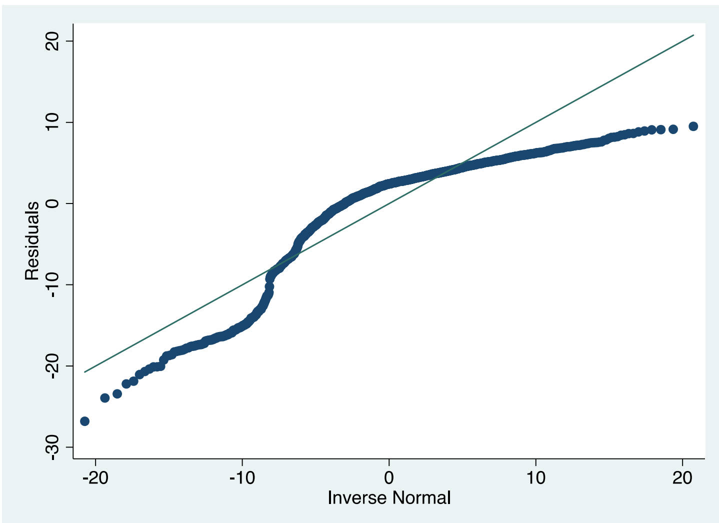


Figure 17: Qnorm Plot of Residuals (Age Cohort 49 to 104)

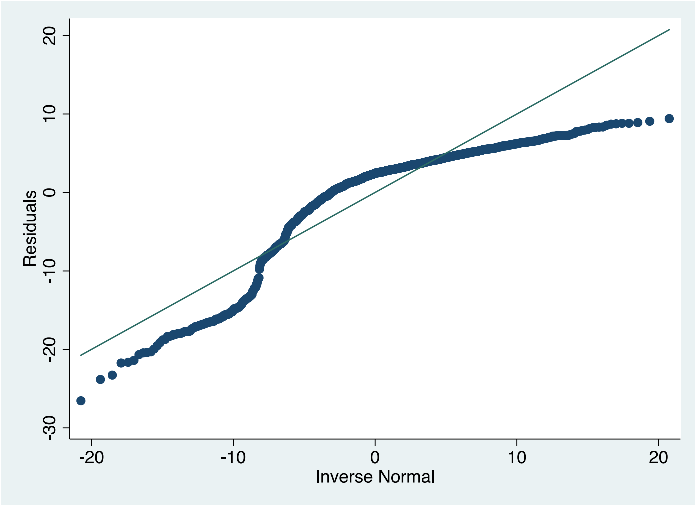


Figure 18: Pnorm Plot of Residuals (Age Quadratic)

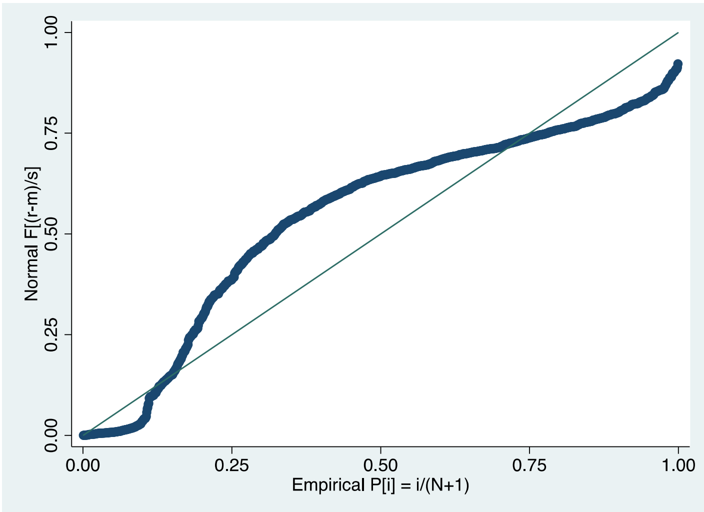


Figure 19: Pnorm Plot of Residuals (Age Cohort 15 to 34)

