Research Paper 49 **| 2017**

# EVALUATING THE PROPENSITY TO SAVE IN SOUTH AFRICA USING WEATHER-INCOME RELATIONSHIP

Helena TING, Martina BOZZOLA, Timothy SWANSON

# Evaluating the propensity to save in South Africa using weather-income relationship

Helena Ting* Martina Bozzola† Tim Swanson‡

**Abstract**

We study the propensity to save on permanent and transitory income based on a comprehensive household panel data-set from South Africa. We decompose income into permanent and transitory components, and proxy transitory income by weather deviations from thirty-year normal. We evaluate the propensity to save by OLS and by median regressions. By OLS, we find that the propensity to save on transitory income is not significantly different from one, while the propensity to save on permanent income is significantly different from zero. This finding is in alignment with previous studies. By median regression, we find that the propensity to save by permanent or transitory income are significantly different from either zero or one, meaning that the propensity to save represents some fractions of permanent and transitory income. We also evaluate the propensity to save by income quintiles using median regression. We find that the propensity to save from permanent and transitory income are only significantly different in the highest quintile when durable goods are considered as saving, while not significantly different in the lower 80% of the income distribution. In the top 20%, we find the propensity to save from transitory income is significantly higher than that of permanent income, although it still does not reach one.

*This research is funded by FOODSECURE - An interdisciplinary research project to explore the future of global food and nutrition security. We thank Dr. Jean-Louis Arcand and Dr. Martina Viarengo for helpful comments. We thank the Centre for International Environmental Studies for their ongoing support. We thank the data center at National Income Dynamic Study and European Centre for Medium Range Forecasting for helpful assistance in data access.
Corresponding author. Centre for International Environmental Studies, Graduate Institute of International and Development Studies. Email:
†helena.ting@graduateinstitute.ch
‡School of Biological Sciences, Queen's University Belfast. Email: m.bozzola@qub.ac.uk
Centre for International Environmental Studies, Graduate Institute of International and Development Studies. Email: tim.swanson@graduateinstitute.ch

## 1. Introduction

Weather can affect household saving in multiple ways. It directly impacts income for those sectors where there exists income sensitivity to weather. For example, agricultural income for those households engaged in rain-fed agriculture has been shown to be affected by weather. Weather has also been shown to negatively impact other economic activities that are sensitive to outdoor temperature (e.g., construction or mining) (Hsiang et al., 2017; Zivin & Neidell, 2014). Weather may also affect saving through household consumption, by influencing local and regional food prices (von Braun, 2009). In economic analyses, weather has been used as an exogenous predictor of transitory income, particularly in developing countries where large proportions of population are engaged in agriculture. Previous research has dis-aggregate income into its two components, permanent and transitory, using weather as a predictor of transitory income (Hirvonen, 2016; Paxson, 1992). This method has been used to evaluate the effect of weather on consumption, saving and migration, among other topics.

In this paper, we evaluate weather as an exogenous predictor of transitory income, and examine household's propensity to save from permanent and transitory income with recent, nation-wide data from South Africa. We decompose income into permanent and transitory components where permanent income is proxied by household demographic structure and characteristics such as asset base, and transitory income is predicted with seasonal weather variations from 30-year climate normals (Paxson, 1992). We develop point estimates of propensity to save from both OLS and median regressions. We further evaluate the propensities to save by income distribution in order to identify any differential effect of income inequality on household propensity to save.

We are able to conduct such analysis with availability of the nationally representative panel data from the National Income Dynamic Study of South Africa (NIDS, 2008). The four waves we use cover the period from 2008 to 2014. This survey includes responses from over 6,000 households across 9 provinces and 52 districts of South Africa. The sampling is statistically weighted to represent the distribution of South Africa's population. We combine this data-set with daily weather data (e.g., temperature and precipitation) from the European Centre for Medium-Range Weather Forecasts' ERA-Interim data-set (Dee et al., 2011). We link the data with household geopositioning at the district level, which is the most dis-aggregated level that is publicly available.

Our data come from a country with high levels of income inequality. According to World Bank and Statistics South Africa, the country has a Gini coefficient of 0.63 in 2015, one of the highest in the world. The bottom 60% of the South African population hold 7% of total wealth, while the wealthiest 10% hold about 71% (WorldBank, n.d.-b). Thus, this presents an interesting opportunity to study the distributional attributes of propensity to save. We link granular daily weather deviations as an exogenous source of heterogeneity to household transitory income. With temperature and precipitation fluctuations in this time period that is related to cycles of El Nino Southern Oscillation (ENSO), we observe particular household heterogeneity as they relate to income that is affected by daily weather variations. We compare propensity to save from linear regression versus median regression, taking into account the non-normality of income distribution.

We find that the propensity to save by linear regression corresponds well with theoretical predictions and previously reported results, where marginal saving out of transitory income is close to one (when durable goods are considered saving). The marginal saving out of permanent income is not zero, indicating some portion of the permanent income is also saved. By median regression, we find a lower propensity to save than linear regression. We also find that the propensity to save from transitory income is significantly different from one. By income groups, our finding suggests that the treatment of durable goods leads to different results. When durable goods are considered as consumption, the propensities to save do not differ significantly across quintiles, leading to a consistent marginal savings rate across the income distribution. When durable goods are considered as saving, we find a higher marginal saving rate for the higher income quintile versus the lower income group.

For the remainder of this paper, we discuss a review of relevant literature in Section 2. We present a summary of the theoretical model of inter-temporal consumption and saving to guide our empirical analysis in Section 3. In section 4, we discuss our data. In section 5, we outline our empirical approach. Section 6 is the presentation of our results, and section 7 concludes.

## 2. Literature review

The literature on the propensity to consume and save is vast and well established. In this section, we selectively highlight some of the results, as they are relevant to the theoretical understanding of our analysis.

*Weather as a predictor of income*

The use of weather as an exogenous predictor of income is well documented. The most recognized pathway by which weather can affect income is through its direct impact on agricultural conditions, which affects yield. A recent systematic review of changes in crop yields under climate change across Africa and South Asia found that yields may decline by 8 percent across both regions by 2050 due to changes in weather patterns (Knox, Hess, & Daccache, 2012). In economic evaluations, Wolpin (1982) used long-term annual rainfall data in India to estimate the permanent income elasticity of consumption, arguing that long-term stable weather conditions is a source of heterogeneity in permanent income among traditional farming households. Paxson (1992) used time-series information on deviation from average of regional rainfall as an explicit measure of transitory income in her evaluation of the propensity to save, while she measured permanent income with household demographic characteristics and asset levels. Hirvonen (2016) estimated coefficients of temperature and rainfall on log-consumption in his evaluation of the effect of temperature change on consumption and migration. A less studied pathway by which weather can affect income and consumption is through its indirect effect on food prices in the local, regional, or even global markets (Wheeler & von Braun, 2013). It is postulated that small shocks on the supply or demand of food can have large impacts on prices (von Braun, 2009). This, however, has not been sufficiently studied, both in macroeconomic modelling and microeconomic studies based on panel datasets (Headey & Fan, 2008). Using weather as a predictor of transitory income has the limitation of ignoring unpaid family labor in an agricultural setting. Such omission may present a biased estimate of income effects on consumption in favor accepting the null hypothesis of perfect consumption smoothing (Rosenzweig, 2001).

In countries like South Africa, where weather-impacted labor activities such as construction and mining represent a significant portion of household livelihood, increased weather variability represent a particular source of vulnerability to stable income and consumption (Bryan, Deressa, Gbetiboubo, & Ringler, 2009; Hsiang et al., 2017; Karfakis, Lipper, & Smulders, 2012). Weather variability affects farmers engaged in rain-fed crop production, which is predominant in many African countries (Bozzola, Smale, & DiFalco, 2016; Hirvonen, 2016; Kurukulasuriya & Mendelsohn, 2008). In a low-income setting, Rosenzweig and Wolpin (1993) showed that without insurance and constrained by credit, those with low average income (i.e., small and medium size farms) cannot effectively smooth consumption based on asset transactions (i.e., bullocks). In rural Burkina Faso during a period of severe drought (1981 to 1985), for example, it was found that rural households exhibited little consumption smoothing and households relied almost exclusively on self-insurance in the form of grain stock adjustments (Kazianga & Udry, 2006). Studying the effect of weather on consumption and the consequent impact on labor migration in Tanzania, Hirvonen (2016) finds that a standard deviation increase in mean temperature of the previous growing season led to a 5% decrease in household consumption, and a 13% decrease in male migration. This reduction in migration is attributed to potential liquidity constraints due to temperature change.

*Consumption and saving responses to income*

Our theoretical framework is related to consumption and saving responses to income change as put forth in the standard permanent income hypothesis (Jappelli & Pistaferri, 2010) and buffer stock model (Deaton, 1989, 1991). If real interest rate is equal to the discount rate, and agents exhibit quadratic preferences and are risk neutral, the optimal consumption path is one of consumption smoothing, where consumption in each period is equivalent to lifetime average of future income. If income can be decomposed into a permanent, deterministic component, and an independent and identically distributed (i.i.d) transitory component, then the propensity to consume out of permanent income should be close to one, while the propensity to consume out of transitory income should be close to zero (Jappelli & Pistaferri, 2010). The buffer-stock model introduces impatience, prudence and credit constraint, by assuming that the discount rate is greater than the real interest rate, and agents exhibit utility functions such that the marginal utility is (strictly) convex. When future income follows an i.i.d process, or an autoregressive process, consumption behaves as a non-linear function of assets and income, where the propensity to consumer above the break point is less than one ((Deaton, 1989, 1991).

Early empirical studies of propensity to consume mostly found a coefficient on permanent income of less than one (Bhalla, 1979, 1980; Musgrove, 1979, 1980). Wolpin (1982) finds the permanent income elasticity of consumption to be close to one, measuring household heterogeneity in permanent income with long-term average rainfall in Indian farming households. Paxson (1992) studies the saving behavior of Thai rice farmers using weather variability as a measure of transitory income and finds that a significant portion of transitory income is saved, although some permanent income is also saved. Empirical evidence on the importance of precautionary saving provides mixed results. While some studies have found a limited role of precautionary saving in wealth accumulation, others have found a larger role. Jappelli, Padula and Pistaferri (2008) conducted a test of precautionary motive in Italian households, leveraging a direct question in the survey questionnaire that asked how much savings they think they need for future emergencies. They find that the ratio of covariance between wealth gap and consumption, and that of wealth gap and income is much lower than what would be predicted from the buffer-stock model, both across the whole population, and in groups that would seem to exhibit buffer-stock like behavior. Their finding suggests a more steady consumption path relative to future expectations of income than what would be predicted from the buffer-stock model.

*Heterogeneous responses by income distribution*

A related question is whether differences exist in saving behavior across income quantiles. Several theory extensions seek to explain the complex relationship between income and saving across income distribution, specifically the positive correlation between income and saving at the national level. An early theory (Kaldor, 1957; Lewis, 1954) suggests that individual propensity to save is less than firm's, such that higher aggregate saving reflects firm's higher propensity to save, not individuals. Empirical studies based on cross-sectional micro-data in general have found that saving rates increases with current income while aggregate saving and income studies offer mixed results. Della Valle and Oguchi (1976) finds no statistically significant effect of income distribution on saving, except for OECD countries. Lim (1980) finds that inequality tends to raise aggregate saving rate in a cross-sectional sample of developing countries. Using data from 49 countries, Venieris and Gupta (1986) find that poorer household have the lowest saving propensities while the highest ones corresponds to the middle-income group. Cook (1995) estimates the impact of various inequality measures on the GDS/GDP ratio in 49 least developed countries (LDCs), and finds significant positive effect of inequality on saving. Using annual cross-sectional time-series data for the 1965-1994 period from World Bank, Schmidt-Hebbel and Serven find no significant relationship between income inequality and aggregate saving (Schmidt-Hebbel & Serven, 2000). Bunting (1991) finds evidence that households' marginal propensity to save uniformly increases as their quintile share of income rises. Using data from the US, Dynan (2004) finds strong positive relationship between income quintile and average saving rate, and a weaker, albeit positive relationship with marginal saving rate. Some studies suggest a bequeath motive among the rich which explains the higher saving rate among higher income quantiles (Carroll, 1998). Kopczuk (2007) finds that roughly 75% of elderly households in the US have a bequest motive, and they spend 25% less on personal outlays on average. A further explanation is precautionary saving, where individuals in lower income quartiles depress current consumption, leading to a higher marginal propensity to consume (Carroll & Kimball, 1996).

Rosenzweig et al. (2001) outlines three characteristics where the study of saving in developing countries differ from high income countries. First, due to the lower income level in such countries relative to the minimal standards for living, opportunities for sustained saving is limited. Second, agriculture tends to be a larger share of economic activity, and it is a less predictable income source that is subject to fluctuation in conditions such as weather. Third, there is usually a lack of a well-functioning insurance and credit market. Thus, households in developing economies are faced with the challenges of maintaining consumption while income is low and highly variable.

*Saving behavior in South Africa*

Several studies have recently examined saving behavior in South Africa. Zwane et al. (2016) studied the effect of income on saving in South Africa using the first three waves of NIDS. They included province level fixed effects, and also an instrument variable using lagged income. They find a strong positive effect of income on saving. Bengtsson (2012) estimated the propensity to consume in South Africa out of unearned income, leveraging the introduction of an unconditional child support grant. The marginal propensity to consume was estimated to be 0.7 in this case while the marginal propensity to save seems to be close to zero (or even negative against future grant payments). Ting and Kollamparambil (2015) looked at saving behavior in South African households using the General Household Survey (2002-2004 and 2008-2010). Their findings suggest that government grants not only sustain household consumption, but also household saving, as the consumption-income ratio seem to be smooth over age cohorts.

## 3. Theoretical framework

In this section, we present a summary of the theoretical framework of rational consumption response to income changes for a representative household (Deaton, 1989; Jappelli & Pistaferri, 2010). We assume a unitary household model, and do not consider intra-household dynamics. In this sample of South African households, only about 30% of the individuals have primary income through employment. This implies that household income typically represents income from a minority of household members, while consumption is shared among the household. We do not have sufficiently granular data, however, to identify separate consumption and saving by household member. Therefore, we consider the household as a unit of analysis, with household income and consumption as the main variables of analysis.

Consider the standard model of a household agent who maximizes the expected utility of consumption over some time period, subject to an inter-temporal budget constraint and a terminal condition on wealth. In each period $t$, the household agent $i$ receives income $y_{it} = \overline{y_{it}} + \epsilon_{it}$, where $\overline{y_{it}}$ is deterministic and $\epsilon_{it}$ represent shocks with $E(\epsilon_{it}) = 0$. The agent chooses consumption $c_{it}$ to maximize remaining lifetime expected utility $E_{it}(\sum_t^\infty (1+\delta)^{-t} u(c_{it}))$, subject to the budget constraint $w_{it} = (1+r_t)(w_{i,t-1} + y_{it} - c_{it})$ where $w_t$ is wealth in period $t$, and $y_t$ is income in period $t$. $r$ is the fixed real interest rate, $\delta > 0$ is the rate of time preference, $c_{it}$ is household consumption, and $u(c_{it})$ is the instantaneous utility associated with consumption $c_{it}$. If the utility function is time-separable, then the Euler equation is:

$$u'(c_{it}) = E_t[\frac{(1+r_t)u'(c_{i,t+1})}{(1+\delta)}] \quad (1)$$

As in the standard model, a quadratic utility function, which assumes a constant marginal utility of consumption, gives the permanent income model with certainty equivalence (Campbell, 1987; Flavin, 1981; Jappelli & Pistaferri, 2010). If the discount rate and interest rate are constants and equal to each other, the Euler equation is:

$$c_{i,t+1} = c_{i,t} + \epsilon_{i,t+1} \tag{2}$$

where $\epsilon_{i,t+1} = c_{i,t+1} - E_t c_{i,t+1}$ is the consumption modifier that adjusts based on new information about uncertainties faced by the agent. Assuming labor income uncertainty in the future periods, changes in consumption from period $t+1$ to $t$ becomes:

$$\Delta c_{it} = \frac{r}{1+r}\sum_{\tau=0}^{\infty}(1+r)^{-\tau}(E_t - E_{t-1})y_{i,t+\tau} \tag{3}$$

where $r/(1+r)$ is the annuity factor under the assumption of infinite horizon. If income can be decomposed into a permanent component, $P_{it} = P_{i,t-1} + u_{it}$ and an independent and identically distributed (i.i.d) transitory component $v_{it}$, then

$$\Delta c_{it} = \frac{r}{1+r}v_{it} + u_{it} \tag{4}$$

In this case, the model predicts that consumption responds one-to-one to permanent income shocks but is nearly insensitive to transitory shocks. Furthermore, as shown by Campbell (1987), the saving equation can be written as

$$s_{it} = \frac{1}{1+r}v_{it} \tag{5}$$

This identity implies that saving should respond to changes in transitory income but not permanent income (Jappelli & Pistaferri, 2010).