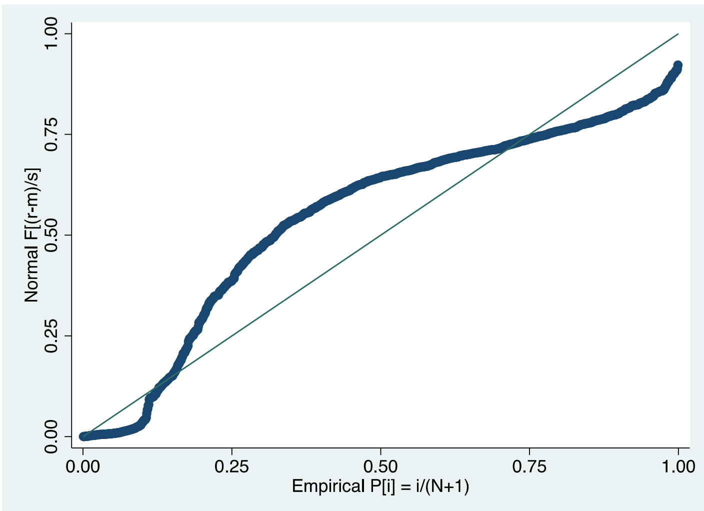


Figure 20: Pnorm Plot of Residuals (Age Cohort 35 to 48)

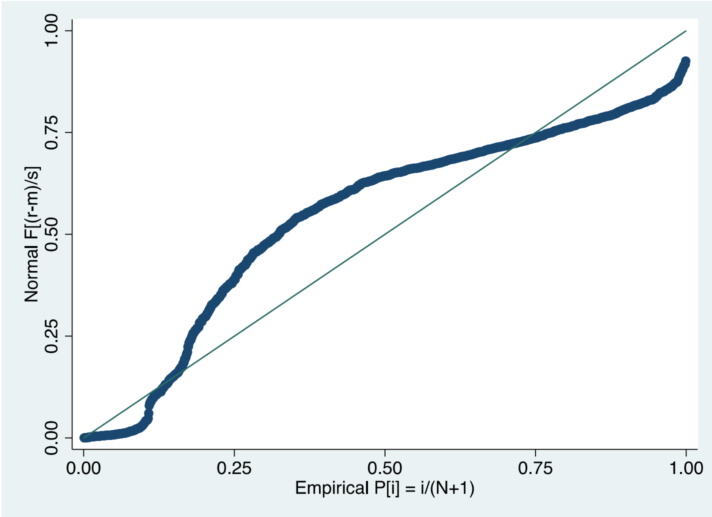
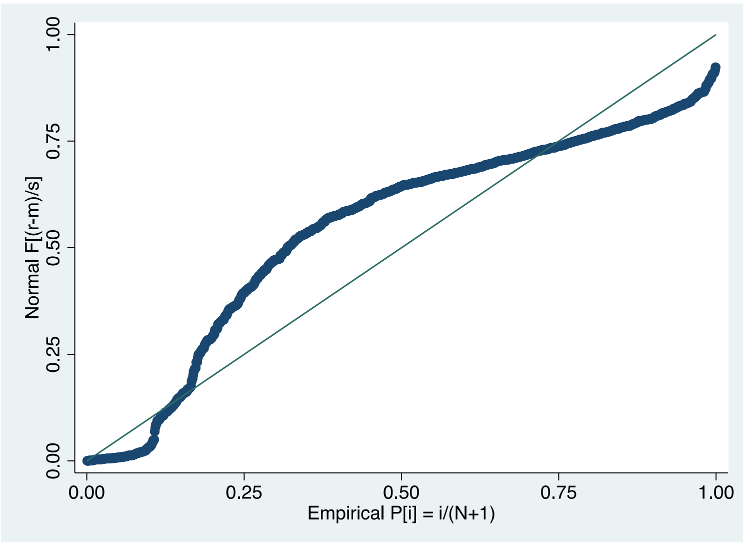


Figure 21: Pnorm Plot of Residuals (Age Cohort 49 to 104)



Assumption 4: Linear in Parameters

The coefficients (parameters) of the model are all linear.

Creating scatter plots of the relationship between the residuals and each non-categorical variable tested the assumption of linearity in the parameters. As shown in Figures 10 through 13, it is evident that none of these plots reveal a linear relationship between the residuals and the independent variables. Thus the assumption is violated and the parameters do not exhibit linearity.

Figure 22: Residuals and Age (Age Quadratic)

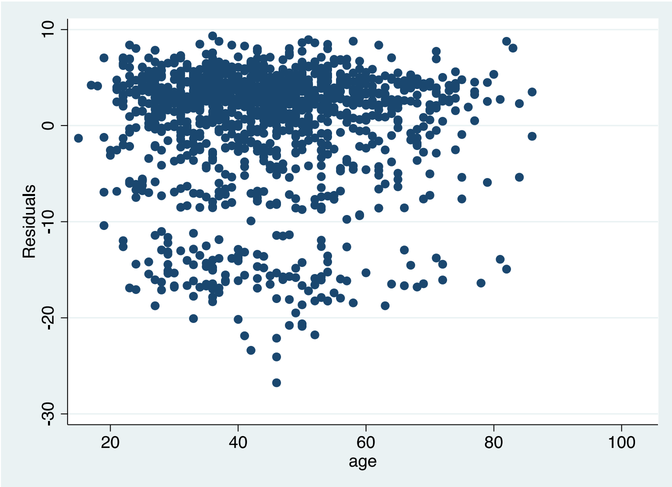


Figure 23: Residuals and Age² (Age Quadratic)

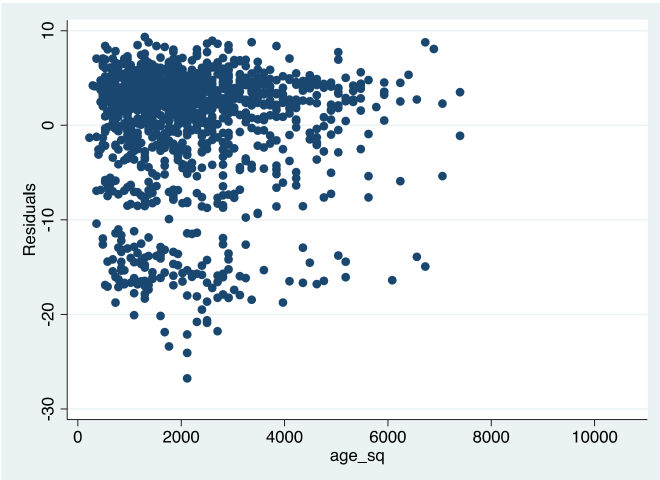


Figure 24: Residuals and Children (Age Quadratic)

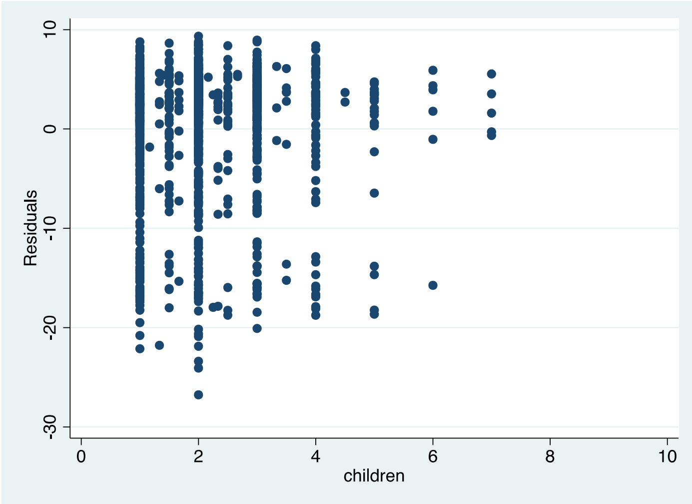


Figure 25: Residuals and Education (Age Quadratic)