To incorporate precautionary saving, assume that the utility function in Equation 1 is isoelastic, so that the agent exhibits constant relative risk aversion. The marginal utility of the isoelastic utility function takes the form of $c^{-\rho}$. The parameter $\rho$ controls the degree of precautionary saving. In this case, the instantaneous utility function $u$ is three times differentiable and satisfies $u' > 0$,$u'' < 0$, and $u''' > 0$. The third derivative condition implies that the agent is a precautionary saver as the marginal utility is (strictly) convex (Berg, 2013; Deaton, 1989). The marginal disutility of consumption losses at near subsistence is greater than gains in marginal utility during times of abundance. Agents will therefore give up some consumption in good times to save for those bad times, even if they are less frequent. Furthermore, assume that the agents are credit constrained, in that she is not able to borrow sufficiently in times of need, such that

$$w_{it} \geq 0 \tag{6}$$

With the assumption that $\delta > r$ such that agents are impatient, the credit constraint will have an effect. Given the presence of credit constraint where the marginal utility of assets and income at time $t$ is greater than the expected marginal utility of consumption at $t+1$, the Euler equation takes on the following form:

$$u'(c_{it}) = \max \{u'(w_{it} + y_{it}), E_t[\frac{(1+r_t)u'(c_{i,t+1})}{(1+\delta)}]\} \tag{7}$$

If income can be assumed to be independent and identically distributed, then consumption graphed against cash-on-hand (assets and permanent income) shows a discontinuity. At lower levels of cash-on-hand, consumption traces cash-on-hand at a steeper slope, resulting in no saving or dis-saving. The shape of the consumption curve as it relates to cash-on-hand is determined by the interaction of the credit constraint and the degree of precautionary saving. A stronger precautionary motive shifts the consumption curve lower (Deaton, 1989). Furthermore, saving will increase with relation to permanent income, leading to a less than one propensity to consume in relation to permanent income (Deaton, 1989). In this model, consumption is a non-linear function of cash-on-hand (Deaton, 1991).

Paxson (1992) specifies a testable form of the saving equation that is linear in permanent income, transitory income, and the variance of income. It is written as:

$$S_{irt} = \alpha_0 + \alpha_1 Y^P_{irt} + \alpha_2 Y^T_{irt} + \alpha_3 \text{VAR}_{ir} + \alpha_4 W_{irt} + \epsilon_{irt} \tag{8}$$

where $S_{irt}$ is saving for individual $i$, in region $r$, at time $t$. $Y^P_{irt}$ represents the permanent portion of the individual's income, while $Y^T_{irt}$ represents the transitory portion. $\text{VAR}_{ir}$ represents income variation of individual $i$ in region $r$, and $W_{irt}$ is a set of household lifecycle characteristics.

The standard model predicts that $\alpha_1$ should be close to zero, while $\alpha_2$ should be close to one. The prediction of $\alpha_3$ should be close to zero with a quadratic utility function as income variances do not factor into the saving equation with this utility function. If $\alpha_3$ is greater than zero, then that is an indication of risk-aversion.

## 4. Data

We leverage two separate data-sets for this analysis. The weather data is comprehensive climate data from 1979 to present day from European Centre for Medium-Range Weather Forecasts' ERA-Interim dataset (Dee et al., 2011). The household data is the National Income Dynamic Study of South Africa (NIDS, 2008).

### *4.1 Weather*

We compile mean daily temperature and total daily precipitation data from ERA-Interim data (Dee et al., 2011). This data provides a reanalysis of global atmosphere since 1979, on a 0.75-by-0.75 degrees resolution grid (about 80km by 60km). The ERA-Interim reanalysis is constructed from actual observations to provide a spatially complete and coherent record of global atmospheric circulation. This type of data is increasingly used by economists (Burgess, Deschenes, Donaldson, & Greenstone, 2014; Colmer, 2016a, 2016b) as it can be a valuable alternative when studying geographies where weather observation stations are scarce and scattered with inconsistent reporting (Berrisford et al., 2011; Dee et al., 2011).

We construct two sets of rainfall variables and three sets of temperature variables, matching weather data to each of South Africa's 52 districts using GPS coordinates. See Table 2 and Appendix 8.1 for descriptive summary of weather. We construct weather variables from GPS-matched climate data, according to planting, growing and non-agriculture seasons in South Africa (VAM, 2016). For rainfall, we calculate seasonal deviation from long-term normals as the difference between total seasonal rainfall and climatic normals. Normals are calculated as the average of total seasonal rainfall of the previous 30 years (from 1979 to 2008). We include a square term to account for non-linearity. We also construct the coefficient of variation by season to account for variation in the standard deviation, where coefficient of variation is the standard deviation of seasonal rainfall, divided by the mean. Similarly for temperature, we calculate the difference between mean seasonal temperature and temperature normals, and coefficient of deviation by season. We also include extreme events, captured by a variable that accounts for the number of days in the growing season when the temperature is above 34 degrees Celsius. Temperature above this level has been shown to be detrimental to crop growth (Hirvonen, 2016).

### *4.2 Household Panel*

The National Income Dynamic Study (NIDS) of South Africa collects representative household consumption and income information in two-year increments, starting in 2008 (NIDS, 2008). The data is rich in individual and household characteristics, spending pattern, and health and education. This data has been used to study gender effects and food adequacy of subsistence farming in South Africa (Tibesigwa & Visser, 2016; Tibesigwa, Visser, & Turpie, 2015), among others. The website of NIDS keeps an ongoing list of working papers and publications using this data.

The NIDS follows a stratified, two-stage clustered sample design. Leibbrandt (Leibbrandt, Woolard, & de Villiers, 2009) provides detailed description of the sampling and data collection methodology. We present a summary in Appendix 8.2. Survey instruments are developed by South African Labor and Development Research Unit at School of Economics at the University of Cape Town. Quality control, such as back-to-the field rework, is built into the data collection process. The attrition rate (relative to the previous wave) for Wave 2 is 22%, Wave 3 is 16% and Wave 4 is 14% (Chinhema et al., 2016).

Definition of key variables and descriptive statistics are provided in Table 1 and Table 2. Mean income for South African households over the study period is 3,708 South African Rand (SFR), with a standard deviation of 9,100 SFR (see Table 2). Median income is 1,834 SFR. Values for 25th percentile and 75th percentile are 975 SFR and 3,568 SFR, respectively.

Table 1: Definition of variables

| Variable | Definition |
|---|---|
| *Outcome variables* | |
| Income | Total HH income from all sources in last 30 days |
| Consumption | Total consumption, including durable and non-durable, in last 30 days |
| Durable consumption | Expenditure in HH maintenance, kitchen, furniture, clothing, etc. in last 30 days |
| Non-durable consumption | Expenditure in food, utilities, and other personal consumption in last 30 days |

| | |
|---|---|
| Saving | Version 1: Income minus total consumption; Version 2: Income minus non-durable consumption |
| *Permanent income factors* | |
| Assets (in deciles) | Market value of owned house (rent=0), inheritances, and household reported net assets in deciles |
| Under 5yo | No of HH members under 5 years old |
| 6-11yo, male / female | No of HH male/female members from 6 to 11 years old |
| 12-17yo, male / female | No of HH male/female members from 12 to 17 years old |
| 18-64yo, male / female, $<$ 7 yrs of education | No of HH male/female members from 18 to 64 years old, with less than 7 years of education |
| 18-64yo, male / female, 7-9 yrs of education | No of HH male/female members from 18 to 64 years old, with 7 to 9 years of education |
| 18-64yo, male / female, $>$9 yrs of education | No of HH male/female members from 18 to 64 years old, with more than 9 years of education |
| $>$65yo, male / female | No of HH male/female members $>$65 years old |
| *Transitory income factors* | |
| Rainfall, deviation from mean (mm) | Total rainfall deviation from climate normal by season |
| Rainfall, coefficient of variation ($\sigma/\mu$) | Rainfall std. dev. divided by mean in the same season |
| Temperature, deviation from mean (mm) | Seasonal mean deviation from climate normal |
| Temperature, coefficient of variation | Temperature std. dev. divided by mean in the same season |
| Days in growing season over 34deg C | No days in the growing season warmer than 34deg C |

In the survey, each household is asked to estimate the income generated by members of the household in the previous 30 days. This includes all of household member's salaries and wages, grants, interest, rental income and income from agriculture. Household either report the actual income figure, or indicate the range in which the household income falls. When income range is indicated, we take the median value for each range, and cap the income at the highest range.

Average "missingness" across all four waves of the survey is 20.5%, and is within the range of typical household surveys (Kim, Egerter, Cubbin, Takahashi, & Braveman, 2007). Where household income is missing, we proxy by adding the individual income from members within the household.

We inflation-adjust income and consumption to take into account the fact that South Africa experienced significant inflation from the period of 2008 to 2014. By province, overall CPI ranges from 111.9 to 113.5 in 2010, 123.1 to 125.5 in 2012, and 131.7 to 134.3 in 2014, with 2008 as the baseline (StatsSA, 2014). Data from three households were identified as outliers and removed from the dataset.

We construct two saving measures. The first measure (version 1) is simply income minus consumption. This measure is likely to underestimate saving since durable goods are considered as consumption. In the second measure (version 2), we exclude durable goods, which comprises of items such as household maintenance, furniture, and clothing, in consumption. Hence the saving measure is income, minus expenditure, plus durable goods.

Table 2: Descriptive statistics

| Variable | Obs. | Mean | S.D. | Min | Max |
|---|---|---|---|---|---|
| *Outcome variables* | | | | | |
| Income | 29,750 | 3,756 | 9,155 | 1 | 440,529 |
| Consumption | 31,467 | 2,716 | 7,562 | 3 | 589,749 |
| Durable consumption | 8,949 | 1,580 | 5,963 | 1 | 308,884 |
| Non-durable consumption | 31,062 | 1,994 | 3,820 | 3 | 377,114 |
| Saving (version 1) | 29,622 | 1,008 | 9,560 | -549,749 | 426,167 |
| Saving (version 2) | 29,622 | 1,466 | 9,244 | -549,749 | 429,515 |
| *Permanent income factors* | | | | | |
| Assets (in deciles) | 16,355 | 5.5 | 2.9 | 1 | 10 |
| Under 5yo | 31,595 | 0.5 | 0.8 | 0 | 13 |
| 6-11yo, male | 31,595 | 0.3 | 0.6 | 0 | 5 |
| 6-11yo, female | 31,595 | 0.3 | 0.4 | 0 | 6 |
| 12-17yo, male | 31,595 | 0.2 | 0.5 | 0 | 5 |
| 12-17yo, female | 31,595 | 0.3 | 0.5 | 0 | 5 |
| 18-64yo, male, $<$ 7 yrs edu. | 31,595 | 0.1 | 0.4 | 0 | 6 |
| 18-64yo, female, $<$ 7 yrs edu. | 31,595 | 0.2 | 0.5 | 0 | 5 |
| 18-64yo, male, 7-9 yrs edu. | 31,595 | 0.4 | 0.6 | 0 | 7 |

| | | | | | |
|---|---|---|---|---|---|
| 18-64yo, female, 7-9 yrs edu. | 31,595 | 0.2 | 0.5 | 0 | 7 |
| 18-64yo, male, >9 yrs edu. | 31,595 | 0.2 | 0.5 | 0 | 6 |
| 18-64yo, female, >9 yrs edu. | 31,595 | 0.6 | 0.8 | 0 | 7 |
| >65yo, male | 31,627 | 0.1 | 0.3 | 0 | 3 |
| >65yo, female | 31,627 | 0.1 | 0.4 | 0 | 3 |
| *Transitory income factors* | | | | | |
| Rainfall, deviation from mean (mm) | | | | | |
| Planting season | 31,214 | 11.5 | 60.8 | -98.8 | 219.3 |
| Growing season | 31,214 | -2.2 | 69.3 | -155.5 | 198.2 |
| Rest of year | 31,214 | -32.2 | 60.8 | -167.9 | 156.3 |
| Rainfall, coefficient of variation ($\sigma/\mu$) | | | | | |
| Planting season | 31,214 | 2.1 | 0.7 | 1.2 | 6.2 |
| Growing season | 31,214 | 2.0 | 0.7 | 1.2 | 5.3 |
| Rest of year | 31,214 | 3.7 | 1.2 | 1.7 | 8.5 |
| Temperature, deviation from mean (deg) | | | | | |
| Planting season | 31,214 | -0.1 | 1.0 | -2.8 | 3.3 |
| Growing season | 31,214 | 0.2 | 1.2 | -3.5 | 2.9 |
| Rest of year | 31,214 | 0.5 | 0.7 | -1.9 | 2.5 |
| Temperature, coefficient of variation ($\sigma/\mu$) | | | | | |
| Planting season | 31,214 | 0.1 | 0.03 | 0.08 | 0.2 |
| Growing season | 31,214 | 0.1 | 0.03 | 0.04 | 0.2 |
| Rest of year | 31,214 | 0.2 | 0.06 | 0.08 | 0.4 |
| Days in growing season over 34 deg $C$ | 31,214 | 36 | 27 | 0 | 88 |

Note: For saving, negative values refer to dissaving in last 30 days, according to survey output. We retain these values in order to have minimum interference with raw data. For weather variables, in South Africa, planting season refers to October to December of previous year, growing season refers to January to March, and rest of the year refers to April to September.

We also assess the distributional effect on propensity to save. The gross domestic savings (as % of GDP) in South Africa in this period ranges from 21.5 in 2008 to 19.1 in 2014 (WorldBank, n.d.-a). In our sample, the average saving rate is 27% when durable goods are considered consumption which is higher than the aggregate number reported by the World Bank. We provide the summary statistics on the outcome variables by income quintile is presented in Table 3.

Table 3: Summary of income and saving by quintile

| Income quintiles | Income (SFR) | Saving (v.1) (SFR) | Saving (v.1) (% of inc.) | Saving (v.2) (SFR) | Saving (v.2) (% of inc.) |
|---|---|---|---|---|---|
| Lowest 20% | 581 | -387 | -66% | -255 | -44% |
| 20% to 40% | 1,153 | 35 | 3% | 138 | 12% |
| 40% to 60% | 1,860 | 283 | 15% | 489 | 26% |
| 60% to 80% | 3,124 | 837 | 27% | 1,149 | 37% |
| Highest 20% | 12,096 | 4,281 | 35% | 5,819 | 48% |

## 5. Empirical approach

Starting with Equation 8, we make a slight modification by including time heterogeneity in the variance term. The empirical equation is

$$S_{irt} = \alpha_0 + \alpha_1 Y^P_{irt} + \alpha_2 Y^T_{irt} + \alpha_3 \mathrm{VAR}_{irt} + \alpha_4 W_{irt} + \epsilon_{irt} \tag{9}$$

As presented in the theoretical framework, the standard model would predict that $\alpha_1$ should be close to zero, while $\alpha_2$ should be close to one in the case of certainty equivalence. In the case of a buffer stock model, $\alpha_1$ should be greater than zero. The estimated $\alpha_3$ is a measure of risk aversion and should be greater than zero without certainty equivalence.

Following Paxson (1992), we estimate permanent income as

$$Y^P_{irt} = \beta^P_{ir} + \beta_1 X^P_{irt} + u^P_{irt} \tag{10}$$

where $X^P_{irt}$ includes a set of variables that capture the household's assets and demographic characteristics. We use market value of the dwelling (owned houses) as an indication of the asset level of the household. If the household rents, then the asset level is zero. Otherwise, asset level is categorized into deciles in order to reduce potential measurement error in the explanatory variables. Furthermore, we construct household demographic structure by categorizing household members by age, gender and education levels. The demographic structure of the households is divided into 13 categories, as described in Appendix 8.3. $\beta^P_{ir}$ captures household fixed effects, and $u^P_{irt}$ is the stochastic error term.

We estimate transitory income using temperature and precipitation. These weather variables are likely correlated, and including one without the other may lead to omitted variable problems

(Auffhammer, Hsiang, Schlenker, & Sobel, 2013). We estimate the following linear expression for transitory income.

$$Y_{irt}^T = \beta_t^T + \beta_2 X_{irt}^T + u_{irt}^T \tag{11}$$

where $X_{irt}^T$ represent a set of parameters that characterize temperature, rainfall and extreme degree-days. $\beta_t^T$ is year fixed effects to capture the year-to-year variation in transitory income not captured by weather, and $u_{irt}^T$ is the error term.