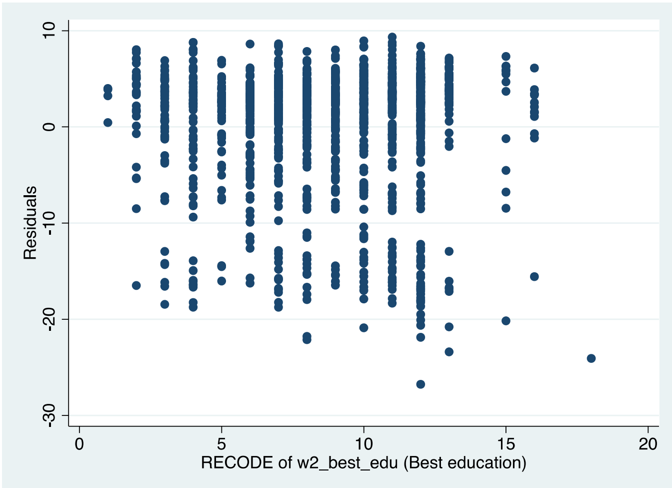


Figure 26: Residuals and Education (Age Cohort 15 to 34)

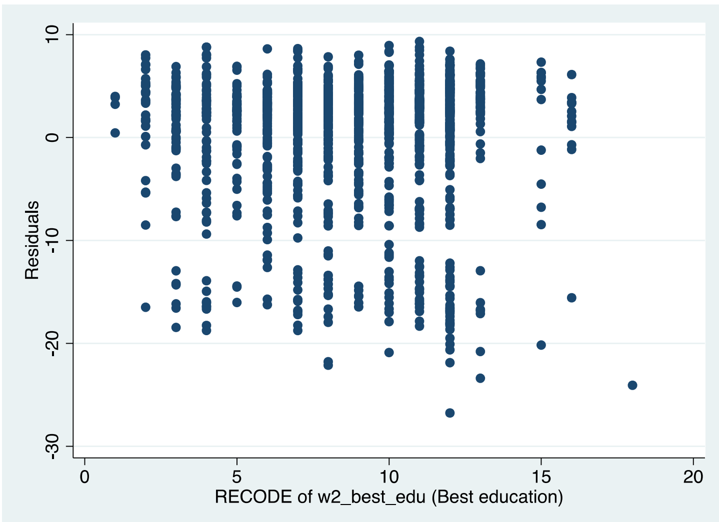


Figure 27: Residuals and Children (Age Cohort 15 to 34)

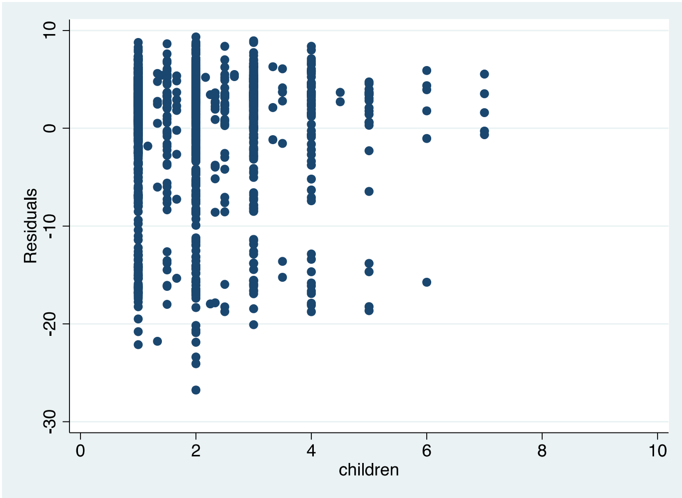


Figure 28: Residuals and Education (Age Cohort 35 to 48)