Given the expressions of permanent and transitory income, total income can be expressed as the sum of the two:

$$Y_{irt} = \beta_t^T + \beta_{ir}^P + \beta_1 X_{irt}^P + \beta_2 X_{irt}^T + u_{irt} \tag{12}$$

The saving equation can therefore be expressed as:

$$S_{irt} = \alpha_0 + \alpha_1(\beta_{ir}^P + \beta_1 X_{irt}^P + u_{irt}^P) + \alpha_2(\beta_t^T + \beta_2 X_{irt}^T + u_{irt}^T) + \alpha_3 \text{VAR}_{irt} + \alpha_4 W_{irt} + \epsilon_{irt} \tag{13}$$

After simplification, the saving equation becomes:

$$S_{irt} = \gamma_t + \gamma_{ir} + \gamma_1 X_{irt}^P + \gamma_2 X_{irt}^T + \gamma_3 \text{VAR}_{irt} + v_{irt} \tag{14}$$

where $\gamma_t$ is time fixed effects, $\alpha_2\beta_t^T$, $\gamma_{ir}$ is the constant $\alpha_0$ and household fixed effects $\alpha_1\beta_{ir}^P$ , $\gamma_1$ is $\alpha_1\beta_1$, $\gamma_2$ is $\alpha_2\beta_2$, and $\gamma_3$ is $\alpha_3$. $W_{irt}$ does not appear in the reduced form equation as it is collinear with determinants of permanent income.

We first undertake the regression of saving equation as specified in Equation 14. We test the joint significance of permanent and transitory factors on saving (i.e., $\gamma_1$ and $\gamma_2$.) We should see that $\gamma_1$ is not significantly different from zero, while $\gamma_2$ is significantly different by the permanent income hypothesis. The buffer stock model would predict that $\gamma_1$ is different from zero. The coefficient $\gamma_3$ represents the significance of risk measure in the saving equation.

Next, we undertake a procedure to estimate the propensities to save out of permanent and transitory income (i.e., $\alpha_1$, $\alpha_2$ from Equations 9 and 13). In the first step, we regress permanent and transitory factors on income as specified in Equation 13. We use the estimated coefficients to construct permanent income ($\hat{Y^P_{irt}}$), and transitory income ($\hat{Y^T_{irt}}$). In the second step, we estimate the saving equation using the fitted values.

$$S_{irt} = \alpha_1 \hat{Y^P_{irt}} + \alpha_2 \hat{Y^T_{irt}} + \alpha_3 \text{VAR}_{irt} + \alpha_4 W_{irt} + \alpha_5 \hat{u_{irt}} + w_{ir} + v_t + \epsilon_{irt} \tag{15}$$

We include the residual from the income regression, $\hat{u_{irt}}$, in the estimation of the saving equation, as income residual is often interpreted as transitory component of the income (Paxson, 1992). For household life-cycle factors ($W_{irt}$), we use the categories of demographic factors presented in Table 1, Table 2 and in Appendix 8.3.

We test the joint significance of coefficient for income variance ($\alpha_3$) in the saving equation (Equation 15). Income variance is represented by coefficient of variation of the weather variables. If we assume that the underlying utility function is isoelastic, we expect that this coefficient is significantly different from zero.

We compare the result of the OLS regression in the second stage with results of a median regression. Linear regression estimate the relationship between a set of regressors and the outcome variable based on the conditional mean $E(y|x)$. Median regression, or least absolute deviation (LAD) regression, consider the relationship using the conditional median function $Q_{50}(y|x)$. If $\epsilon_i$ is the model prediction error, OLS minimizes $\sum_i e_i^2$ while median regression minimizes $\sum_i |e_i|$. It minimizes a sum that gives asymmetric penalties $0.5|e_i|$ for overprediction and $0.5|e_i|$ for underprediction. While OLS can be inefficient if the errors are highly non-normal, median regressions are more robust to non-normal errors and outliers (Wooldridge, 2012).

To assess the distributional effects on the propensity to save, we divide income into quintiles, and assess the result of median regressions by each quintile. We construct dummy variables for each quintile, and allow the estimate of the median regression to vary with each quintile, while the constant remains the same. This is similar to the methodology used by Dynan (2004) using data from US Survey of Consumer Finances in their investigation of the relationship between saving rate and current income. We report the result from the full set of income and saving data, using both saving (v.1) and saving (v.2), which investigates the role of durable goods as a saving mechanism.

In the next section, we report the main findings of our empirical study.

## 6. Results

### *6.1 Propensity to save with OLS and median regressions*

Table 4 presents the results of the joint significance tests from Equations 12 and 14. Standard errors are clustered at districts, which is aggregation level of weather data. We present results from three sets of tests. The first one is the regression of permanent and transitory income factors on income, following Equation 12. The next two are the regression of permanent and transitory income factors on saving (version 1 and version 2), following Equation 14.

As predicted from previous studies (Paxson, 1992), we find significant effect of asset, demographic, and weather variables on income, suggesting that the underlying factors that drive permanent income (e.g., assets and demographics) and transitory income (e.g., weather) are predictors of household income. From the standard model, we would expect to find the coefficients of permanent income on saving to be non-significant, and transitory income on saving to be significant. In this sample, we find that the coefficients on transitory income is significant. We find that the coefficients of permanent income on saving is significant when durable consumption is excluded as saving, but not significant when durable consumption is included. In addition, the coefficients on income variance is not significant when durable goods are not considered as saving, and weakly significant when durable goods are considered as saving.

Table 4: Test of joint significance on income and saving

| F-statistic | Income | Saving (v.1) | Saving (v.2) |
|---|---|---|---|
| Permanent income ($X^P_{irt}$) | 3.35*** | 2.51** | 1.53 |
| Transitory income ($X^T_{irt}$) | 2.28* | 4.52*** | 4.62*** |
| Income variance ($VAR_{irt}$) | - | 1.68 | 1.99* |
| *Controls* | | | |
| Household fixed effects | Yes | Yes | Yes |
| Year fixed effects | Yes | Yes | Yes |
| Observations | 15,588 | 15,559 | 15,559 |

*** $p < 0.001$ ** $p < 0.01$ * $p < 0.05$. Standard errors clustered at districts.

Using permanent and transitory income factors, we construct estimated permanent and transitory income, $\hat{Y}^P_{irt}$ and $\hat{Y}^T_{irt}$. Table 5 outlines the second stage result of the regression from Equation 15. We first use a pooled OLS model for the second stage regression. We find that $\hat{\sigma}(u_i) = 0$, which indicates that the standard deviation of the individual effects is zero. In this case, the GLS is equivalent to OLS and there is no heterogeneity in individual fixed effect.

For every unit increase in estimated transitory income, saving (v.1) increases by 0.77 units. For every unit increase in estimated permanent income, saving (v.1) increased by 0.71 units. There is no significant differences between the coefficients of permanent income and transitory income. However, while we can strongly reject that hypothesis that the coefficient on permanent income is equal to one, we can weakly reject this hypothesis for transitory income. When durable goods are considered as saving, we cannot reject the hypothesis that the marginal saving out of transitory income is equal to one. While the propensity to save from permanent income is not zero, it is also not close to one. Some portion of permanent income is also saved. This result is in alignment with result from Paxson (1992) (although the marginal saving on permanent income is higher than that found in Paxson's study).

We also present the results of median regression in Table 5. The difference between the marginal saving from OLS versus median regression reflect the difference between the estimator of the expected conditional mean and the expected conditional median. Since the residual from OLS is not normally distributed, median regression using least absolute deviation is more robust to outliers.

Estimated marginal saving using median regression is smaller in magnitude than OLS. We reject both null hypotheses that the marginal saving from permanent and transitory saving is equal to one.

Table 5: Estimated propensity to save

| | OLS | | Median | |
|---|---|---|---|---|
| Coefficients | Saving (v.1) | Saving (v.2) | Saving (v.1) | Saving (v.2) |
| Permanent income ($\hat{Y}^P_{irt}$) | 0.71*** | 0.77*** | 0.52*** | 0.64*** |
| Transitory income ($\hat{Y}^T_{irt}$) | 0.77*** | 0.90*** | 0.55*** | 0.59*** |
| Residual ($\hat{u}^T_{irt}$) | 0.93*** | 0.99*** | 0.67*** | 0.81*** |
| *Controls* | | | | |
| Household fixed effects | No | No | No | No |
| Year fixed effects | Yes | Yes | Yes | Yes |
| Household demographic | Yes | Yes | Yes | Yes |
| Income variation | Yes | Yes | Yes | Yes |
| *Test statistics* | | | | |
| $\hat{Y}^P_{irt} = \hat{Y}^T_{irt}$ | 0.24 | 1.56 | 4.21* | 0.50 |
| $\hat{Y}^P_{irt} = 1$ | 33.99*** | 32.49*** | 889*** | 271*** |
| $\hat{Y}^T_{irt} = 1$ | 5.37* | 1.31 | 377*** | 25*** |
| $VAR_{irt} = 0$ | 6.36*** | 7.26*** | 5.95*** | 3.07*** |
| $R^2$ overall | 0.44 | 0.56 | 0.43 | 0.55 |
| No. of observations | 15,559 | 15,559 | 15,559 | 15,559 |
| *Distribution of residuals* | | | | |
| Mean | 4.83E-7 | 3.82E-7 | - | - |
| Std. Dev. | 9,142 | 7,811 | - | - |
| Shapiro-Francia | $p < 0.001$ | $p < 0.001$ | - | - |
| Kolmogorov-Smirnov | $p < 0.001$ | $p < 0.001$ | - | - |

*** $p < 0.001$ ** $p < 0.01$ * $p < 0.05$. Standard errors clustered at districts. Shapiro-Francia and Kolmogorov-Smirnov are normality tests which assess whether the residual from the OLS regression follow a normal distribution.

*6.2 Propensity to save by income quintiles*

We present the propensity to save by income quintiles in two subsections. In the first subsection, we include the full data set and evaluate the result using saving v.1 where durable goods is considered as part of consumption. In the second subsection, we evaluate the results using saving v.2, where durable goods is considered as part of saving.

#### 6.2.1. Durable goods as consumption

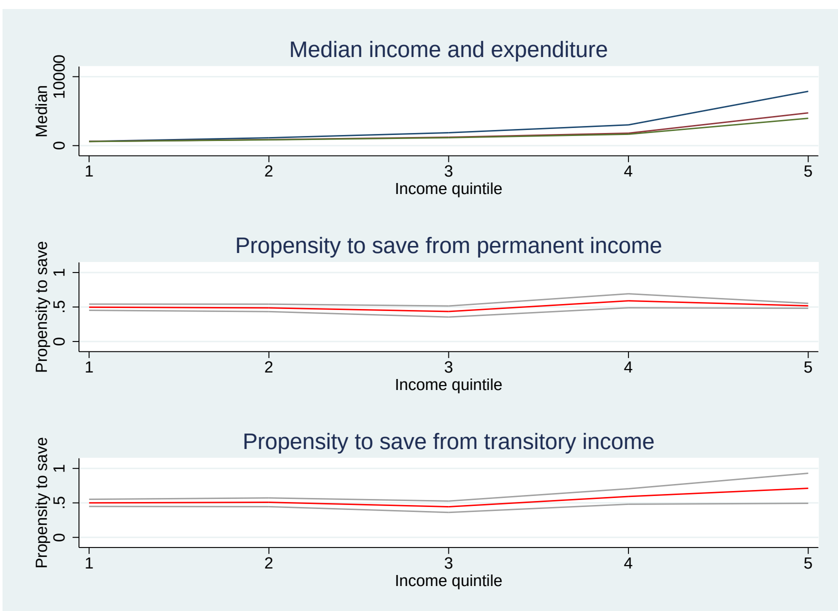
Table 6 presents the result of median regressions of saving (v.1) by income quintiles, and Figure 1 presents the results graphically. When durable goods are considered as consumption, saving is total income subtracting total expenditure including durable goods. With this outcome specification, we find no difference in marginal saving rate across income quintile. While there is a slight trend towards higher marginal saving for the highest 20% of the income distribution, we cannot reject the null hypothesis that it is the same as the lowest 20%.

Table 6: Estimated propensity to save by income quintiles

| | Income quintiles | | | | |
|---|---|---|---|---|---|
| Coefficients | 20% | 40% | 60% | 80% | 100% |
| Permanent income ($\hat{Y}^P_{irt}$) | 0.50*** | 0.49*** | 0.44*** | 0.59*** | 0.52*** |
| Transitory income ($\hat{Y}^T_{irt}$) | 0.50*** | 0.51*** | 0.44*** | 0.59*** | 0.71*** |
| Residual ($\hat{u}^T_{irt}$) | 0.54*** | 0.51*** | 0.52*** | 0.75*** | 0.79*** |
| *Controls* | | | | | |
| Household fixed effects | No | No | No | No | No |
| Year fixed effects | Yes | Yes | Yes | Yes | Yes |
| Household demographics | Yes | Yes | Yes | Yes | Yes |
| Income variation | Yes | Yes | Yes | Yes | Yes |
| *Test statistics* | | | | | |
| $\hat{Y}^P_{irt} = \hat{Y}^T_{irt}$ | 0.04 | 2.28 | 0.26 | 0.00 | 2.92 |
| $\hat{Y}^P_{irt} = 1$ | 477*** | 346*** | 190*** | 62*** | 695*** |
| $\hat{Y}^T_{irt} = 1$ | 358*** | 230*** | 175*** | 51*** | 7** |
| $\hat{Y}^P_{irt}$ of bottom 20% = top 20% | | | 0.38 | | |
| $\hat{Y}^T_{irt}$ of bottom 20% = top 20% | | | 2.95 | | |

*** $p < 0.001$ ** $p < 0.01$ * $p < 0.05$. Standard error clustered at districts.

Figure 1: Propensity to save by quantiles, full data set




Note: Top graph panel represent median income and expenditure (total and non-durable expenditure). Red lines in graph panels below represent median regression estimates. Gray lines represent 95% confidence intervals.

#### 6.2.2. Durable goods as saving

Table 7 presents the result of median regressions of saving (v.2) by income quintiles, and Figure 2 presents the results graphically. In this case, the outcome variable, saving, include durable consumption of the period. Consumption in this period include non-durable consumption and housing. If we consider the purchase of a durable good as a saving mechanism, then we would consider this measure of saving as a more accurate reflection of saving. In this case, we reject the null hypothesis that the marginal saving of the poorest 20% is the same as the richest 20% for both permanent and transitory income. There is an increasing trend in marginal saving for both permanent income and transitory income. In addition, for the top 20% of income distribution, the marginal saving from transitory income is significantly different from the marginal saving from permanent income, although both are different from one.

Table 7: Estimated propensity to save by income quintiles

| | Income quintiles | | | | |
|---|---|---|---|---|---|
| Coefficients | 20% | 40% | 60% | 80% | 100% |
| Permanent income ($\hat{Y}^P_{irt}$) | 0.56*** | 0.56*** | 0.51*** | 0.66*** | 0.64*** |
| Transitory income ($\hat{Y}^T_{irt}$) | 0.57*** | 0.58*** | 0.53*** | 0.67*** | 0.80*** |
| Residual ($\hat{u}^T_{irt}$) | 0.59*** | 0.58*** | 0.59*** | 0.75*** | 0.88*** |
| *Controls* | | | | | |
| Household fixed effects | No | No | No | No | No |
| Year fixed effects | Yes | Yes | Yes | Yes | Yes |
| Household demographics | Yes | Yes | Yes | Yes | Yes |
| Income variation | Yes | Yes | Yes | Yes | Yes |
| *Test statistics* | | | | | |
| $\hat{Y}^P_{irt} = \hat{Y}^T_{irt}$ | 1.50 | 2.59 | 1.38 | 0.09 | 10.13** |
| $\hat{Y}^P_{irt} = 1$ | 553*** | 358*** | 192*** | 75*** | 557*** |
| $\hat{Y}^T_{irt} = 1$ | 349*** | 209*** | 159*** | 44*** | 15*** |
| $\hat{Y}^P_{irt}$ of bottom 20% = top 20% | | | 8.76** | | |
| $\hat{Y}^T_{irt}$ of bottom 20% = top 20% | | | 12.29*** | | |

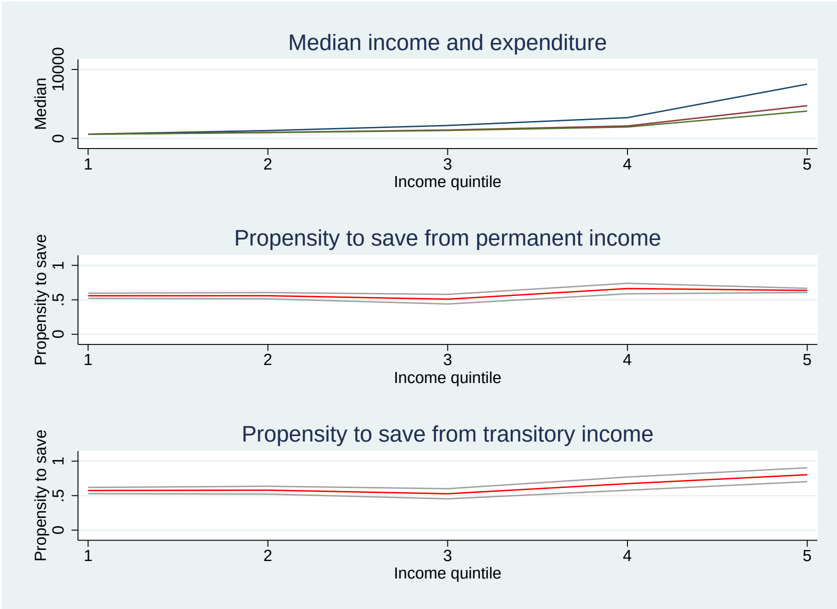
*** $p < 0.001$ ** $p < 0.01$ * $p < 0.05$. Standard error clustered at districts.