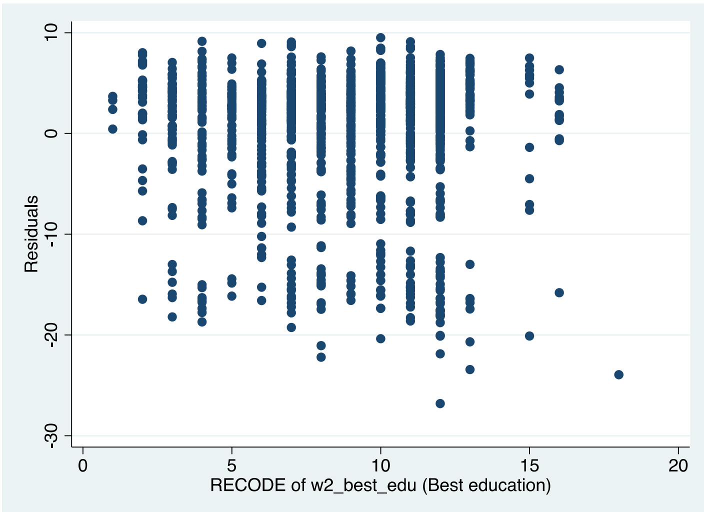


Figure 29: Residuals and Children (Age Cohort 35 to 48)

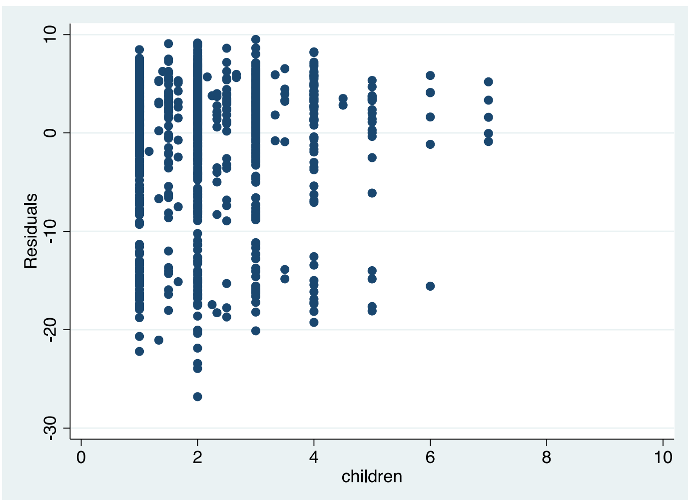


Figure 30: Residuals and Education (Age Cohort 49 to 104)