## 7. Conclusion

We conduct a decomposition of income into permanent and transitory income, using a newly available, detailed household panel data-set from South Africa. We evaluate the propensity to save on income, using a set of factors to proxy for permanent and transitory income. For permanent income, we use household demographic categories and household asset levels. For transitory income, we use district-level rainfall and temperature data adjusted for agricultural seasons. We include both deviation from long term climate norms, and coefficient of deviation of the weather data. We test whether the propensity to save from permanent income should be close to unity, while the propensity to save from transitory income should be close to zero. We also test the propensity to save by income quintiles in order to identify any differences in saving behavior among income categories.

Following Paxson (1992), we find a significant effect of transitory income as proxied by weather

Figure 2: Propensity to save by quantiles, full data set




Note: Top graph panel represent median income and expenditure (total and non-durable expenditure). Red lines in graph panels below represent median regression estimates. Gray lines represent 95% confidence intervals.

variables, on household saving, in a reduced form fixed effect framework. In the two-step analysis, an OLS framework in the second step yield a propensity to save on permanent income of 0.71, and a propensity to save on transitory income of 0.77. While the propensity to save on permanent income is significantly rejected from unity, the null cannot be rejected for the propensity to save from transitory income. This confirms the predictions of the permanent income hypothesis that propensity to save from transitory income is close to unity. Similar to the finding by Paxson (1992), however, some portion of permanent income is also saved.

We proceed to conduct a series of median regressions, in order to account for potential outliers due to distribution of the income and saving variables. The median propensity to save from permanent income and transitory income are on the order of 0.51 and 0.52. The null hypothesis that either propensities to save is close to one is now strongly rejected.

We conduct a series of analysis on the median propensity to save by income quintiles, in order to examine differences in saving behavior at lower end and higher end of the income distribution. We consider two cases. One is when durable consumption is considered as part of consumption, and thus not included in saving. In this case, we find that the propensity to save is not significantly different between the bottom 20% and top 20% of the income distribution. When we include durable consumption as part of saving, we find that the propensity to save from transitory income significantly differs from permanent income. There is a trend of increasing saving with transitory income, however, it still does not reach one for the highest quintile.

Overall, our findings seem to suggest that households save from both permanent income and transitory income. There is a difference in the propensity to save derived from OLS and median analyses. The distribution of our income and saving data seem to point to median regression as a more appropriate model, where the propensity to save is estimated at around 0.50 from both permanent and transitory income. When we account for the distributional differences in income, we find that durable goods makes a difference in the resulting propensities to save. When durable goods are considered as consumption, there is no difference in propensities to save between the bottom 20% and top 20% of the distribution. However, when durable goods are considered as saving, there is a significant difference between the propensity to save between the bottom 20% and top 20%. In addition, there is a trend of increasing saving on transitory income, although the propensity to save on transitory income does not reach one even in the highest quintile.

## 8. Appendices

### *8.1 Summary figures on weather variables*

Figure 3: Cumulative distribution of temperature (1980-1989 and 2010-2015, unit: deg $K$)

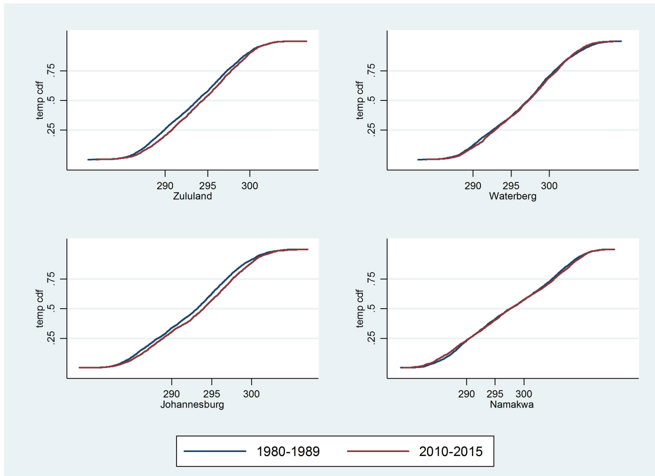
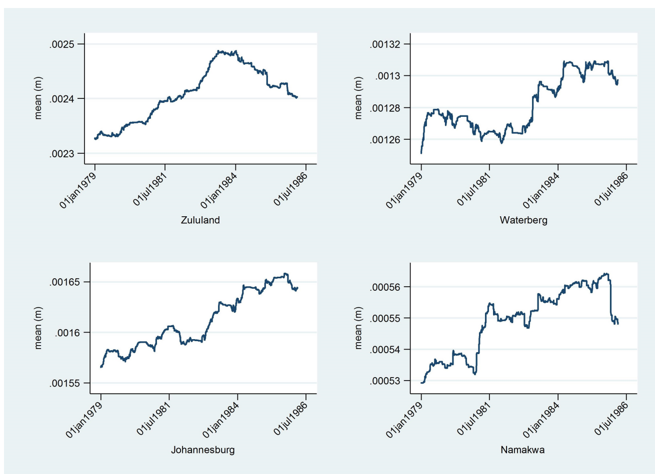



Figure 4: Mean daily precipitation (rolling 30-years, unit: m)




### *8.2 Household panel data collection summary*

In the first stage, 400 primary sampling units (PSUs) were selected from Stats SA's Master Sample of 3,000 PSUs. This master sample is used for its Labour Force Survey and General Household Survey between 2004 and 2007, and for the Income and Expenditure Survey in 2005-2006. The target population includes private households in all nine provinces of South Africa and residents in worker's hostels, convents and monasteries. The frame excluded students' hostels, old age homes, hospitals, prisons and military barracks. The sample was proportionally allocated to strata based on master sample district council PSU allocations, and was not designed to be representative at the province level. Within each PSU, non-overlapping samples (clusters) of dwelling units were systematically drawn. This distribution of PSU clusters for NIDS is comparable to the Master Sample. An initial sample of 9,600 dwelling were drawn. In phase 1 of baseline collection, 6,498 households were successfully interviewed. Since the target response rate of 83% was not achieved, phase 2 data collection was undertaken. A 43% response rate was achieved in phase 2, resulting in an additional 807 successful households. Therefore, a total of 7,305 households were successfully interviewed in 2008 for the baseline, consisting of 28,255 individuals (Leibbrandt et al., 2009). Two sets of weights were calculated for the sample. The design weights were calculated as the inverse of the probability of inclusion, and the post-stratification weights adjust the age-sex-race marginal total in the NIDS data to match the population estimates produced by Stats SA for mid-year population estimates for 2008. Constraints were imposed so that population distribution by province corresponds to population estimates and total weights add up to estimated population of 48,687,000. The in-field call-backs revealed that upper income households were reluctant to participate due to concerns of privacy and questions about the legitimacy of the study. Poorer households were more willing to participate because of availability (e.g., unemployment) and experience with similar previous community studies.

### *8.3 Demographic categories*

| Category | Age | Gender | Education |
|---|---|---|---|
| 1 | less than 5yo | | |
| 2 | 6 to 11 yo | male | |
| 3 | 6 to 11 yo | female | |
| 4 | 12 to 17 yo | male | |
| 5 | 12 to 17 yo | female | |
| 6 | 18 to 64 yo | male | less than 7 years |
| 7 | 18 to 64 yo | male | 7 to 9 years |
| 8 | 18 to 64 yo | male | greater than 9 years |
| 9 | 18 to 64 yo | female | less than 7 years |
| 10 | 18 to 64 yo | female | 7 to 9 years |
| 11 | 18 to 64 yo | female | greater than 9 years |
| 12 | greater than 65 yo | male | |
| 13 | greater than 65 yo | female | |

## References

Auffhammer, M., Hsiang, S., Schlenker, W., & Sobel, A. (2013). Using weather data and climate model output in economic analyses of climate change. *Review of Environmental Economics and Policy*, *7*(2), 181-198.

Bengtsson, N. (2012). The marginal propensity to earn and consume out of unearned income: evidence using an unusually large cash grant reform. *The Scandinavian Journal of Economics*, *114*(4), 1393-1413.

Berg, E. (2013). Are poor people credit-constrained or myopic? Evidence from a South African panel. *Journal of Development Economics*, *101*, 195-205.

Berrisford, P., Dee, D., Poli, P., Brugge, R., Fielding, K., Fuentes, M., ... Simmons, A. (2011). The ERA-interim archive: version 2.0. *ERA Report Series*.

Bhalla, S. S. (1979). Measurement errors and the permanent income hypothesis: evidence from rural India. *American Economic Review*, *69*(3), 295-307.

Bhalla, S. S. (1980). The measurement of permanent income and its application to savings behavior. *Journal of Political Economy*, *88*(4), 722-744.

Bozzola, M., Smale, M., & DiFalco, S. (2016). Climate shocks, weather and maize intensification decisions in rural Kenya. *FOODSECURE working paper no. 39*.

Bryan, E., Deressa, T., Gbetiboubo, G., & Ringler, C. (2009). Adaptation to climate change in Ethiopia and South Africa: options and constraints. *Environmental Science and Policy*, *12*(4), 413-426.

Bunting, D. (1991). Savings and distribution of income. *Journal of Post Keynesian Economics*, *14*, 3-22.

Burgess, R., Deschenes, O., Donaldson, D., & Greenstone, M. (2014). The unequal effects of weather and climate change: evidence from mortality in India. *Unpublished working paper*.

Campbell, J. (1987). Does saving anticipate declining labor income - an alternative test of permanent income hypothesis. *Econometrica*, *55*(6), 1249-1273.

Carroll, C. (1998). Why do the rich save so much? In J. Slemrod (Ed.), *Does Atlas shrug? Economic consequences of taxing the rich.* London: Cambridge University Press.

Carroll, C., & Kimball, M. (1996). On the concavity of the consumption function. *Econometrica*, *64*(4), 981-992.

Chinhema, M., Brophy, T., Brown, M., Leibbrandt, M., Mlatsheni, C., & Woolard, I. (2016). National Income Dynamics Study panel user manual. *NIDS*.

Colmer, J. (2016a). Parental income uncertainty, child labour, and human capital accumulation.

*Under review*.

Colmer, J. (2016b). Weather, labour reallocation and industrial production: evidence from India. *Job market paper*.

Cook, C. (1995). Saving rates and income distribution: further evidence from LDCs. *Applied Economics*, *27*, 71-82.

Deaton, A. (1989). Saving in developing countries: theory and review. *Proceedings of the World Bank annual conference on Development Economics*, 61-96.

Deaton, A. (1991). Saving and liquidity constraints. *Econometrica*, *59*(5), 1221-1248.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., … Vitart, F. (2011). The ERA-interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553-597.

DellaValle, P., & Oguchi, N. (1976). Distribution, the aggregate consumption function, and the level of economic development: some cross-country results. *Journal of Political Economy*, *84*, 1325-1334.

Dynan, K., Skinner, J., & Zeldes, S. (2004). Do the rich save more? *Journal of Political Economy*, *112*(2), 397-444.

Flavin, M. (1981). The adjustment of consumption to changing expectations about future income. *Journal of Political Economy*, *89*(5), 974-1009.

Headey, D., & Fan, S. (2008). Anatomy of a crisis: the causes and consequences of surging food prices. *Agricultural Economics*, *39*(1), 375-391.

Hirvonen, K. (2016). Temperature changes, household consumption and internal migration: evidence from Tanzania. *American Journal of Agricultural Economics*, *98*(4), 1230-1249.

Hsiang, S., Kopp, R., Jina, A., Rising, J., Delgado, M., Mohan, S., … Houser, T. (2017). Estimating economic damage from climate change in the united states. *Science*, *356*.

Japelli, T., Padula, M., & Pistaferri, L. (2008). A direct test of the buffer-stock model of saving. *Journal of the European Economic Association*, *6*(6), 1186-1210.

Jappelli, T., & Pistaferri, L. (2010). The consumption response to income changes. *Annual Review of Economics*, *2*, 479-506.

Kaldor, N. (1957). A model of economic growth. *Economic Journal*, *57*.

Karfakis, P., Lipper, L., & Smulders, M. (2012). The assessment of the socio-economic impacts of climate change at household level and policy implications. *Building resilience for adaptation to climate change in the agricultural sector: Proceedings of a joint FAO/OECD Workshop*, 133-150.

Kazianga, H., & Udry, C. (2006). Consumption smoothing? Livestock, insurance and draught in rural Burkina Faso. *Journal of Development Economics*, *79*(2), 413-446.

Kim, S., Egerter, S., Cubbin, C., Takahashi, E., & Braveman, P. (2007). Potential implications of missing income data in population-based surveys: an example from a postpartum survey in California. *Public Health Reports*, *122*, 753-763.

Knox, J., Hess, T., & Daccache, A. (2012). Climate change impacts on crop productivity in Africa and South Asia. *Environmental Research Letters*, *7*(3), 034032.

Kopczuk, W., & Lupton, J. (2007). To leave or not to leave: the distribution of bequest motives. *The Review of Economic Studies*, *74*(1), 207-235.

Kurukulasuriya, P., & Mendelsohn, R. (2008). A Ricardian analysis of the impact of climate change on African cropland. *The African Journal of Agricultural and Resource Economics: AfJARE*, *2*(1), 1-23.

Leibbrandt, M., Woolard, I., & de Villiers, L. (2009). Methodology: report on NIDS wave 1. *NIDS Technical Paper No. 1*.

Lewis, W. (1954). Economic development with unlimited supplies of labour. *The Manchester School*, *22*(2), 139-191.

Lim, D. (1980). Income distribution, export instability and saving behavior. *Economic development and cultural change*, *26*, 356-364.

Musgrove, P. (1979). Permanent household income and consumption in urban South America. *American Economic Review*, *69*(3), 355-368.

Musgrove, P. (1980). Income distribution and the aggregate consumption function. *Journal of Political Economy*, *88*(3), 504-525.

NIDS. (2008). *National Income Dynamic Study, 2008-2014, Southern Africa Labour and Development Research Unit, Cape Town, South Africa.* Cape Town, South Africa: Southern Africa Labour and Development Research Unit.

Paxson, C. (1992). Using weather variability to estimate the response of saving to transitory income in Thailand. *American Economic Review*, *82*(1), 15-33.

Rosenzweig, M. (2001). Savings behaviour in low-income countries. *Oxford Review of Economic Policy*, *17*(1), 40-54.

Rosenzweig, M., & Wolpin, K. (1993). Credit market constraints and the accumulation of durable production assets in low-income countries: investments in bullocks. *Journal of Political Economy*, *101*(2), 223-244.

Schmidt-Hebbel, K., & Serven, L. (2000). Does income inequality raise aggregate saving? *Journal* *of Development Economics*, *61*, 417-446.

StatsSA. (2014). Consumer price index, 2010, 2012, 2014. *Statistics South Africa*.

Tibesigwa, B., & Visser, M. (2016). Assessing gender inequality in food security among small-holder farm households in urban and rural South Africa. *World Development*, *88*, 33-49.

Tibesigwa, B., Visser, M., & Turpie, J. (2015). The impact of climate change on net revenue and food adequacy of subsistence farming households in South Africa. *Environment and Development Economics*, *20*(3), 327-353.

Ting, L., & Kollamparambil, U. (2015). Nature and determinants of household retirement saving behavior in South Africa. *Development South Africa*, *32*, 675-696.

VAM. (2016). Southern africa growing season, 2015-2016. *WFP VAM Food Security Analysis*.

Venieris, Y., & Gupta, D. (1986). Income distribution and sociopolitical instablity as determinants of savings: a cross-sectional model. *Journal of Political Economy*, *94*, 873-883.

von Braun, J. (2009). Addressing the food crisis: governance, market functing, and investment in public goods. *Food Security*, *1*, 9-15.

Wheeler, T., & von Braun, J. (2013). Climate change impacts on global food security. *Science*, *341*(6145), 508-513.

Wolpin, K. (1982). A new test of the Permanent Income Hypothesis: the impact of weather on income and consumption of farm households in India. *International Economic Review*, *23*(3), 583-594.

Wooldridge, J. (2012). *Introductory Econometrics: A Modern Approach.* Mason, OH: South-Western College Pub.

WorldBank. (n.d.-a). Gross domestic savings (percent of gdp). *The World Bank DataBank*.

WorldBank. (n.d.-b). The world bank in South Africa. *The World Bank*.

Zivin, J., & Neidell, M. (2014). Temperature and the allocation of time: implications for climate change. *Journal of Labor Economics*, *32*(1).

Zwane, T., Greyling, L., & Maleka, M. (2016). The determinants of household savings in South Africa: a panel data approach. *International Business and Economics Research Journal*, *15*(4), 209-218.