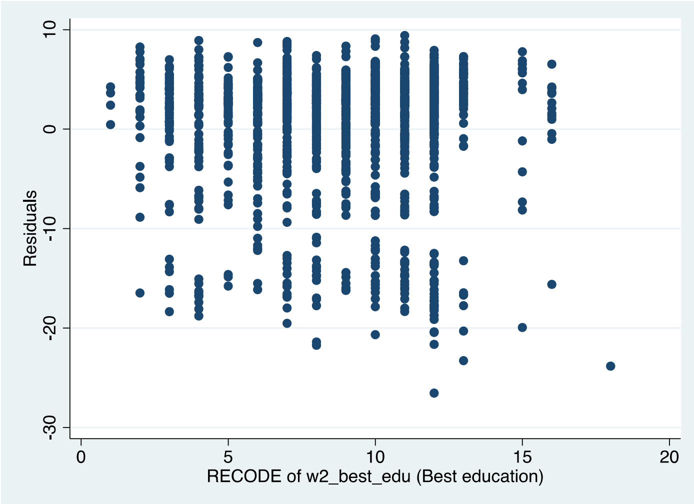
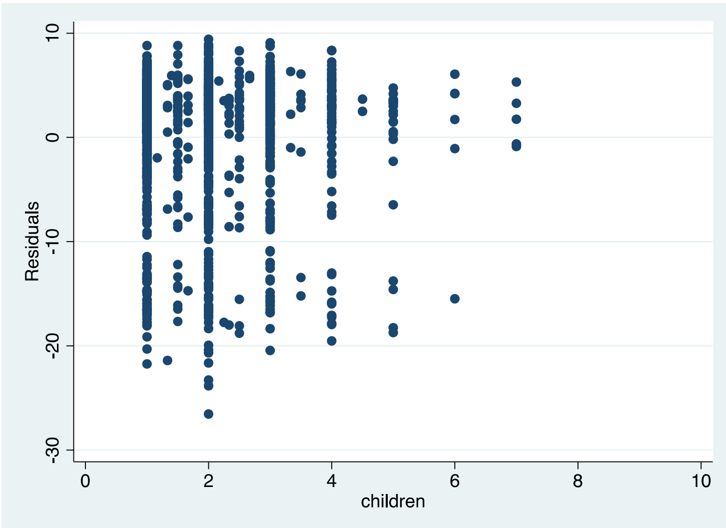


Figure 31: Residuals and Children (Age Cohort 49 to 104)



Assumption 5: Random Sampling

The data is a random sample of the population.

This assumption does not require explicit testing. Although there is attrition present in NIDS, as a result of household nonresponse between waves, the correct weights were applied to correct for this matter. Thus the assumption is not violated.

Assumption 6: Zero Conditional Mean

The error has an expected value of zero, given any value of the explanatory variables.

The testing of this assumption is contentious, as there is no way to truly test the expected value of the error in the sample. We plot the predicted values against the fitted values and in order for the assumption to be met; the values should cluster around zero. Figure 32 through 35 below showed that the values do not cluster around zero, in fact they are closer to ten, as such the assumption is not met.

Figure 32: Testing Zero Conditional Mean Assumption (Age Quadratic)

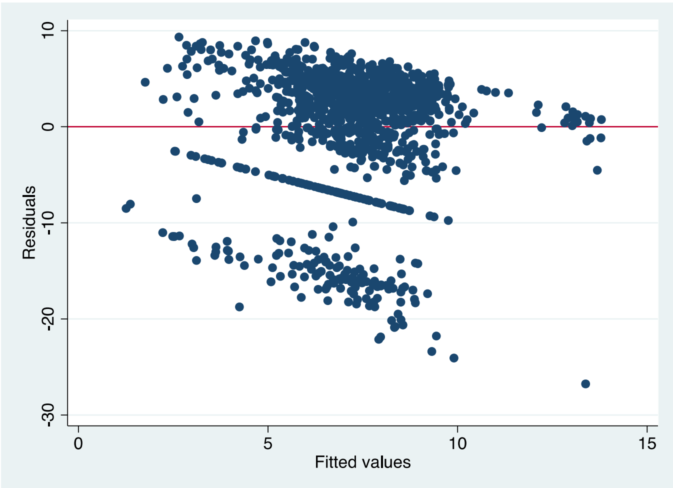


Figure 33: Testing Zero Conditional Mean Assumption (Age Quadratic)

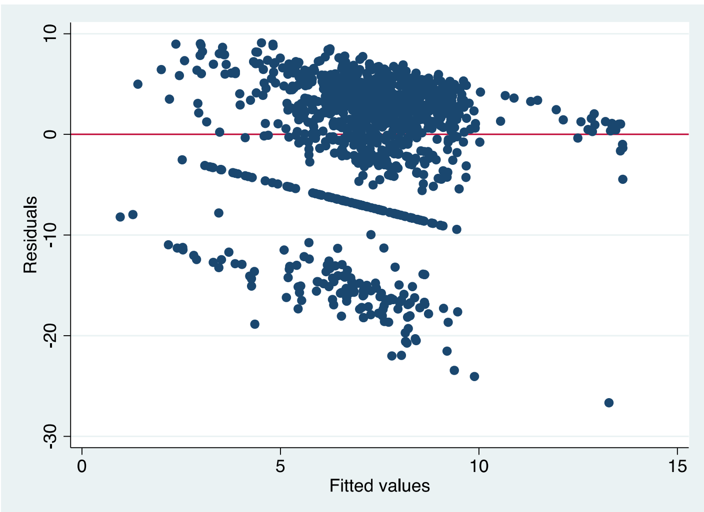


Figure 34: Testing Zero Conditional Mean Assumption (Age Quadratic)

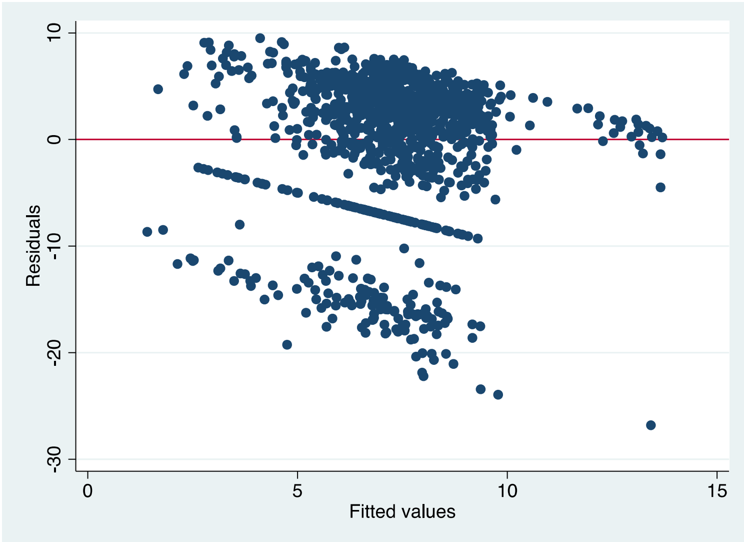
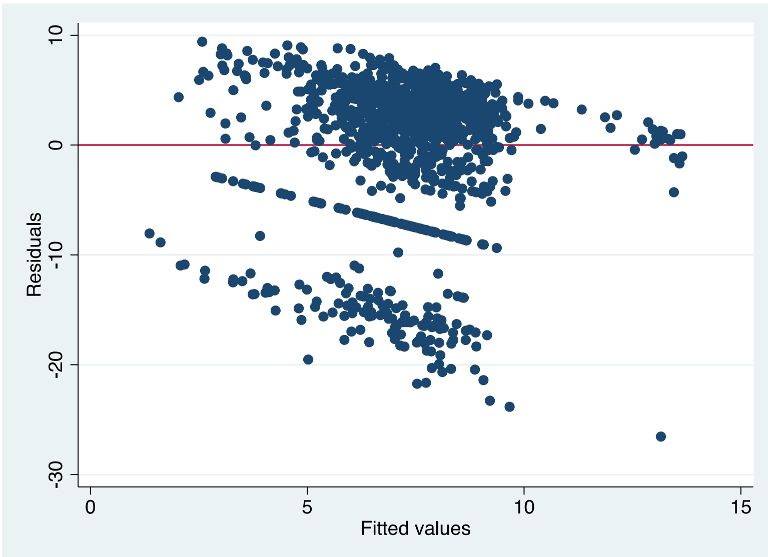


Figure 35: Testing Zero Conditional Mean Assumption (Age Quadratic)



These preceding tests of the OLS assumptions found that not all assumptions were met. In particular, the residuals were found to be heteroskedastic, non-normally distributed, and evidence of non-linearity in the parameters was found. Finally, the zero conditional mean assumption was violated. As such the OLS estimation models are misspecified and inadequate for the estimation of determinants of wealth in South Africa